Reducing bias in the measurement of selection

Rodney Mauricio and Laura E. Mojonnier

Evolutionary ecologists have long been interested in describing and quantifying the action of selection in natural populations. Although examples of the measurement of natural selection date to the turn of the century[1,2], Lande and Arnold's[3] paper in 1983 launched a renewed interest in the measurement of selection in field populations of both animals and plants. This landmark paper appeared at a time when a rigorous theoretical framework for the study of the evolution of complex phenotypes was being developed[4,5].

Lande and Arnold's method allowed field biologists to obtain quantitative estimates of the form and magnitude of selection acting directly on each of a set of possibly correlated traits by simply measuring phenotypic values of the traits and fitness for the members of a population, and then performing a multiple regression of relative fitness on the trait values (Box 1). Using this method, selection could be characterized using data collected within a single generation. In addition, one could quantitatively predict the evolutionary response of the population to selection by combining these estimates of selection with estimates of the genetic variation and covariation of the traits obtained separately[3,4] (Box 1). Since 1983, other advances have been made in the measurement of natural selection, most notably the use of nonparametric analyses and graphical representations of selection[6–8] (see Ref. 9 for a recent and thorough review).

Although these techniques have proved extremely useful, several authors[10–16] have pointed out the following potentially severe limitation: phenotypic correlations between trait values and fitness (which these techniques rely on to characterize selection) do not give an accurate view of the causal relationship between the phenotype and fitness when environmental factors also contribute to covariances between trait values and fitness (Box 2). Here, we describe how environmentally induced covariances can result in severely biased estimates of selection from standard phenotypic methods. We then describe methods for dealing with this problem, focusing on a 'genetic' method for estimating selection (recently developed by Rausher[15]) which can eliminate bias from environmental covariances, even if the environmental factors responsible for the bias are unknown.

The problem of environmental covariances

Environmental covariances between traits and fitness can arise whenever environmental factors that influence trait values also affect fitness, either directly or through an unmeasured trait (Fig. 1). For example, consider a hypothetical

Selection on quantitative characters is commonly measured in natural populations using regression techniques based on phenotypic covariances between traits and fitness. However, such methods do not give an accurate view of the causal relationship between the phenotype and fitness if environmental factors also contribute to covariances between traits and fitness. A recently developed method for estimating selection eliminates the problem of bias resulting from environmental covariances. This underappreciated method represents a significant addition to the toolbox of the evolutionary ecologist.

Rodney Mauricio is at the Dept of Ecology and Evolution, University of Chicago, Chicago, IL 60637, USA (mauricio@uchicago.edu); Laura Mojonnier is at the Dept of Zoology, Duke University, Durham, NC 27708-0325, USA (laura@mendel.zoo.duke.edu).

population of a species of plant that exhibits genetic variation for the mean size of its flowers, but in which genetically based differences in flower size have no effect on fitness. Flower size is, however, affected by an environmental factor, soil water content, with the result that plants produce small flowers when grown in dry soil and large flowers when grown in wet soil. Because water is a limiting resource in this population, plants growing in relatively wet soil also tend to have higher fitness than plants growing in drier soil.

If we were to examine these plants in an area heterogeneous for soil water content, we would find a positive relationship between the mean flower size of a plant and its fitness (Fig. 2). Since standard measurements of selection are based on the relationship between phenotypic values of fitness and traits, we would conclude that directional selection is acting to favor a larger flower size. Since there is genetic variation for flower size, we would also predict an increase in flower size in the next generation (Box 1). However, the positive relationship between flower size and fitness is due entirely to differences in the soil water available to individual plants, rather than any causal effect of flower size on fitness. Essentially, water availability is affecting fitness and flower size is simply a marker for soil water content.

An evolutionary response to selection requires that there be additive genetic covariance between the character and fitness[18,19]. Since flower size does not affect fitness, genotypes that differ in flower size should not differ in fitness, and there can be no genetic covariance between flower size and fitness. Therefore, flower size will not respond to selection. The environmental covariance between flower size and fitness gives a false appearance of selection, resulting in false predictions of evolutionary change.

