

Strength and tempo of directional selection in the wild

H. E. Hoekstra^{*,†,‡}, J. M. Hoekstra^{*}, D. Berrigan^{*§}, S. N. Vignieri^{*†}, A. Hoang^{*}, C. E. Hill^{*¶}, P. Beerli^{||}, and J. G. Kingsolver^{*,**}

^{*}Department of Zoology, [†]Burke Museum, and [‡]Department of Genetics, University of Washington, Seattle, WA 98195

Communicated by Gordon H. Orians, University of Washington, Seattle, WA, June 5, 2001 (received for review November 20, 2000)

Directional selection is a major force driving adaptation and evolutionary change. However, the distribution, strength, and tempo of phenotypic selection acting on quantitative traits in natural populations remain unclear across different study systems. We reviewed the literature (1984–1997) that reported the strength of directional selection as indexed by standardized linear selection gradients (β). We asked how strong are viability and sexual selection, and whether strength of selection is correlated with the time scale over which it was measured. Estimates of the magnitude of directional selection ($|\beta|$) were exponentially distributed, with few estimates greater than 0.50 and most estimates less than 0.15. Sexual selection (measured by mating success) appeared stronger than viability selection (measured by survival). Viability selection that was measured over short periods (days) was typically stronger than selection measured over longer periods (months and years), but the strength of sexual selection did not vary with duration of selection episodes; as a result, sexual selection was stronger than viability selection over longer time scales (months and years), but not over short time scales (days).

natural selection | sexual selection | selection gradients | quantitative traits | natural populations

How strong is selection on quantitative traits in nature? Is strong directional selection common? Is viability selection typically stronger or weaker than sexual selection? Is selection strength correlated with the time period over which it is measured? The answers to these and similar questions are fundamental to understanding how selection determines evolutionary change and adaptation. Whereas theoretical analyses have typically modeled selection as a weak evolutionary process (1, 2), some recent empirical studies demonstrated that strong selection (3, 4) and rapid evolution (5–8) of discrete and quantitative traits may sometimes occur in natural populations. Reconciling these apparent differences through knowledge of the distribution of selection strengths will allow us to more accurately explore both the theoretical and practical consequences of selection.

An early synthesis by Endler (4) revealed two important patterns about the distribution of selection strengths measured by selection differentials. First, selection differentials exhibited a roughly exponential frequency distribution, suggesting that the strength of selection in the wild varied substantially such that strong selection was not necessarily rare. Second, nonmortality and mortality components of selection on quantitative traits yielded similar selection differentials, but nonmortality components yielded higher selection coefficients (S) for discrete polymorphic traits (4). Although provocative, Endler's work used data from just 25 species and, for quantitative traits, relied on selection differentials that did not correct for indirect or correlative selection. Quantitative estimates of the strength of selection using selection gradients, which are standardized and correct for phenotypic correlations among traits (9), have become available only during the past 15 years.

To compile these newer selection measures, we reviewed published studies from 1984 through 1997 that estimated selection on natural variation in quantitative traits within unmanipulated field populations (10). Here we focus on studies that

estimated directional selection as measured by standardized linear selection gradients (β) (9). Selection gradients relate relative fitness to variation in a quantitative trait measured in standard deviation units, and provide a standard metric for comparisons among traits and among systems. In addition, selection gradients are directly relevant to quantitative genetic models for the evolution of quantitative traits (1, 2). In our analyses, we categorized each selection gradient estimate in terms of the type of trait (e.g., morphological, phenological) and the type of fitness component (e.g., survival, mating success) that was measured. We also noted whether each estimate of β was reported to be significantly different from zero at the significance level $P < 0.05$. Finally, we categorized the time scale (days, months, or years) over which β was estimated. Here we use these data to consider two issues: (i) does the strength of selection differ when measured for survival (viability selection) versus mating success (sexual selection)? (ii) Does the strength of selection differ when measured over shorter versus longer time periods?

