SHORT COMMUNICATION

Interspecific patterns of phenotypic selection do not predict intraspecific patterns

J. R. OTT* & S. P. EGAN†‡

*Department of Biology, Population and Conservation Biology Program, Texas State University, San Marcos, TX, USA
†Department of Biological Sciences, University of Notre Dame, Notre Dame, IN, USA
‡Department of Ecology and Evolutionary Biology, Rice University, Houston, TX, USA

Abstract

We estimated linear (β) and nonlinear (γ) selection gradients to quantify host plant-mediated selection on the trait gall size in each of 22 unequally sampled subpopulations of the cynipid gall wasp *Belonocnema treatae*. We characterized the relationship between variation in subpopulation sample size and the magnitude of and the variance among selection gradients. We then tested the hypothesis that the intraspecific patterns we observed would follow two patterns that have emerged from published estimates of linear and nonlinear selection gradients compiled across species, namely that the average magnitude of β and γ and the variance among estimated β and γ decrease with increasing sample size. For both β and γ, intraspecific patterns of phenotypic selection in relation to sample size were not predicted by interspecific patterns. Thus, our results suggest that when selection is heterogeneous among subpopulations, variation in the biological basis for selection is more important in influencing estimates of selection than is variation in study size. Our study highlights the value of inspecting selection in relation to sampling effort at the level at which understanding the sources of variation in selection is most important, among populations within species.

Introduction

By compiling studies of selection conducted in natural populations, Kingsolver *et al.* (2001) assessed the magnitude and statistical significance of both linear and nonlinear selection, estimated by selection differentials and selection gradients (β and γ) (Lande & Arnold, 1983), for a variety of traits and fitness components. The review produced a number of patterns that have since influenced perceptions of the magnitude of selection in natural populations (Conner, 2001; Hereford *et al.*, 2004; Hendry, 2005; Kingsolver & Pfennig, 2007). Two patterns in particular emerged from inspection of estimates of β and γ in relation to sample size: (i) the variance among estimates of both β and γ decreased with increasing sample size, particularly for $Ns > 1000$ and (ii) the average magnitude of both β and γ

decreased with increasing sample size. Importantly, the two patterns are linked as the first, viewed through the lens of sampling error in conjunction with publication bias (Rosenthal, 1979), provides a hypothesis to explain the convergence of β and γ on zero as sample size increases.

The among-species review conducted by Kingsolver *et al.* (2001) inspired our research question: does the negative association between sample size and (i) the magnitude and (ii) the variance in selection found in interspecific comparisons predict intraspecific patterns in relation to sample size? Thus, we treat the patterns revealed by the analysis of Kingsolver *et al.* (2001) as hypotheses to be evaluated within species. Specifically, we estimated selection on a single trait, gall size, in each of 22 unequally sampled subpopulations of the gall-forming insect *Belonocnema treatae* Mayr, 1881 (Hymenoptera: Cynipidae) to test the predictions that both the magnitude of and variance among estimates of β and γ decrease as study sample size increases. Such a design excludes publication bias as a source of variation

Correspondence: James R. Ott, Biology Department, Texas State University, 601 University Drive, San Marcos, TX 78666, USA.
Tel.: +512 245 2321; fax: 512 245 8713; e-mail: jimott@txstate.edu

among selection gradients while incorporating two important and co-occurring sources of variation present in studies of natural populations: biological variation due to realized differences in the selection process among populations and sampling error due to unequal sampling of populations. To the best of our knowledge, this approach has not been explored to date. We test each hypothesis using the published data set presented in Egan *et al.* (2011) (Dryad doi: 10.5061/dryad.1js1n). The results of our analyses demonstrate that the magnitude of selection and the variance among estimates of selection are independent of variation in sample size and highlight the value of studies aimed at assessing both the roles of sampling variation and process variation in generating patterns of selection among populations within species.