When natural selection is actually acting on the trait of interest, environmental covariances can result in a distorted view of the form and magnitude of selection. For example, consider a hypothetical study designed to measure selection on adult cuticle pigmentation in a species of beetle whose larvae feed on two different species of host plant (species A and B). The two host plant species differ in nutritional quality, such that beetles that feed mainly on species A as larvae tend to have higher fitness as adults than beetles that feed mainly on plant species B. In addition to being of higher nutritional quality, host plant species A contains a higher concentration of a chemical that adult beetles need for pigment production, so that, in addition to having higher fitness, beetles that eat a higher proportion of host plant

 433

species A tend to be darker as adults. Differences in the composition of the larval diet may thus induce a positive environmental covariance between cuticle pigmentation and fitness.

Such environmental covariance may bias estimates of selection from the Lande–Arnold method. For example, consider a situation in which darker beetles tend to be more visible to predators, and thus less likely to survive to reproduce than lighter beetles. Natural selection, as mediated through the behavior of predators, would then result in a negative relationship between cuticle pigmentation and fitness. The overall phenotypic covariance between cuticle pigmentation and fitness would thus be the sum of a negative covariance reflecting the action of natural selection, and a positive environmental covariance resulting from differences in the composition of the larval diet. Depending on the relative magnitudes of these two covariances, a researcher measuring selection based on the phenotypic covariance between cuticle pigmentation and fitness might conclude that selection favored darker pigmentation, lighter pigmentation, or did not act on cuticle pigmentation at all.

Although the preceding examples are hypothetical, environmentally induced covariances between a trait and fitness may be common in nature, especially for traits that (like fitness) vary depending on the 'condition' of the organism. For example, environmental covariances caused by effects of nutritional state on both traits and fitness have been proposed to mask the pattern of selection on several traits in birds, including breeding date[11], clutch size[12] and tarsus length[13]. Recent studies of phenotypic selection on plant

characters that vary with environmental conditions (e.g. floral traits[20], flowering time[21] and physiological traits[22]), may also be subject to bias from environmental covariances between traits and fitness.

Solutions

A number of methods may be used to reduce bias from environmental covariances. For some traits, environmental covariances between trait values and fitness can be removed by experimental manipulation[14,23]. For example, by physically manipulating tail length in male swallows (*Hirundo rustica*), Møller was able to remove positive effects of nutrition on both tail length and fitness components, and thus uncover the actual pattern of selection on this trait[14,24-26].

When measuring selection on naturally occurring phenotypic variation, several techniques may be used to remove environmental covariances statistically. If the factors responsible for an environmental covariance between trait values and fitness are known, it may be possible to avoid bias by measuring the level of each environmental factor experienced by each individual, and then including these environmental factors as covariates in a standard Lande–Arnold type analysis[10].

Environmental factors often may not affect fitness directly, but instead affect characters that in turn affect fitness. In such cases, bias from environmental covariances can be considered to result from the failure to include all relevant characters in the analysis (Box 2), and can be avoided by identifying and including these intervening characters. When known, the confounding effects of environmental factors or intervening characters may also be removed using path-analytical techniques[27].

Frequently, however, numerous environmental factors may contribute to environmental covariances in subtle and complex ways, and identifying and including all relevant environmental factors and/or correlated characters may be extremely difficult, if not impossible. Such cases require a method for obtaining estimates of selection that are free from bias due to environmental covariances, even if the environmental factors responsible for these covariances remain unknown. Two such methods have been developed[14,15]. The first method, proposed by Schluter et al.[14], involves performing a linear regression of the trait values of members of the population of interest on the trait values of relatives (e.g. parent–offspring regression), and then including the residuals as an independent variable in a standard Lande–Arnold analysis. The second method, developed by Rausher[15], involves the use of breeding values in the regression analysis, and will serve as the focus of the rest of this review.

A 'genetic' method for measuring selection

Rausher's[15] 'genetic' method is identical to the multivariate regression method proposed by Lande and Arnold[3] (the 'phenotypic method'; Box 1) except that estimates of additive genetic (or 'breeding') values are used in place of phenotypic values. Directional selection gradients are estimated as the partial regression coefficients from a linear multiple regression of breeding values for relative fitness on breeding values of the characters, and non-linear selection gradients are estimated as the partial regression coefficients associated with the second-order coefficients from a quadratic multiple regression using breeding values. The estimation of breeding values requires that members of the population fall into known family groups (e.g. paternal half-sib families). Breeding values can be calculated from such family-structured data using techniques developed in plant and animal breeding[28-30]. Alternatively, family means may be

used to approximate breeding values, given certain assumptions[15].