Methods

We reviewed articles published in 1984 through 1997 in the following journals: *American Journal of Botany*, *American Naturalist*, *Biological Journal of the Linnean Society*, *Ecology*, *Evolution* (Lawrence, Kans.), *Genetics*, *Journal of Evolutionary Biology*, *Philosophical Transactions of the Royal Society of London*, and *Proceedings of the Royal Society of London*. The following journals were reviewed but no appropriate articles were found: *Annals of the Entomological Society of America*, *Ecological Entomology*, *Environmental Entomology*, *Evolutionary Ecology*, *Heredity*, *Nature* (London), and *Science*. Details of the methods and resulting database are described in Kingsolver *et al.* (10). Articles were included if they reported standardized linear selection gradients measured in unmanipulated field populations. Studies were excluded from our database if they manipulated the environment (e.g., in cages, greenhouse, laboratory), population (e.g., with experimental, inbred, or domesticated species), or traits (i.e., phenotypic manipulation); if they used discrete or categorical traits, or nonstandardized selection measures. We also excluded a single study that reported two outlying estimates where the magnitude of directional selection ($|\beta|$) > 5 (11). For each linear selection gradient reported in an article, we categorized the fitness component measured as: survival (S), mating success (M), fecundity (F), total fitness (T) or other (O), and recorded the duration (in days) of the selection episode. We

^{*}Present address: Department of Ecology and Evolutionary Biology, University of Arizona, Tucson, AZ 85721.

[§]Present address: Cancer Prevention Fellowship Program, Rockville, MD 20852-7105.

[¶]Present address: Department of Biology, Coastal Carolina University, Conway, SC 29526.

^{**}To whom reprint requests should be sent at present address: Department of Biology, University of North Carolina, Chapel Hill, NC 27516. E-mail: jgking@bio.unc.edu.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

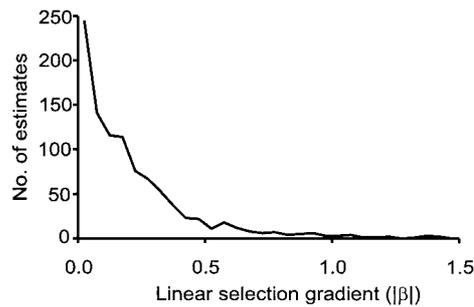


Fig. 1. Frequency distribution of estimated linear selection gradients ($|\beta|$) grouped by 0.05 β units. Selection gradients ($n = 993$) were compiled from 63 studies (median $|\beta| = 0.150$; mean $|\beta| = 0.220$). Two outlying estimates of $|\beta| > 5$ were excluded from analyses. This distribution was similar to an exponential distribution ($P = 0.091$).

then categorized the time scale of each selection episode as day (1–30 days), month (31–364 days) or year (365 + days). The article citations and coded data included in our study can be found in our database at <http://www.bio.unc.edu/faculty/kingsolver/>.

To test our hypotheses regarding the relative strength of viability versus sexual selection and the effect of the duration of selection episodes, we compared the magnitude of β estimates by using appropriate nonparametric statistical tests. The specific tests used for each analysis are identified along with the presentation of results below. Although it was desirable, we did not conduct a formal metaanalysis of the data. Metaanalyses of partial regression coefficients, such as β , require consideration of the entire phenotypic variance–covariance matrix (10, 12, 13). However, only a small minority of studies reported this information.

