Differences in the temporal dynamics of phenotypic selection among fitness components in the wild

Adam M. Siepielski1,*, Joseph D. DiBattista2, Jeffrey A. Evans1 and Stephanie M. Carlson3

1Department of Biological Sciences, Dartmouth College, Hanover, NH 03755, USA
2Hawaii Institute of Marine Biology, University of Hawaii at Manoa, Kane‘ohe, HI 96744, USA
3Department of Environmental Science, Policy, and Management, University of California, Berkeley, CA 94720, USA

The balance of selection acting through different fitness components (e.g. fecundity, mating success, survival) determines the potential tempo and trajectory of adaptive evolution. Yet the extent to which the temporal dynamics of phenotypic selection may vary among fitness components is poorly understood. Here, we compiled a database of 3978 linear selection coefficients from temporally replicated studies of selection in wild populations to address this question. Across studies, we find that multi-year selection through mating success and fecundity is stronger than selection through survival, but varies less in direction. We also report that selection through mating success varies more in long-term average strength than selection through either survival or fecundity. The consistency in direction and stronger long-term average strength of selection through mating success and fecundity suggests that selection through these fitness components should cause more persistent directional evolution relative to selection through survival. Similar patterns were apparent for the subset of studies that evaluated the temporal dynamics of selection on traits simultaneously using several different fitness components, but few such studies exist. Taken together, these results reveal key differences in the temporal dynamics of selection acting through different fitness components, but they also reveal important limitations in our understanding of how selection drives adaptive evolution.

Keywords: adaptation; fecundity; mating success; natural selection; sexual selection; survival

1. INTRODUCTION

Selection is a driving evolutionary force, having a central role in shaping much of the phenotypic and ultimately species diversity observed in nature [1–4]. Studies of selection often distinguish between ‘natural’ and ‘sexual’ selection because of the focus on different fitness components—natural selection often acting through survival or fecundity and sexual selection acting through mating success (e.g. via male–male competition or female choice [5–7]). Darwin [5] was the first to suggest that sexual selection was distinct from natural selection when attempting to explain the evolution of exaggerated secondary sexual characters, such as the archetypal peacock’s tail, which seemed at odds with the theory of natural selection because such traits would presumably carry survival or other fitness costs.

Selection also exhibits considerable temporal variation in nature owing, in large part, to temporal variation in environmental conditions [6,7]. Although temporal variation in natural selection has long been appreciated [8], studies of sexual selection have also increasingly shown it to vary temporally and, much like natural selection, to be driven by variation in environmental conditions [9]. For example, temporal variation in sexual selection has been attributed to variation in ecological factors such as density [10–12], climate [13] and sex ratio [10,12], and many of these same factors may also impact temporal variation in selection through survival or fecundity [6,7]. In fact, the dynamics of sexual selection are especially prone to temporal environmental fluctuations because of the often environmentally context-dependent nature of female preferences [14–16], and environmentally dependent expression of many sexually selected traits [17,18].

While researchers may focus on specific fitness components in studies of adaptive evolutionary change, such evolutionary responses are due to the combined effect of selection on multiple fitness components. Indeed, the relative strength of selection acting among different fitness components should ultimately determine the potential tempo and direction of adaptive evolution [5,9,10,15–25]. Consequently, understanding differences in the temporal dynamics of selection among different fitness components is highly relevant for studies of adaptive evolutionary change. Yet the extent to which the temporal dynamics of phenotypic selection vary among fitness components is largely unexplored.

To illuminate patterns of temporal variation in selection among fitness components, we gathered temporally...
replicated (2 or more years) estimates of selection coefficients on quantitative traits from wild, unmanipulated populations published in the primary literature between 1983 and 2008 [6]. We only included studies estimating standardized selection coefficients [26], which relate relative fitness to a quantitative trait measured in standard deviation units to facilitate comparisons among taxa, traits and fitness components. The resulting database contained a total of 5519 estimates of selection, and we focus primarily on the subset of this database that are linear (directional) coefficients, including 3978 linear gradients (corrected for correlations among traits) or differentials (uncorrected for correlations among traits). We categorized each of these coefficients with regard to fitness component (fecundity, mating success or survival), and then used these data to address the following questions. (i) How does the multi-year strength of selection on traits vary among fitness components? (ii) Does temporal variation in selection (including overall variation as well as variation in strength and direction) on traits differ among fitness components?