Materials and methods

Study system and data sets

Egan *et al.* (2011) provided estimates of β and γ for each of 22 subpopulations in a study system that circumvents many problems associated with detecting and measuring selection. A single trait closely tied to fitness, the size of the gall produced by *B. treatae* on the leaves of plateau live oak, *Quercus fusiformis* (Muller, 1961), was studied. The fitness for each gall former was unambiguously assessed (binomial outcome survive = 1, die = 0) as a function of gall size for samples of gallers that developed on each of 22 naturally occurring plateau live oak trees at Texas State University's Freeman Ranch in Hays County, Texas, USA (29°55′ N, 98° 00′ W). Selection was analysed at the level of the subpopulation (tree) because (i) gall formers exhibit local adaptation to individual trees (Egan & Ott, 2007) (ii) both mean survival and mean gall size differ among trees (Egan & Ott, 2007; Hood & Ott, 2010) and (iii) the fitness function describing the relationship between gall size and percentage survival varies among gall-former subpopulations (Egan *et al.*, 2011). Selection was measured for gall formers that developed within enclosures that protected gallers from natural enemy attack. Thus, the estimated selection gradients represent mortality selection arising from the intimate interaction of the galler and the host plant. This host plant-mediated selection is distinct from selection on gall size that can be imposed when natural enemies are present (Reynolds, 2000; Stone & Schönrogge, 2003). Both β and γ (and corresponding standard errors, SEs) were estimated for each of the 22 subpopulations. Sample sizes ranged from 39 to 2090, with 11 subpopulations having Ns > 500 and 6 subpopulations with Ns > 1000. The six subpopulations with the largest sample sizes substantially add to the number of studies reported by Kingsolver *et al.* (2001) with an $N > 1000$ in which stabilizing selection has

been studied. Details of the design of the studies producing the survival data and estimation of β and γ, as well as of the spatial distribution of gall-former subpopulations, are given in Egan & Ott (2007) and Egan *et al.* (2011).

Tests of hypotheses

Estimates of β and γ were used to test a series of hypotheses of the distributional properties of β and γ and hypotheses of the magnitude and variation among estimates in relation to sample size. We first tested three hypotheses of the distribution of β and γ that emerged from the among-species review of Kingsolver *et al.* (2001): (i) the distribution of γ is centred on zero (i.e. stabilizing and disruptive selection are equally prevalent) (ii) observed values of β and γ indicate selection of relatively low magnitude and (iii) β and γ are distributed exponentially (i.e. a skewed density distribution reflecting the prevalence of low values). To determine whether the distribution of γ was centred on zero, we tested the hypothesis that mean $\gamma = 0$ using a *t*-test. To assess the shape of the distribution of β and γ, we tested the observed distributions of $|\beta|$ and $|\gamma|$ for fit both to the negative exponential distribution by means of a Kolmogorov–Smirnov test and to the half-normal distribution by means of the Shapiro–Wilks test.

We next used the variation in the number of gall formers monitored in each of the 22 subpopulations and the observed values of $|\beta|$ and $|\gamma|$ to examine the relationship between sample size and (i) the incidence of selection (i.e. the probability of detecting statistically significant selection) (ii) the magnitude of selection and (iii) the variance among estimates of β and γ. To assess the incidence of selection, we scored linear and nonlinear selection as present or absent per subpopulation based on the log-likelihood ratio tests of significance presented in Table 3 of Egan *et al.* (2011). We then used this binomial response variable as the dependent variable in a logistic regression with sample size as the independent variable. Linear regression was then used to determine whether the magnitudes of $|\beta|$ and $|\gamma|$ were dependent on sample size per subpopulation. Finally, to test the hypothesis that variation among estimated selection gradients decreases with increasing sample size, we binned the 22 subpopulations into three $\log_{10}$ categories ($N = 0$–99, $N = 100$–999 and $N \geq 1000$) following Kingsolver *et al.* (2001) and compared the variance among estimates of β and γ, respectively, among sample size categories by means of Levene's test for homogeneity of variances. To determine whether the outcome of the above test was robust to the manner of grouping, we also grouped β and γ according to four natural breaks in sample size observed among the surveyed subpopulations ($N \leq 75$, $N = 138$–192, $N = 469$–808 or $N \geq 987$) and repeated the analysis.