Rausher[15] provides a formal mathematical demonstration that the use of breeding values in place of phenotypic values eliminates bias due to environmental covariances. Intuitively, one can think of the calculation of family means or breeding values as averaging out the effects of environmental factors, with the result that any correlation between the means of fitness and trait values should reflect causal relationships only.
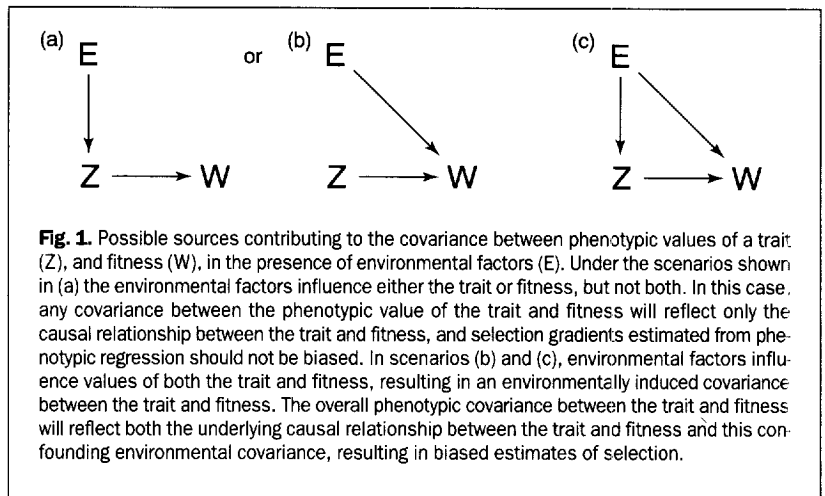
Comparison of the genetic and phenotypic methods

By obscuring the actual causal relationship between phenotype and fitness, environmental covariances can greatly bias estimates of selection from standard phenotypic methods. Thus, when environmental covariances are strong, some method of reducing bias is needed to characterize selection accurately and to predict evolutionary change. The genetic method can eliminate bias from environmental covariances, but may not be practical under some circumstances. Because data consist of family means or breeding values rather than of individual phenotypic values, the use of the genetic method reduces the effective number of observations for a given size of experiment and therefore reduces statistical power. This reduction in sample size may be especially problematic when trying to detect curvature in the relationship between trait values and fitness, which can require many observations[3,9]. Also, error in the estimation of breeding values may make the form of selection more difficult to detect.

In addition to requiring a larger number of individuals to estimate selection, obtaining accurate estimates of breeding values generally requires crosses to be performed; this may not be practical in some species or in non-experimental populations. Not infrequently, however, researchers measure genetic variances and covariances in the same population in which they measure selection. In such cases, it should be possible to estimate breeding values and measure selection using the genetic method.

Finally, the genetic method requires that the traits of interest be genetically variable, while the phenotypic method can provide estimates of selection even in the absence of genetic variation. That the genetic method involves the measurement of selection using genetic rather than phenotypic variation may seem counterintuitive given the common view that natural selection acts directly on the phenotype and only indirectly on the genotype (although many evolutionary biologists explicitly define selection as differential reproductive success of genotypes[31-33]). However, the genetic and phenotypic methods give identical estimates of selection gradients, provided that environmental covariances are absent[15]. Estimates of selection from the two methods will only differ when environmental covariances have caused bias in estimates from the phenotypic method.