2. MATERIAL AND METHODS
(a) Data
We reviewed articles published between 1983 and 2008 in an attempt to conduct an exhaustive review of temporally replicated estimates of selection. We included studies that (i) calculated standardized selection differentials and gradients (sensu [26]), (ii) considered quantitative traits showing continuous variation, (iii) focused on wild, unmanipulated populations, and (iv) estimated selection in at least 2 years. Selection differentials reveal total (indirect plus direct) selection on a trait, whereas selection gradients reveal direct selection on a trait after taking into account underlying phenotypic correlations among traits included in the regression models. Complete details on our methods for identifying pertinent papers and the resulting database are described in [6].

We reviewed a total of 1569 studies. Of these, 89 studies met the above criteria and were included in the database, yielding 3978 estimates of linear selection [6]. The database is biased in favour of vertebrates (especially birds) and morphological traits. The number of temporal replicates ranged between 2 and 45 years, with a mean of 7.6 years among studies. We have posted the database to the DRYAD website (http://www.datadryad.org/repo/). Upon entering the data into the database, we first used the authors’ original assignments of fitness components but also checked for consistency in these assignments among authors. When discrepancies emerged, which were rare, we (A.M.S., J.D.D. and S.M.C.) made a collective decision as to how to assign the fitness component. Although most measures of fitness components were fairly standard among authors, fecundity was represented by several measures (often varying among taxa), but most measures reflected some value related to the number of offspring or potential offspring produced (e.g., number of eggs produced in the case of birds or fish, or number of seeds or fruits in the case of plants). As a result, some of these values incorporate elements of mating success while other values may be closer to a metric of reproductive effort independent of mating success.

Although we focus our analysis on linear components of selection, we also used estimates of nonlinear selection ($\gamma$) to parametrize a simple quantitative genetics model in the interest of estimating how far traits responding to variation in mating success, fecundity and survival are from phenotypic optima (electronic supplementary material). Recently, Stinchcombe et al. [27] revealed that estimates of $\gamma$ may frequently be underestimated by one-half. However, we have no reason to suspect a systematic tendency for researchers to have underestimated $\gamma$ for one fitness component more than the others. Thus, although the absolute values of the distance to optima may be affected (overestimated), the relative differences among fitness components should not.

(b) Statistical analyses
We quantified four components of temporal variation in selection on traits. First, we considered how the multi-year strength of selection (AVG) varied in relation to fitness components by calculating the mean of the absolute values of the annual selection coefficients across all years. Second, we used the standard deviation (s.d.) of the raw selection coefficients across all years to provide an overall measure of temporal variation in selection on traits. This measure reflects variation in selection on traits owing to both changes in the strength and direction of selection. The standard deviation is an ideal measure because its properties are well known, it describes the distribution of selection coefficients and it describes variation in the original variable’s unit of measure (selection coefficients). Third, to isolate temporal variation in the strength of selection, we also calculated the standard deviation of the absolute values of the annual selection coefficient (s.d.abs). Finally, to quantify temporal variation in the direction of selection, we calculated the frequency of changes in the direction of selection. We estimated this value as the number of changes in the direction of selection between successive temporal intervals relative to $n - 1$, where $n$ is the total number of years of the study and $n - 1$ is the total number of potential changes in the direction of selection. In addition, we present frequency histograms for each of these different components of temporal variation to visually assess differences among fitness components.

We used linear mixed models (LMMs) to investigate temporal variation in selection on traits among fitness components (factor in the models) across studies to draw generalities regarding the temporal dynamics of selection in nature. We used the AVG, s.d., s.d.abs and frequency of changes in the direction of selection as the response variables in the models. LMMs were appropriate because multiple measures of fitness, or selection on several traits, were often reported for a given study and so the data were not independent; the LMM approach allowed us to take into account such underlying covariances. We included trait and fitness measure (both nested within species) as random effects, with species treated as a repeated factor and the number of years in the study as a covariate. This model structure also helped to guard against the fact that we often had multiple traits for some species and only one trait for others, and occasionally multiple estimates of fitness for the same trait within a species (e.g. [28]). If multiple populations were present in a study, each was designated as its own independent sample. Because variation in selection on traits is composed of two parts—‘real’ variation among years and variation imposed by random sampling error of the selection coefficients [6]—we included mean sample size (across years) as a weighting factor because variation imposed by sampling error is expected to be greater when small sample sizes are
used to estimate selection coefficients \([6]\). Data were log\(_{10}\)-transformed prior to analysis because distributions of the above measures of temporal variation in selection were right-skewed. Throughout, we present back-transformed ls means.