Results

We identified 63 published studies that reported 993 different estimates of β on quantitative traits in 62 different species of plants (40% of the estimates), invertebrates (31%), and vertebrates (29%) (10). More than 95% of the estimates involved morphological, life history, or phenological traits. The frequency distribution of the magnitude of β declined monotonically as $|\beta|$ increased (Fig. 1). Because the sign of β simply reflected the chosen scale of measurement, we considered all estimates in terms of absolute values, $|\beta|$. The overall distribution of $|\beta|$ had a median of 0.150 and a mean of 0.220, and was similar to an exponential distribution (one-sample Kolmogorov–Smirnov test, $ks = 0.0395$, $P = 0.091$). Despite some reports of very strong selection ($\beta > 1$) (11) and extremely rapid evolution (reviewed in ref. 5), only 10% of the estimates of $|\beta|$ exceeded 0.5 in magnitude. This general frequency distribution pattern in $|\beta|$ was independent of sample sizes used in the studies, although variation in $|\beta|$ decreased with increasing sample size as expected (14, 15). Only 26% of the β estimates were significantly different from zero at the $P < 0.05$ level (as reported in each study), reflecting the low power of most studies to detect “typical” levels of selection (see ref. 10). All estimates of β were used in subsequent analyses, regardless of their significance; only 4% of β estimates did not report a significance value.

To determine the relative strength of viability versus sexual selection, we identified 284 estimates of β based on survival (viability selection) and 394 estimates based on mating success (sexual selection). Measures of fecundity often reflected both mating success and survival, and were therefore not used in this comparison. Estimates of $|\beta|$ for sexual selection were significantly greater than for viability selection (Fig. 2; Wilcoxon rank sum test, $Z = 6.45$, $P < 0.001$). The median $|\beta|$ for sexual

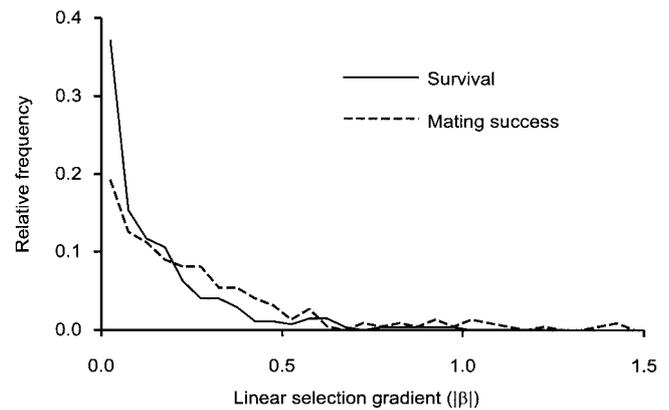


Fig. 2. Frequency distributions showing relative strength of sexual and natural selection. Magnitudes of linear selection gradients were grouped by 0.05 $|\beta|$ units. Frequency distributions (standardized to account for sample size differences between groups) were significantly different ($P < 0.001$). More than 33% of studies that used survival as a fitness component reported $|\beta| \leq 0.05$. Median and mean $|\beta|$ based on measures of survival were 0.153 and 0.088, respectively; median and mean $|\beta|$ based on mating success were 0.250 and 0.180, respectively.

selection was 0.180, and that for viability selection was 0.088. Means were 0.250 and 0.153, respectively. This difference was due primarily to the high frequency of $|\beta|$ values less than 0.1 for viability selection. This pattern occurred within different taxonomic groups (e.g., vertebrates, plants) and for different types of traits (e.g., morphological, life history).

We next examined the effect of the time scale over which selection was measured (episode duration) on the strength of selection, $|\beta|$. This effect was analyzed separately for 284 selection gradient estimates based on survival and 380 estimates based on mating success. Episode duration for each estimate of β was categorized into time scales of days (≤ 30 days), months (31–364 days), or years (≥ 365 days). Among selection gradient estimates based on survival, 30 were measured over days, 55 over months, and 199 over years. Among estimates based on mating success, sample sizes were 119, 25, and 236, respectively. The strength of viability selection (measured by survival) differed significantly among time scale classes (Fig. 3a, Kruskal–Wallis rank sum test, $\chi^2 = 12.68$, $P < 0.002$). Median values of $|\beta|$ were 0.179, 0.112, and 0.070 for episode durations of days, months, and years, respectively. Mean values were 0.222, 0.129, and 0.150, respectively. Pairwise multiple comparisons indicated that the strength of selection measured over days was significantly greater than that measured over years (Tukey-type nonparametric comparison, $Q = 3.42$, $P < 0.002$). This general pattern occurred within different taxonomic groups (e.g., invertebrates, vertebrates) for which generation times differed substantially. In contrast, the strength of sexual selection (measured by mating success) did not differ with the duration of selection episodes (Fig. 3b, Kruskal–Wallis rank sum test, $\chi^2 = 0.52$, $P = 0.77$). Median $|\beta|$ values were 0.174, 0.199, and 0.184, for episode durations of days, months, and years, respectively. Mean values were 0.283, 0.221, and 0.235, respectively.