Results

Distribution and magnitude of selection gradients

Across all subpopulations, the median observed values of $|\beta|$ and $|\gamma|$ were 0.06 and 0.29, respectively. However, the distribution and magnitude of β and γ characterizing selection on gall size varied widely among subpopulations (Fig. 1a,b). The 22 estimates of γ were not centred on zero ($t_{21} = 6.18$, $P < 0.000$; Fig. 1b); instead, all but two estimates of γ were negative, including all 17 statistically significant estimates, which indicates a preponderance of stabilizing selection. The hypothesis that $|\beta|$ was distributed normally was rejected ($W = 0.81$, $P < 0.005$, $N = 22$); however, the hypothesis that $|\beta|$ was distributed as a negative exponential could not be rejected ($D = 0.15$, $P > 0.15$, $N = 22$). Similarly, the distribution of $|\gamma|$ did not fit a half-normal distribution ($W = 0.87$, $P < 0.006$, $N = 22$), but the hypothesis that $|\gamma|$ was distributed as a negative exponential could not be rejected ($D = 0.21$, $P > 0.09$, $N = 22$). Overall, the range of variation among estimates of selection in this single study system span a significant portion of the range reported in the among-species comparison.

Patterns of selection in relation to sample size

The incidence of linear selection was not simply a function of sample size ($\chi_1^2 = 1.01$, $P = 0.3155$). Instead, all

estimates of $|\beta| > 0.12$, regardless of sample size, were statistically significant, whereas only 2 of 14 $|\beta|$ estimates < 0.10 were found to be significant (Fig. 1a). These latter two estimates, with Ns of 808 and 1315, respectively, were characterized by smaller SEs but not higher $|\beta|$s relative to the 12 other nonsignificant estimates. The hypothesis that the magnitude of $|\beta|$ should decrease with increasing sample size was not supported as linear regression demonstrated that $|\beta|$ was independent of sample size ($F_{1,21} = 0.18$, $P = 0.6711$; Fig. 1a). For nonlinear selection, in contrast to the prediction that the incidence of selection should decrease with increasing sample size, the probability of detecting statistically significant selection increased with sample size ($\chi_1^2 = 9.08$, $P = 0.0026$). Also in contrast to our prediction that $|\gamma|$ decreases with sample size, linear regression showed that $|\gamma|$ actually increased with increasing sample size ($F_{1,21} = 6.26$, $P = 0.021$; Fig. 1b). The negative relationship between sample size per subpopulation and the SE of β and γ expected from sampling theory are evident in Fig. 1a,b; however, for both β and γ, we found no support for our hypothesis that variation among estimated selection gradients decreases with increasing sample size as variation among estimates of both β and γ did not differ among the three $\log_{10}$ sample size categories (β: $F_{2,19} = 0.26$, $P = 0.77$; γ: $F_{2,19} = 1.58$, $P = 0.23$; Fig. 1a,b). This result was robust to the manner of grouping (β: $F_{3,18} = 1.38$, $P = 0.28$; γ: $F_{3,18} = 2.27$, $P = 0.11$).
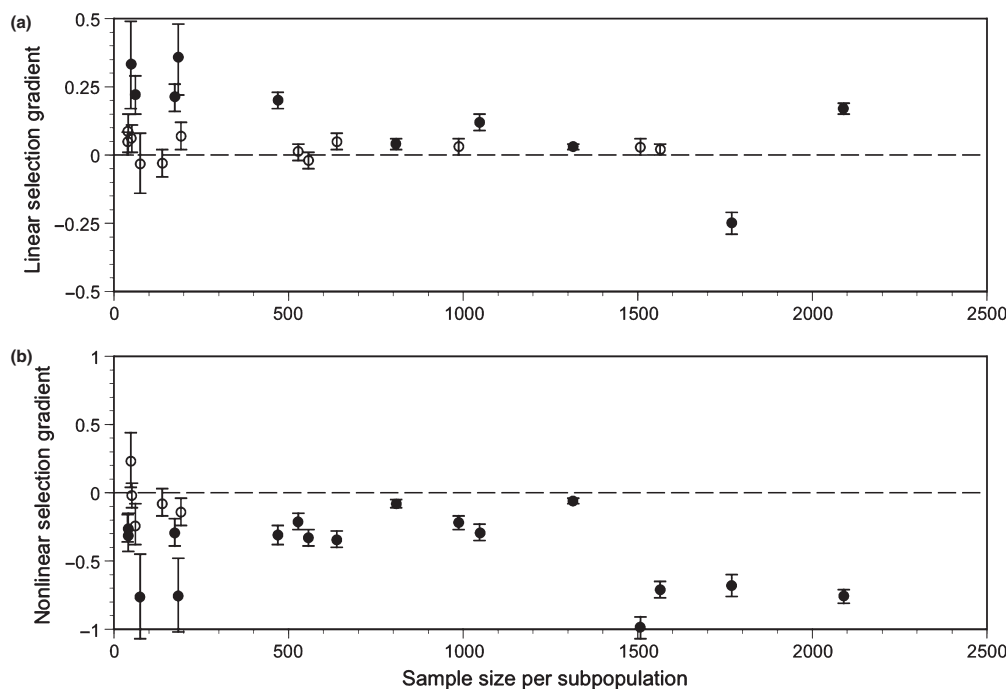