We also investigated temporal variation in the dynamics of selection within studies by focusing on the small number of studies \((n = 11)\) that estimated phenotypic selection on the same trait using more than one fitness component. This allowed us to examine whether the patterns we observed across studies were also reflected within studies. Because our measures of temporal variation in selection were based on values averaged over the entire study (i.e. leaving no error d.f. for statistical comparisons), we did not perform formal statistical analyses here. Instead, we used a simple graphical approach by creating scatter plots of the AVG, s.d. and the frequency of changes in the direction of selection for each trait among the different fitness components.

**3. RESULTS**

Significant differences existed in the multi-year strength of selection among fitness components for both linear gradients \((F_{2,152} = 7.96, p = 0.0005)\) and linear differentials \((F_{2,108} = 4.91, p = 0.009)\), with the general patterns being the same (figures 1 and 2). For linear gradients, directional selection was roughly twice as strong on mating success \((t_{152} = 2.94, p = 0.004)\) and fecundity \((t_{152} = 3.67, p = 0.0003)\) when compared with survivorship. Similarly, for linear differentials, selection was about twice as strong on fecundity as on survivorship \((t_{108} = 3.02, p = 0.003)\), and
marginally stronger on mating success than on survival (figure 1b; $t_{108} = 1.84$, $p = 0.068$). No significant differences existed between fecundity and mating success for gradients ($t_{152} = 0.29$, $p = 0.769$) or differentials ($t_{108} = 1.20$, $p = 0.231$). Comparisons of the frequency distributions of the strength of selection reveal similar patterns (figure 2a,b), but also show that mating success exhibits bouts of very strong selection.

Overall temporal variation in selection, as quantified by the standard deviation of selection coefficients among years, also differed among fitness components for linear gradients ($F_{2,152} = 8.35$, $p = 0.004$) but not linear differentials ($F_{2,108} = 0.21$, $p = 0.808$; figures 1c and 2). For gradients, the standard deviation among years was roughly twice as large for mating success relative to either survival ($t_{152} = 3.80$, $p = 0.0002$) or fecundity ($t_{152} = 3.50$, $p = 0.0006$; figure 1c). No significance was found when comparing the standard deviation between survival and fecundity (figure 1c; $t_{152} = 0.05$, $p = 0.963$). Comparing the patterns for gradients (figure 1c) with differentials (figure 1d) suggests that direct selection on traits (gradients) is more temporally variable than direct and indirect selection combined (differentials), at least for mating success. Nearly identical results were found when we used the standard deviation of the absolute value of the selection coefficients among years, which is a measure of the frequency of changes in direction (see figure 2e,f).
figure S1), which isolated variation in the strength of selection.

Finally, significant differences existed in the frequency of changes in the direction of selection among fitness components for both linear gradients ($F_{2,152}^2 = 7.93, p = 0.02$) and linear differentials ($F_{2,108}^2 = 7.03, p = 0.001$), and the general patterns were similar (figures 1 and 2). For linear gradients, changes in the direction of selection were marginally more common for survivorship relative to either mating success (figure 1e; $t_{152} = 1.77, p = 0.07$) or fecundity (figure 1e; $t_{152} = 1.70, p = 0.0001$). Similarly, for linear differentials, changes in the direction of selection were about twice as common for survivorship relative to either mating success (figure 1f; $t_{108} = 3.01, p = 0.003$) or fecundity (figure 1f; $t_{108} = 3.25, p = 0.001$). Changes in the direction of selection did not differ significantly between mating success and fecundity for either gradients ($t_{152} = 1.70, p = 0.091$) or differentials ($t_{108} = 0.47, p = 0.639$).

Comparisons of the frequency distributions of changes in the direction of selection reveal similar patterns (figure 2). For simplicity, and because analyses of gradients and differentials generated largely similar...
results, we do not distinguish between these here. Mating success was consistently stronger than survival selection (figure 3a) and tended to be stronger than fecundity (figure 3d), while fecundity selection tended to be slightly stronger than survival selection (figure 3g). Similarly, mating success was much more variable overall relative to survival (figure 3b), but comparable to fecundity (figure 3e); fecundity and survival selection were similarly variable (figure 3h). Unlike the overall comparisons (figure 1), mating success and survival were equally likely to exhibit changes in the direction of selection (figure 3c), whereas fecundity selection was more likely to change direction than either mating success (figure 3f) or survival (figure 3i).