Inspection of these results suggested that the strength of viability and sexual selection were similar when measured over short time scales but diverged when measured over longer intervals. To test this, we compared estimates of $|\beta|$ based on survival versus mating success for each of the time scale categories. The strength of sexual selection did not differ from that of viability selection when measured over days (Wilcoxon rank sum test, $Z = 0.30$, $P = 0.77$), but sexual selection was significantly stronger when measured over months (Wilcoxon rank sum

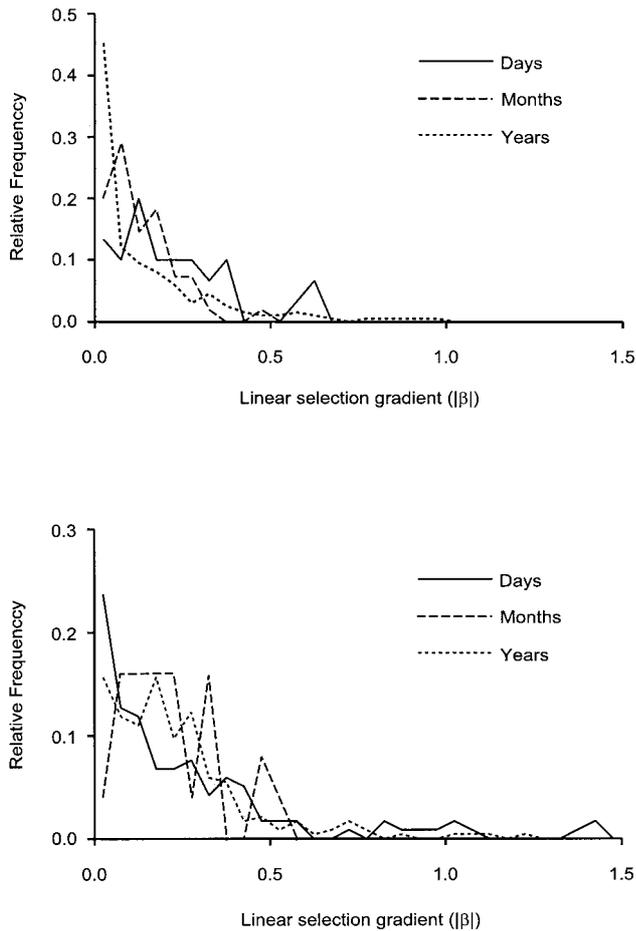


Fig. 3. (a) Frequency distributions of strengths of selection based on survival measured over episode lengths of days (<30 days), months (31–364 days), and years (>365 days). Frequency distributions (standardized to account for sample size differences among groups) of linear selection gradients ($|\beta|$) differed significantly ($P < 0.002$). More than half (54%) of β s measured over a year reported $|\beta| \leq 0.1$, whereas only 36% of β s measured over days reported $|\beta| \leq 0.1$. (b) Frequency distributions of strengths of selection based on mating success measured over the same episode length categories. Frequency distributions (standardized to account for sample size differences among groups) of linear selection gradients ($|\beta|$) were not significantly different ($P > 0.75$). In both figures, selection gradients were grouped by 0.05 $|\beta|$ units.

test, $Z = 3.40$, $P < 0.001$) and years (Wilcoxon rank sum test, $Z = 6.37$, $P < 0.0001$).