Fig. 1 (a) Linear (β) and (b) nonlinear (γ) selection gradients ($\pm$ SE) as a function of sample size for each of 22 subpopulations of the gall former *Belonocnema treatae*. Filled circles denote estimates that are statistically significant at the level of $P \leq 0.05$, whereas open circles denote estimates that are not significant.

Discussion

The review by Kingsolver et al. (2001) of the distribution of selection gradients estimated across species in natural populations has been accompanied by vigorous discussion and insightful extensions (Conner, 2001; Hoekstra et al., 2001; Hereford et al., 2004; Hersch & Phillips, 2004; Knapczyk & Conner, 2007; Siepielski et al., 2009; Morrissey & Hadfield, 2011). Our analysis of patterns of phenotypic selection among 22 unequally sampled subpopulations of a single species allows us to contribute to this discussion. By treating the gall formers resident on individual trees as subpopulations and estimating host plant-mediated selection on the trait gall size at the level of the subpopulation (Egan et al., 2011), we have compared patterns of selection gradients within this single species to patterns that have emerged from the among-species review of Kingsolver et al. (2001). The patterns of selection evident in our within-species analysis differ markedly from those revealed from among-species comparisons. Our results thus call into question the general applicability of patterns of selection in relation to variation in sample size that emerge from among-species comparisons for predicting patterns of selection in single-species studies.

Kingsolver et al. (2001) reported that only 25% and 16% of estimates of β and γ, respectively, differed from 0 at the 95% probability level. In contrast, in the *B. treatae* – plateau live oak system, selection occurs frequently: statistically significant estimates of β and γ were observed in, respectively, 45% and 77% of subpopulations surveyed. The median values of $|\beta|$ and $|\gamma|$ we observed across all subpopulations were, respectively, consistent with and greater than the median values of $|\beta| = 0.09$ and $|\gamma| = 0.02$ reported by Kingsolver et al. (2001) for selection assessed by differential survivorship. Also consistent with Kingsolver et al. (2001), $|\beta|$ and $|\gamma|$ were distributed exponentially, indicating that most subpopulations were characterized by selection of low magnitude with specific cases of selection of much greater magnitude. However, we rejected the hypothesis that the distribution of γ is centred on zero. Instead, nonlinear selection in this study system was consistently negative. Thus, the observation of Kingsolver et al. (2001) that stabilizing selection is no more commonly observed than is disruptive selection, putatively due to low effect sizes in combination with sampling error, did not hold for the 22 *B. treatae* subpopulations examined.