An additional advantage of the genetic method over the standard phenotypic method is that estimates of selection gradients are less likely to be biased by the failure to identify and measure important characters. With the phenotypic method, the omission of any character that is both phenotypically correlated with the measured characters and subject to selection can result in biased estimates of the forces of selection on the measured characters. In contrast, with the genetic method, bias will not result through the omission of phenotypically correlated traits that are neither genetically

Fig. 1. Possible sources contributing to the covariance between phenotypic values of a trait (Z), and fitness (W), in the presence of environmental factors (E). Under the scenarios shown in (a) the environmental factors influence either the trait or fitness, but not both. In this case, any covariance between the phenotypic value of the trait and fitness will reflect only the causal relationship between the trait and fitness, and selection gradients estimated from phenotypic regression should not be biased. In scenarios (b) and (c), environmental factors influence values of both the trait and fitness, resulting in an environmentally induced covariance between the trait and fitness. The overall phenotypic covariance between the trait and fitness will reflect both the underlying causal relationship between the trait and fitness and this confounding environmental covariance, resulting in biased estimates of selection.

variable nor genetically correlated with the measured traits. However, both phenotypic and genetic analyses will be biased by the omission of traits that are genetically correlated with the measured traits and that experience selection; these analyses are subject to several other assumptions and limitations common to regression methods for measuring selection (Box 2).

Under what circumstances are environmental covariances likely to be strong enough to significantly bias estimates of selection from standard Lande–Arnold analyses? In general, traits that vary with the condition of the organism (such as flower size in the example given earlier) should be more likely to be subject to environmental covariances with fitness than traits that are fixed early in development (e.g. vertebrae number in garter snakes[34]). Spatial heterogeneity in environmental conditions may also be more likely to result in strong environmental covariances in sessile organisms, such as plants, than in organisms that are mobile, such as many animals. However, there are likely to be many exceptions to these general rules.
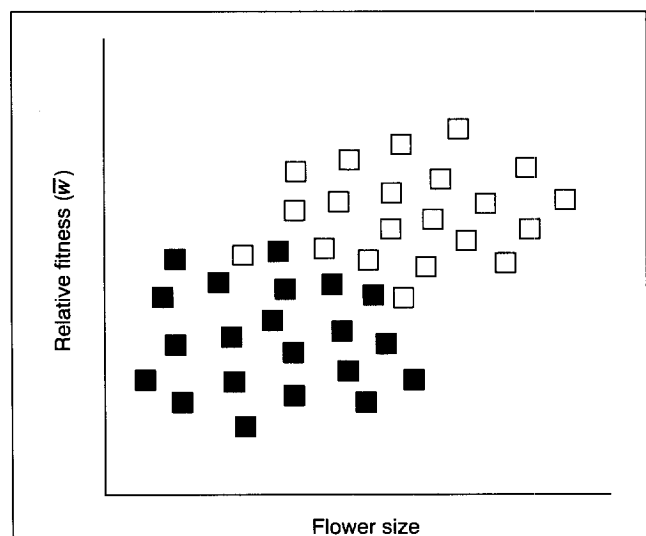

Fig. 2. The relationship between flower size and relative fitness ($\bar{w}$) Plants growing in dry soil (black squares) tend to have low fitness because of lack of water. Plants growing in wet soil (white squares) tend to have high fitness because water is beneficial. Within each of these groups, there is no significant relationship between flower size and relative fitness, reflecting a lack of any causal effect of flower size on fitness. However, because flower size depends on water content of the soil, there appears to be a significant positive relationship between relative fitness and flower size. Thus, the environmental covariance between flower size and fitness yields a false appearance of selection.

A straightforward way to determine definitively whether the phenotypic method provides an accurate view of the forces of selection is to perform both the genetic and phenotypic analyses on the same data, and then to compare the results. Rausher[15] provides a statistical technique for determining whether selection gradients calculated using the phenotypic and genetic methods are significantly different. If estimates of selection from the two techniques are sufficiently similar, the phenotypic method can be used with confidence.

Despite its limitations, the Lande–Arnold method represents a powerful tool for studying natural selection. Along with nonparametric and graphical analyses, the genetic method of estimating selection is yet another tool that evolutionary biologists can apply to describe the action of selection in nature. Combined with experimental manipulations, descriptions of selection from these methods can yield powerful inferences about the operation of selection in natural populations[10,35].