4. DISCUSSION

Our study reveals key differences in the temporal dynamics of selection among fitness components (figures 1 and 2). We found that selection through mating success and fecundity are comparable in strength, but stronger than survival selection. Selection through mating success was more variable overall among years (at least for selection gradients), particularly in strength (electronic supplementary material, figure S1), than either fecundity or survival. Selection through survival was more variable in direction than mating success and fecundity, which were similar. These major features of the temporal dynamics of selection among fitness components across studies are quite similar to the patterns observed within the subset of studies that quantified selection on the same traits using multiple fitness components (figure 3).

The finding that selection through mating success and fecundity is stronger than selection through survival is consistent with earlier work that compared the strength of selection for studies spanning mostly a single year [20,29,30]. Thus, even when considering temporally replicated estimates of selection, selection through survival is consistently weaker on average than selection through either mating success or fecundity (figures 1 and 2). Why might this be? Estes & Arnold [31] suggested one possibility: weak selection might be expected if populations are being maintained near a long-term equilibrium by small changes in the direction of selection. In support of this idea, we found that selection through survival is roughly half as strong as selection through either mating success or fecundity, but that changes in the direction of survival selection are nearly twice as common as changes in the direction of selection through mating success and fecundity (figure 1). This pattern of fluctuating selection is probably a consequence of inter-annual variation in environmental conditions, which impose selection for different phenotypes in different years [6,7]. These environmental conditions encompass a variety of abiotic and biotic factors, but are often linked with climatic changes. In Darwin’s finches, change in the direction of survival selection on beak size is associated with annual variation in rainfall and the resulting size of seeds that are available to the birds [8]. In contrast, survival selection on body size of salmon attempting to reach their stream breeding grounds varies in strength, but not direction, owing to the water level at the mouth of the creek, which differs among years owing to annual variation in temperature and rainfall [32]. Long-term observational studies such as these provide us with an opportunity to understand what might be the causal agents of selection [33]. However, experimental studies are ultimately needed to demonstrate causality (e.g. [34]). For example, experimental studies have shown that temporal variation in the direction of survival selection on limb size in Anoles lizards can be attributed to altered habitat use imposed by predators [35].

An alternative explanation for weaker survival selection relative to fecundity and mating success may be due, in part, to the time interval over which these fitness components are quantified. For example, Hoekstra et al. [20] found that over short time intervals (e.g. days) the strength of selection through survival was comparable to the strength of selection through mating success, but over long time scales (e.g. years) sexual selection was stronger. In contrast, one of the few studies that quantified selection through survival and mating success over the same interval found that selection through mating success is stronger than survival selection (e.g. [36]), which suggests this observation is not simply a sampling artefact. As noted by Kingsolver et al. [30], stronger apparent selection through variation in mating success could also be due to choice of traits to study. Investigators studying sexual selection may often choose specific traits they think a priori are under strong selection (i.e. ornamental traits in males), whereas studies of survival or fecundity selection may choose to include a number of traits because it is perhaps less clear which traits may influence these fitness components. Additionally, estimates of selection among these fitness components may be biased statistically. Kingsolver et al. [30] noted that mating success (and fecundity) is often based on discrete counts, whereas survival selection is based on binary data, which may overestimate the former, but should not consistently bias the latter towards smaller values when a linear regression is applied to estimate selection coefficients.