Selection and sample size

Kingsolver et al. (2001) provided a necessary update to the initial inspection by Endler (1986) of patterns of natural selection across a variety of species. The analyses conducted, the patterns revealed and the questions raised have subsequently provided the framework for investigating the magnitude and frequency of selection in natural populations (Conner, 2001; Hoekstra et al., 2001; Kingsolver & Pfennig, 2007; Siepielski et al., 2009; Morrissey & Hadfield, 2011) and have spurred research aimed at understanding both sources of error associated with estimating selection gradients (Stinchcombe et al., 2002, 2008; Blows & Brooks, 2003; Blows, 2007) and patterns of selection among species (Hereford et al., 2004; Hersch & Phillips, 2004; Knapczyk & Conner, 2007). In particular, the apparent conflation of the magnitude of and variation among estimates of selection with sample size has motivated research into the statistical basis of these patterns. Subsequently, Hersch & Phillips (2004) and Knapczyk & Conner (2007), using different approaches, have, respectively, supported and refuted the hypothesis of publication bias. Additionally, Hersch & Phillips (2004) documented that many studies have lacked the power to detect selection of the estimated magnitude due to small sample sizes and, as a consequence, the estimates reported by Kingsolver et al. (2001) may actually overestimate the median magnitude of selection. In addition, the mechanics and hence interpretation of selection gradient estimation have been called into question as Hereford et al. (2004) have suggested standardizing estimates of selection using the mean value of traits to redefine the scale by which the magnitude of selection is assessed. However, while rescaling increases the perceived magnitude of selection, it does not alter the fundamental patterns noted above between the mean and variation among estimates of selection in relation to sample size. As a consequence of inherent problems in estimating selection gradients, authors from Kingsolver et al. (2001) onwards have implored empiricists aiming to measure selection to (i) identify study systems that afford the opportunity for robust sample sizes (ii) conduct studies that circumvent problems associated with single populations and simultaneous selection on correlated characters and (iii) identify study systems in which stabilizing selection is likely. The *B. treatae* – plateau live oak system provides such a system for analysis.

At the core of the among-species review are the observations that the magnitude and variation among estimated β and γ decrease with increasing sample size. These patterns have been used as the basis for the supposition that studies seeking to estimate β and γ in natural populations will likely reveal that linear and nonlinear selection are of low magnitude or not different from zero given sufficient sample size (Kingsolver & Pfennig, 2007). We cannot directly compare the functional forms of the relationships between β and γ in relation to sample size between the among-species analysis and our within-species analysis because the forms of relationships are not described in the among-species analysis of Kingsolver et al. (2001). Instead, using our data, we tested the null hypotheses of no

relationship between sample size and (i) the ability to detect either form of selection (incidence) (ii) the magnitude of selection, and (iii) the variance among estimates of selection. The relationships between both β and γ and sample size depicted in Kingsolver *et al.* (2001) (Figs 1 and 7, respectively) suggest that the incidence of selection (i.e. the likelihood of detecting a statistically significant selection gradient) should decrease with increasing sample size, particularly for γ. In our study system, however, for neither β nor γ did the incidence of selection follow this prediction. Instead, the incidence of linear selection was independent of sample size; thus, when directional selection occurred, it was of sufficient magnitude to be detected, even in subpopulations with small sample sizes. The importance of heterogeneity in selection is further emphasized by our finding that nonlinear selection was of sufficient magnitude to be detected as being statistically significant in three of the six subpopulations with the smallest sample sizes surveyed. In addition, our observation of an overall positive relationship between the incidence of nonlinear selection and sample size runs counter to the pattern present in the among-species analysis and likely reflects both the sensitivity of estimates of nonlinear selection to sample size (Blows & Brooks, 2003) and the relationship of SE and sample size. Finally, our estimates of both β and γ did not converge on zero with increasing sample size, as suggested by the patterns evident in Figs 1 and 7 of Kingsolver *et al.* (2001). In fact, γ increased with sample size, thus driving the positive relationship between the incidence of nonlinear selection and sample size.