Discussion

Three important results emerged from our analyses of the strength of directional selection on quantitative traits in natural populations. First, the frequency distribution of selection gradient estimates was remarkably similar to the distribution of selection differentials reported by Endler (4), even though the estimates in our data have removed the effects of indirect selection. This finding suggests that indirect selection may not play a major role in determining selection strength on a particular trait, as suggested by the analyses of Kingsolver *et al.* (10). This observation also affirmed Endler's conclusion that natural selection in the wild can be quite strong. For example, in a population in which a trait has heritability of 0.5 and lacks genetic correlation to other traits, sustained directional selection

of “typical” magnitude^{††} (median $|\beta| = 0.150$) could change the mean phenotype by one standard deviation in only 16 generations. In fewer than 50 generations, the population mean would exceed the current range of variation of the population. Alternatively, consider that estimates of selection gradients implicitly scale fitness to 1. A $|\beta|$ value of 0.150 suggests that the relative fitness of an individual with a trait value one standard deviation above the mean would be 1.15. Still, the exponential distribution of selection gradients (Fig. 1) indicates that most values of $|\beta|$ cluster near zero (10). The second major result of our analyses was that sexual selection on quantitative traits that influenced mating success was consistently stronger than viability selection estimated by survival. This result was consistent with previously reported patterns of nonmortality versus mortality measures for discrete polymorphic traits (4), and with suggestions from *Drosophila* (16), although not plants (17). This result was also consistent with theoretical models that suggest that sexual selection and mate choice may be important for rapid adaptive divergence (18, 19). Still, a mechanistic explanation for this systematic difference was not immediately apparent, although our third result revealed some potential clues.

The most surprising result from our analyses was that the strength of viability selection varied inversely with the duration of the episode over which it was measured, whereas estimates based on mating success were insensitive to episode duration. The largest estimates of $|\beta|$ based on survival were made when viability selection was measured over short time intervals (days), suggesting that strong directional selection acted only briefly. Over longer time intervals, bursts of strong directional selection may have been tempered by periods of reversal or stasis (5, 20). For example, environmental variation caused selection on Darwin's finches to switch direction over time both within and between generations (e.g., fluctuating selection) such that selection appeared stronger in the short term than in the long (21, 22). In contrast, estimates of $|\beta|$ based on mating success did not vary with episode duration, suggesting that sexual selection is less time-sensitive. Importantly, the median magnitudes of selection by means of survival and mate success were similar for short time intervals (days), and diverged only at larger time scales (Fig. 3).

This latter insight suggests a possible explanation for the greater magnitude of sexual selection versus viability selection. Mating success, although often episodic, is essentially an additive metric that does not necessarily depend on time (i.e., success is usually measured as some discrete number of mates rather than as mates per unit time). Survival, on the other hand, is a multiplicative rate metric that is consequently sensitive to the time interval over which it is measured. For example, assuming type I survivorship, estimates made over longer time intervals will consistently underestimate the instantaneous rate. In addition, mating success depends largely on the social environment, whereas survival is affected by diverse external factors in the environment. If social environments, and thus sexual selection gradients, were more constant through time whereas the strength of viability selection varied stochastically, then one would expect to estimate relatively smaller strengths of viability selection when measured over longer time intervals.

Over the last 15 years, the availability of standardized estimates of the strength of selection on quantitative traits in wild populations has increased tremendously. Our synthesis of these estimates affirmed earlier observations that the strength of

^{††}Published studies included here encompassed a wide range of taxa, habitats, and traits, but probably did not represent an unbiased sample of traits and study systems. Many traits and organisms in these studies were likely chosen with the expectation that selection was probably occurring. In addition, studies that failed to detect significant selection have been less likely to be published, and thus would not be represented in our sample of studies. As a result, our analyses may have slightly overestimated the overall magnitude of selection in nature.

selection varies substantially and that selection in the wild could be quite strong (4, 10). We also found that selection gradient estimates for quantitative traits based on mating success were generally greater than those based on survival. This result suggests that the strength of sexual and viability selection may differ systematically (4, 10). Furthermore, we discovered that the estimated strength of viability selection was greatest when measured over shorter time intervals, but that episode duration did not affect estimated strengths of sexual selection. Consequently, the estimated magnitudes of viability and sexual selection were similar for short selection episodes, but not for longer episodes. This last result adds a curious twist to the observed differences between sexual and viability selection. Elucidating

mechanistic explanations for these patterns should pose an interesting and exciting challenge as evolutionary biologists continue to measure the strength of selection in the wild.

We thank S. A. Combes, G. W. Gilchrist, C. M. Hess, T. R. Hammon, K. M. Kay, J. Ramsey, M. C. Silva, E. K. Steinberg, and K. L. Ward for help with the identification of research articles. D. Schemske provided critical input throughout this project. Helpful discussion was provided by P. Gibert, S. V. Edwards, S. C. Stearns, and R.B. Huey. This manuscript was greatly improved by comments from S. Arnold and two anonymous reviewers. H.E.H. was supported by a Howard Hughes Predoctoral Fellowship. This work was supported in part by the National Science Foundation.

1. Lande, R. & Arnold, S. J. (1983) *Evolution (Lawrence, Kans.)* **37**, 1210–1226.
2. Lande, R. (1979) *Evolution (Lawrence, Kans.)* **33**, 402–416.
3. Boag, P. T. & Grant, P. R. (1981) *Science* **214**, 82–85.
4. Endler, J. A. (1986) *Natural Selection in the Wild* (Princeton Univ. Press, Princeton).
5. Hendry, A. P. & Kinnison, M. T. (1999) *Evolution (Lawrence, Kans.)* **53**, 1637–1653.
6. Losos, J. B., Warheit, K. I. & Schoener, T. W. (1997) *Nature (London)* **387**, 70–73.
7. Reznick, D. N., Shaw, F. H., Rodd, F. H. & Shaw, R. G. (1997) *Science* **275**, 1934–1936.
8. Thompson, J. N. (1998) *Trends Ecol. Evol.* **13**, 329–332.
9. Phillips, P. C. & Arnold, S. J. (1989) *Evolution (Lawrence, Kans.)* **53**, 1506–1515.
10. Kingsolver, J. G., Hoekstra, H. E., Hoekstra, J. M., Berrigan, D., Vignieri, S. N., Hill, C. E., Hoang, A., Gibert, P. & Beerli, P. (2001) *Am. Nat.* **157**, 245–261.
11. Arnold, S. J. & Wade, M. J. (1984) *Evolution (Lawrence, Kans.)* **38**, 72–734.
12. Gurevitch, J. & Hedges, L. V. (1999) *Ecology* **80**, 1142–1149.
13. Hedges, L. V., Gurevitch, J. P. & Curtis, S. (1999) *Ecology* **80**, 1150–1156.
14. Palmer, A. R. (1999) *Am. Nat.* **154**, 220–223.
15. Arnqvist, G. & Wooster, D. (1995) *Trends Ecol. Evol.* **10**, 236–240.
16. Brittnacher, J. G. (1981) *Genetics* **97**, 719–730.
17. Clegg, M. T. & Allard, R. W. (1973) *Science* **181**, 667–668.
18. Lande, R. (1981) *Proc. Natl. Acad. Sci. USA* **78**, 3721–3725.
19. Kirkpatrick, M. (1982) *Evolution (Lawrence, Kans.)* **36**, 1–12.
20. Gingerich, P. D. (1983) *Science* **222**, 159–161.
21. Gibbs, H. L. & Grant, P. R. (1987) *Nature (London)* **327**, 511–513.
22. Grant, P. R. & Grant, B. R. (1995) *Evolution (Lawrence, Kans.)* **49**, 241–251.