The greater estimated frequency of changes in the direction of selection for survival relative to the other fitness components could be due in part to weaker selection through survival (figures 1 and 2). With sampling error of the selection coefficients, weaker selection (e.g. coefficients closer to zero) could result in more frequent inferred shifts between positive and negative values. In our empirical dataset, we did find evidence of negative correlations between the frequency of changes in the direction of selection and the strength of selection among fitness components. However, these correlations are rather weak, with the average correlation $r = -0.210$ for both gradients and differentials (electronic supplementary material, figure S2). To try and gauge how sampling error alone may produce inferred changes in the direction of selection, we also conducted a simple simulation study. We created simulated datasets in which we fixed the value of the selection coefficient, and then sampled from the distribution around that coefficient; the variation across a range of standard errors (electronic supplementary material, Methods). We then calculated inferred changes in the direction of selection that would be due to sampling error alone. Overall, this exercise revealed that inferred changes in the direction of selection tend to increase as the standard error increases relative to
derive an equation that could be used to estimate how far a population was from its phenotypic optima for a given quantitative trait (eqn 7 in [31]). Under the assumption that a quadratic function approximates the individual fitness surface [31,39], the absolute value of the ratio of directional selection gradients to their respective quadratic gradients ($|\beta|/|\gamma|$ where $\beta$ are the linear selection gradients and $\gamma$ the quadratic selection gradients) gives an estimate of how far a population trait mean is from its optimum in phenotypic standard deviations [39]. Estes & Arnold [39] provide a lucid description of the theoretical basis for this estimate, and we refer the reader to that paper for details. Using our database to estimate this ratio for selection coefficients via mating success, fecundity and survival selection revealed that, consistent with the above prediction, traits experiencing selection through mating success (14.59 phenotypic s.d. $\pm$ 9.14(s.e.)) and fecundity (11.62 $\pm$ 5.48(s.e.)) are further from their phenotypic optima than traits experiencing selection through survival (2.28 $\pm$ 0.54(s.e.)), which would favour more consistent (i.e. fewer reversals) directional sexual and fecundity selection such as that reported here (figures 1 and 2).

While our analysis suggests important differences in the temporal dynamics of selection among different fitness components, this analysis also reiterates an important limitation in our understanding of selection—namely that very few studies have quantified selection on the same trait using more than one fitness component (e.g. figure 3 [19,41]). Selection is a complex process, and can act on multiple traits and multiple fitness components sequentially or concurrently. Sexual organisms have to survive long enough to choose (or be chosen by) a mate, to mate, and to then produce offspring. Conflict in selection among different fitness components for a given trait certainly exists [19], and different traits will often, but not always, mediate these different fitness components. It is ultimately the integration of all of these fitness components, however, that matters [5,9,10,15–25]. Indeed, numerous authors have grappled with the problem of combining estimates of multiple bouts of selection through different fitness components (e.g. [42–44]; reviewed in [45]). Most recently, Shaw et al. [44] introduced aster modelling, which allows one to combine estimates of selection acting through different fitness components in a unifying statistical framework to gain insight into overall fitness. Our understanding of adaptive evolutionary change would also be greatly advanced through more temporally replicated studies that incorporate experimental manipulations [35], studies that quantify selection on the same trait using multiple fitness components, studies that control for environmental biases [46], studies that identify the genetic underpinnings of traits [47–50] and studies with sufficient statistical power to detect significant but weak selection [51].

In sum, our analysis has identified differences in the temporal dynamics of selection among fitness components in the wild. These differences have clear implications for models of evolutionary change that focus on traits evolving through selection on different fitness components. Of course, the extent to which evolutionary change may occur in natural systems is potentially limited by a number of factors. For example, selection through different fitness components is often
in conflict (e.g. [5,9,10,15–25]), but can work in unison [4,52], and strong genetic covariances among traits may prevent their independent evolution [53], among other factors limiting evolutionary change [54]. Unfortunately, very few studies have quantified the extent of evolutionary change in traits under sexual selection relative to natural selection within a single system (but see [23]). Ultimately, however, the balance between selection acting through multiple fitness components—which should itself vary with time—will determine the temporal dynamics of adaptive trait change in populations.

We thank C. Benken, R. Cal skeeb, R. Cox, S. Kuchta, D. McDonald, T. Parchman, M. McPeek, L. Rowe, L. Symes and several anonymous reviewers for providing many useful comments on earlier versions of this paper. S. Arnold, A. Hendry, J. Kingsolver, R. Lande, D. Schemske and D. Schluter deserve special thanks for many fine suggestions and insightful comments on earlier portions of this work. A. Ragavendran helped with the statistical models. We are also extremely grateful to several authors who provided us with datasets, and in some cases re-analysed data for us. A.M.S. was supported by the NSF (DBI-0630626, awarded to R. Sloboda) and Dartmouth College (R. Irwin). S.M.C. was supported by the Howard Hughes Medical Institute (awarded to R. Sloboda) and Dartmouth College (R. Irwin). S.M.C. was supported by the NSF (DBI-0630626, DEB-0911573).

REFERENCES


34 Calsbeek, R. & Cox, R. M. Experimentally assessing the relative importance of predation and competition as agents of selection. *Nature* 465, 613–616. (doi:10.1038/nature09020)