Several aspects of the interaction of the gall former with its host plant may directly or indirectly drive the observed increase in nonlinear selection with sample size. For example, variation in the density of gall formers among the trees may provide an explanation to the observed pattern of nonlinear selection. Sample size variation among the 22 subpopulations for which selection was quantified reflected differences in the numbers of galls initiated within experimental enclosures that were stocked with equivalent numbers of gall formers (Egan & Ott, 2007; Hood & Ott, 2010). The density of gallers naturally resident on these same trees also varied dramatically, ranging, for example, from an average low of 3.8 ± 1.2 galls per $0.25\ m^2$ to 63.6 ± 7.2 galls per $0.25\ m^2$ for trees categorized as low and high density by Egan & Ott (2007). It is notable that all four experimental subpopulations with the largest sample size ($N > 1500$) were characterized as high-density trees. These four subpopulations had the strongest signals of nonlinear selection and exert high leverage on the regression shown in Fig. 1b. Thus, variation in natural galler density may underlie the observed relationship between sample size and the magnitude of nonlinear

selection if density (or a correlate of density) influences the form and magnitude of selection. Most important, our results emphasize that when selection is heterogeneous, predictions based on interspecies comparisons and sampling theory are not likely to predict patterns of variation in selection among unequally sampled populations.

Finally, for neither β nor γ did the variance among estimated selection gradients decrease with increasing sample size. Our results were invariant to the manner in which estimates of β and γ were grouped (i.e. following the $\log_{10}$ scale used by Kingsolver *et al.*, 2001 or according to natural breaks in the ranges of sample sizes present among the subpopulations). Thus, our results, unlike those of Kingsolver *et al.* (2001), do not follow expectations based on the relationship of sample size and sampling variance and the contribution of sampling variance to overall variance (Light & Pillemer, 1984) – again emphasizing the role of heterogeneity in selection.

Investigating the effect of sampling error

Two alternative approaches based on data gathered from natural populations of single species may be used to examine the relationship between sample sizes and selection gradients. In the first approach, given one or more populations and following estimation of selection gradients based on observed sample sizes, the individuals that constitute the population, with their associated phenotype and fitness, can be subsampled with replacement. Bootstrap selection gradients with associated SEs can then be computed for any series of simulated hypothetical population sizes. Knapczyk & Conner (2007) employed this approach. Their results questioned the hypothesis that sampling error in conjunction with publication bias produced the pattern of decreasing variation among selection gradients reported by Kingsolver *et al.* (2001).

Our study provides an alternative means of testing the sampling error hypothesis. We considered estimates of β and γ to be representative of the variation in selection gradients present among subpopulations of gall formers at the study site. As the 22 subpopulations differed meaningfully in characteristics of the phenotype distribution and fitness functions (Egan *et al.*, 2011), the total variation in selection gradients among the subpopulations is both a function of unequal sample size and these biologically meaningful differences. Our test of the hypothesized relationships asked the following question: given variation present in selection gradients among a series of randomly sampled subpopulations, are the hypothesized relationships between sample size and both the magnitude of and variance among selection gradients strong enough to override the biological contribution to the observed variation in selection? The answer is no: in our study

system, among-subpopulation differences in selection appear to outweigh variation in sample size as the primary determinant of the observed magnitude of selection. Our results are cautionary because variation in both the magnitude of selection and sampling intensity are likely common if not ubiquitous features in studies of selection in natural populations. When variably sized samples are available for a series of populations experiencing selection of equivalent magnitude, the intrinsic relationship between sampling error and sample size (Knapczyk & Conner, 2007) should predominate in the estimate of total variation among estimated selection gradients and produce a pattern of decreasing variability among estimates as sample size increases. In contrast, when the magnitude of selection differs significantly among populations, the interaction of variable selection and unequal sampling that leads to sampling error can produce patterns of variation in selection in relation to sample size that do not follow the prediction of Kingsolver et al. (2001). We emphasize that when features of the biology of the organism, the spatial distribution of study units and logistics allow selection in multiple populations to be assessed simultaneously, such studies provide the opportunity to evaluate simultaneously the roles of sample size variation and heterogeneity in selection in producing patterns of selection in natural populations.

Acknowledgments

References

Blows, M.W. 2007. A tale of two matrices: multivariate approaches in evolutionary biology. *J. Evol. Biol.* **20**: 1–8.

Blows, M.W. & Brooks, R. 2003. Measuring nonlinear selection. *Am. Nat.* **162**: 815–820.

Conner, J.K. 2001. How strong is natural selection. *Trends Ecol. Evol.* **16**: 215–217.

Egan, S.P. & Ott, J.R. 2007. Host plant quality and local adaptation determine the distribution of a gall-forming herbivore. *Ecology* **88**: 2868–2879.

Egan, S.P., Hood, G.R. & Ott, J.R. 2011. Natural selection on gall size: variable contributions of individual host plants to population-wide patterns. *Evolution* **65**: 3543–3557.

Endler, J.A. 1986. *Natural Selection in the Wild*. Princeton University Press, Princeton, NJ.

Hendry, A.P. 2005. The power of natural selection. *Nature* **433**: 694–695.

Hereford, J., Hansen, T.F. & Houle, D. 2004. Comparing strengths of directional selection: how strong is strong? *Evolution* **58**: 2133–2143.

Hersch, E.I. & Phillips, P.C. 2004. Power and potential bias in field studies of natural selection. *Evolution* **58**: 479–485.

Hoekstra, H.E., Hoekstra, J.M., Berrigan, D., Vignieri, S.N., Hoang, A., Hill, C.E. et al. 2001. Strength and tempo of directional selection in the wild. *Proc. Natl. Acad. Sci. USA* **98**: 9157–9160.

Hood, G.R. & Ott, J.R. 2010. Developmental plasticity and reduced susceptibility to natural enemies following host plant defoliation in a specialized herbivore. *Oecologia* **162**: 673–683.

Kingsolver, J.G. & Pfennig, D.W. 2007. Patterns and power of phenotypic selection in nature. *Bioscience* **57**: 561–572.

Kingsolver, J.G., Hoekstra, H.E., Hoekstra, J.M., Berrigan, D., Vignieri, S.N., Hill, C.E. et al. 2001. The strength of phenotypic selection in natural populations. *Am. Nat.* **157**: 245–261.

Knapczyk, F.N. & Conner, J.K. 2007. Estimates of the average strength of natural selection are not inflated by sampling error or publication bias. *Am. Nat.* **170**: 501–508.

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

Light, R.J. & Pillemer, D.B. 1984. *Summing Up: The Science of Reviewing Research*. Harvard University Press, Cambridge, MA.

Mayr, G. 1881. *Die genera der gallenbewohnden cynipiden. Zwanzigsten Jahresberichte der Communal Oberrealsch nie im*. I. Bezirke, Wien, pp. 4,16,38.

Morrissey, M.B. & Hadfield, J.D. 2011. Directional selection in temporally replicated studies is remarkably consistent. *Evolution* **66**: 435–442.

Muller, C.H. 1961. The live oaks of the series Virentes. *Am. Midl. Nat.* **65**: 17–39.

Reynolds, R. 2000. *The role of Natural Enemies in Determining the Relationship Between Gall Size and Emergence Success of a Host-Specific Cynipid*. Master's thesis, Department of Biology, Texas State University, San Marcos, TX.

Rosenthal, R. 1979. The "file drawer problem" and tolerance for null results. *Psychol. Bull.* **86**: 638–641.

Siepielski, A.M., DiBattista, J.D. & Carlson, S.M. 2009. It's about time: the temporal dynamics of phenotypic selection in the wild. *Ecol. Lett.* **12**: 1261–1276.

Stinchcombe, J.R., Rutter, M.T., Burdick, D.S., Tiffin, P., Rauscher, M.D. & Mauricio, R. 2002. Testing for environmentally induced bias in phenotypic estimates of natural selection: theory and practice. *Am. Nat.* **160**: 511–523.

Stinchcombe, J.R., Agrawal, A.F., Hohenlohe, P.A., Arnold, S.J. & Blows, M.W. 2008. Estimating nonlinear selection gradients using quadratic regression coefficients: double or nothing? *Evolution* **62**: 2435–2440.

Stone, G.N. & Schönrogge, K. 2003. The adaptive significance of insect gall morphology. *Trends Ecol. Evol.* **18**: 512–522.