Acknowledgements

References

1 Bumpus, H.C. (1899) The elimination of the unfit as illustrated by the introduced sparrow, *Passer domesticus*, *Biol. Lect. Marine Biol. Lab. Woods Hole* 6, 209–226

2 Weldon, W.F.R. (1901) A first study of natural selection in *Clausilia laminata* (Montagu), *Biometrika* 1, 109–124

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

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

5 Turelli, M. (1984) Heritable genetic variation via mutation-selection balance: Lerch's zeta meets the abdominal bristle, *Theor. Popul. Biol.* 25, 138–193

6 Schluter, D. (1988) Estimating the form of natural selection on a quantitative trait, *Evolution* 42, 849–861

7 Phillips, P.C. and Arnold, S.J. (1989) Visualizing multivariate selection, *Evolution* 43, 1209–1222

8 Schluter, D. and Nychka, D. (1994) Exploring fitness surfaces, *Am. Nat.* 143, 597–616

9 Brodie, E.D., III, Moore, A.J. and Janzen, F.J. (1995) Visualizing and quantifying natural selection, *Trends Ecol. Evol.* 10, 313–318

10 Mitchell-Olds, T. and Shaw, R.G. (1987) Regression analysis of natural selection: statistical inference and biological interpretation, *Evolution* 41, 1149–1161

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

12 Price, T. and Liou, L. (1989) Selection on clutch size in birds, *Am. Nat.* 134, 950–959

13 Alatalo, R.V., Gustafsson, L. and Lundberg, A. (1990) Phenotypic selection on heritable size traits: environmental variance and genetic response, *Am. Nat.* 135, 464–471

14 Schluter, D., Price, T.D. and Rowe, L. (1991) Conflicting selection pressures and life history trade-offs, *Proc. R. Soc. London Ser. B* 246, 11–17

15 Rausher, M.D. (1992) The measurement of selection on quantitative traits: biases due to the environmental covariances between traits and fitness, *Evolution* 46, 616–626

16 van Tienderen, P.H. and de Jong, G. (1994) A general model of the relation between phenotypic selection and genetic response, *J. Evol. Biol.* 7, 1–12

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

18 Price, G.R. (1970) Selection and covariance, *Nature* 227, 520–521

19 Crow, J.F. and Nagylaki, T. (1976) The rate of change of a character correlated with fitness, *Am. Nat.* 110, 207–213

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

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

22 Dudley, S.A. (1996) Differing selection on plant physiological traits in response to environmental water availability: a test of adaptive hypotheses, *Evolution* 50, 92–102

23 Sinervo, B. *et al.* (1992) Allometric engineering: a causal analysis of natural selection on offspring size, *Science* 258, 1927–1930

24 Møller, A.P. (1988) Female choice selects for male sexual tail ornaments in the monogamous swallow, *Nature* 332, 640–642

25 Møller, A.P. (1989) Viability costs of male tail ornaments in a swallow, *Nature* 339, 132–135

26 Møller, A.P. (1991) Viability is positively related to a degree of ornamentation in male swallows, *Proc. R. Soc. London Ser. B* 243, 145–148

27 Sinervo, B. and DeNardo, D.F. (1996) Costs of reproduction in the wild: path analysis of natural selection and experimental tests of causation, *Evolution* 50, 1299–1313

28 Henderson, C.R. (1963) Selection index and expected genetic advance, in *Statistical Genetics and Plant Breeding* (Hanson, W.D. and Robinson, H.F., eds), pp. 141–163, National Academy of Sciences and National Research Council

29 Shaw, R.G., Platenkamp, G.A.J. and Shaw, F.H. (1995) Quantitative genetics of response to competitors in *Nemophila menziesii*: a field experiment, *Genetics* 139, 397–406

30 Littel, G.A. *et al.* (1996) *SAS System for Mixed Models*, SAS Institute

31 Dobzhansky, T. (1979) *Genetics of the Evolutionary Process*, Columbia University Press

32 Endler, J.A. (1986) *Natural Selection in the Wild*, Princeton University Press

33 Hartl, D.L. and Clark, A.G. (1989) *Principles of Population Genetics*, Sinauer

34 Arnold, S.J. (1988) Quantitative genetics and selection in natural populations: microevolution of vertebral numbers in the garter snake *Thamnophis elegans*, in *Proceedings of the Second International Conference on Quantitative Genetics* (Weir, B.S. *et al.*, eds), pp. 619–636, Sinauer

35 Wade, M.J. and Kalisz, S. (1990) The causes of natural selection, *Evolution* 44, 1947–1955

Coming soon in *TREE*: