

Using path analysis to measure natural selection

S. M. SCHEINER,* R. J. MITCHELL† & H. S. CALLAHAN‡

*Department of Life Sciences, Arizona State University West, PO Box 37100, Phoenix, AZ 85069, USA

†Department of Biology, University of Akron, Akron, OH, 44325-3908, USA

‡Department of Botany, University of Tennessee, Knoxville, TN 37996, USA

Keywords:

Arabidopsis thaliana;
directional selection;
natural selection;
nonlinear selection;
path analysis.

Abstract

We expand current methods for calculating selection coefficients using path analysis and demonstrate how to analyse nonlinear selection. While this incorporation is a straightforward extension of current procedures, the rules for combining these traits to calculate selection coefficients can be complex. We demonstrate our method with an analysis of selection in an experimental population of *Arabidopsis thaliana* consisting of 289 individuals. Multiple regression analyses found positive directional selection and positive nonlinear selection only for inflorescence height. In contrast, the path analyses also revealed positive directional selection for number of rosette leaves and positive nonlinear selection for leaf number and time of inflorescence initiation. These changes in conclusions came about because indirect selection was converted into direct selection with the change in causal structure. Path analysis has great promise for improving our understanding of natural selection but must be used with caution since coefficient estimates depend on the assumed causal structure.

Introduction

The process of evolution by natural selection can be divided into two components, phenotypic selection operating within a single generation and a genetic response operating across generations (Fisher, 1930). Over the past 15 years we have greatly refined our ability to analyse phenotypic selection. Lande & Arnold (1983) presented a fundamental methodological advance by recognizing that phenotypic selection could be modelled as a multivariate analysis of the covariance between trait values and fitness. Their initial effort concerned the estimation of directional (linear) and stabilizing/disruptive (nonlinear) selection on multiple traits. They proposed the use of multiple regression to partition total selection on a trait into components due to direct selection on the trait of interest and those due to correlated responses through selection on other traits

(indirect selection). Since that publication the method has been refined and advanced by partitioning selection into multiple episodes (Arnold & Wade, 1984; Wade & Kalisz, 1989), recognizing nonlinear directional-selection components (Mitchell-Olds & Shaw, 1987), adding measurements of correlational selection (Phillips & Arnold, 1989), analysing selection on breeding values (Rauscher, 1992), and the use of logistic regression (Janzen & Stern, 1998).

This paper extends those efforts to the analysis of selection on complex life histories by the use of path analysis. The critical difference between path analysis and multiple regression is that in the former the analytical model is built around a specific set of causal relationships among traits that determine fitness. In contrast, a multiple regression assumes a simpler causal relationship in which all traits affect fitness directly. As we show, this change in causal structure can substantially alter conclusions about the strength of phenotypic selection on given traits.

Path analysis has been used previously to examine patterns of phenotypic selection (e.g. Arnold, 1983; Crespi & Bookstein, 1989; Mitchell-Olds & Bergelson,

Correspondence: S. M. Scheiner, Division of Environmental Biology, Rm 635, National Science Foundation, 4201 Wilson Blvd, Arlington, VA 22230, USA.
Tel.: +1 703 306 1481; fax: +1 703 306 0817;
e-mail: sscheine@nsf.gov

1990; Conner, 1996; Pugsek & Tomer, 1996; see review in Kingsolver & Schemske, 1991). These authors make a point that we reiterate here. Complex life histories involve traits that are not directly linked to fitness, but rather are linked through other traits. An analysis of selection on these traits requires accounting for these causal relationships. Our paper makes two important advances to these previous efforts. First, we show how a path analysis provides clues about whether important unmeasured traits have been left out of the analysis. The issue of missing traits has been a persistent criticism about multiple regression methods for analysing selection. We do this by examining the partition of total selection into direct and indirect selection components, an issue not examined in previous papers. Note that this partition is different from direct and indirect effects (see below). Second, we explicate methods for analysing nonlinear components of selection, such as stabilizing selection. Most previous path analyses have been limited to linear selection components (but see McCall *et al.*, 1994; Conner, 1996). In Scheiner & Callahan (1999) we extend these methods further to the measurement of selection on phenotypic plasticity. We illustrate these methods using data on selection in an experimental population of *Arabidopsis thaliana*.

Selection in a path-analytical framework

Path analysis is a generalization of multiple regression that allows one to estimate the strength and sign of directional relationships for complicated causal schemes with multiple dependent variables (Wright, 1920; Li, 1975). The causal scheme is usually considered an *a priori* hypothesis of potential effects (but see Shipley, 1997), and alternative hypotheses can be proposed and tested against each other (Mitchell, 1993). Conversely, the *a priori* causal scheme can be taken as a given and used to make predictions about patterns of evolution. It is in this latter sense that we use our method, where the pieces of the causal scheme are built either from first principles or previous experimental and observational studies.

A path diagram (Fig. 1) is a scheme of causal relationships. Consider an annual plant that grows vegetatively for some period of time, then ceases growth, flowers, sets seed and dies. More complex life cycles can be accommodated with this method; the example developed here is for simplicity of presentation. Five traits are measured: cotyledon size (z_1), time of inflorescence initiation (bolting time; z_2), number of rosette leaves at flowering initiation (z_3), inflorescence height (z_4) and number of fruits (z_5). In our causal scheme, cotyledon size affects both time of inflorescence initiation and number of leaves, and both of them affect inflorescence height. Inflorescence height in turn influences fruit production. In this formulation, only first-order effects are included. That is, inflorescence height at the end of the season depends only on timing of inflorescence initiation and

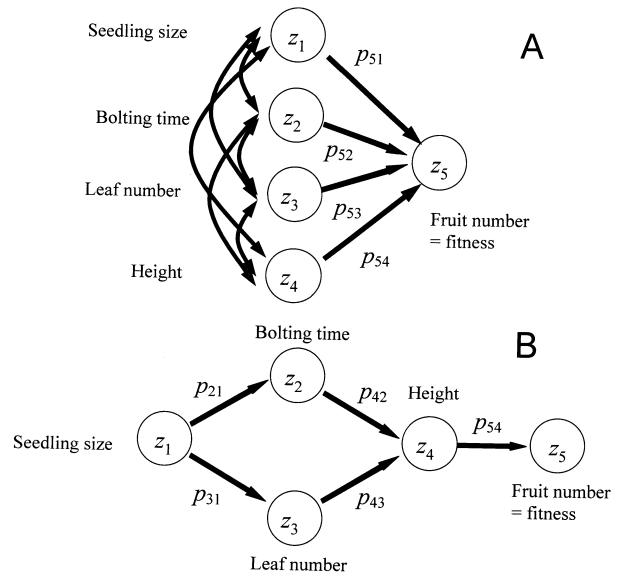


Fig. 1 Two different models of trait effects on fitness. A. Multiple regression model showing each trait operating simultaneously on fitness. B. Path analysis model showing five traits at four time periods. The path coefficients are standardized regression coefficients. Variation due to error (U) is not included for simplicity.

number of rosette leaves, not additionally on cotyledon size. Such second-order effects could be added by including additional paths.

A path diagram, besides showing the nature and direction of causal relationships, also includes estimates of the strength of those relationships, the path coefficients (p). A path coefficient is the standardized slope of the regression of the dependent variable on the independent variable in the context of the other independent variables. For example, inflorescence height (z_4) is regressed on bolting time (z_2). The slope (b_{42}) is then standardized (p_{42}) by multiplying it by the ratio of the standard deviations of the independent and dependent variables, respectively. If there is only a single independent variable, this standardized coefficient is a Pearson product-moment correlation; if there are additional independent variables, it is a standardized partial regression coefficient. The standardization acts to remove differences in scale among variables. Typically these relationships are assumed to be monotonic and linear, possibly after transformation. However, nonmonotonic (e.g. quadratic) relationships can be included by adding squared traits (see below). If a trait is transformed, the same transformation must be used for that trait as both a dependent and an independent variable.

Terminology

When moving from the multiple regression formulation of Lande & Arnold (1983) to a path-analytical

framework, we run into a terminology discrepancy between statisticians and evolutionary biologists. To avoid confusion it is necessary to refine previous word usage. In doing so we bring precision while avoiding unnecessary jargon.

In Lande & Arnold's multiple regression framework there are only two ways a trait can affect fitness, a direct connection between the two traits, and a connection that proceeds backwards through a correlation with another trait and then forward to fitness (Fig. 1A). Lande & Arnold refer to these two possible pathways as direct and indirect effects, respectively. A path-analytical framework adds a third possibility. A trait appearing early in the model may have its effect through an intermediate trait or traits and ultimately influence fitness. In standard path analysis parlance (Pedhazur, 1982; see also Mitchell-Olds & Bergelson, 1990), a forward connection through an intermediate trait is referred to as an indirect effect (or a mediated effect), while a backward connection is a noncausal effect that includes both spurious effects due to shared causes and correlated effects due to correlated causes (e.g. in Fig. 1B, cotyledon size is a shared cause for leaf number and bolt date, while in Fig. 1A fitness is affected by four correlated causes). To add to the confusion, in selection analysis, correlational selection refers to simultaneous selection on a combination of traits (Phillips & Arnold, 1989).

In order to avoid misunderstanding while not straying too far from previous usage, we propose the following terminology which replaces 'effect' with 'selection' when referring to the measures described by Lande & Arnold, and reserves the term 'effect' for describing paths or connections between traits. Direct selection is any forward connection between a trait and fitness, whether through an intermediate trait or not. In path analysis terminology this is the sum of all direct and indirect effects associated with a given trait. Indirect selection is any backward connection between a trait and fitness. In path analysis this is often called the noncausal effect. Total selection is the sum of direct and indirect selection. Correlational selection is defined as above (Phillips & Arnold, 1989). A direct effect is any direct connection between traits. An indirect effect is any forward connection between two traits that goes through an intermedi-

ate trait. The causal effect is the sum of the direct and indirect effects (also known as the total effect or the total causal effect of Mitchell-Olds & Bergelson, 1990). A noncausal effect is any backward connection between two traits, including both spurious and correlated effects.

Partitioning the selection differential

The covariance between a trait and fitness is the selection differential (*s*). Such a covariance might develop for several reasons (Arnold, 1983). In a multiple regression framework, there might be direct or indirect selection on the trait, i.e. *s* describes total selection. In a path-analytical framework, it might develop because of direct effects, indirect effects or noncausal effects (Pedhazur, 1982; Schemske & Horvitz, 1988). Given the definitions mentioned earlier, it is therefore possible to use path analysis to partition the selection differential into direct selection (direct + indirect effects) and indirect selection (noncausal effects).

Partitioning an observed selection differential into these components depends on the causal model applied to the system. Consider the multiple regression model used in a typical selection analysis (Fig. 1A). In this model there is no hierarchy of relationships among traits; all four of the observed traits influence fitness directly, and are correlated with one another. This model therefore only allows direct and noncausal effects on fitness, since there are no intermediate traits through which indirect effects might arise. Contrast the path model (Fig. 1B). In this model only one trait (height) has a path leading directly to fitness with no intermediate steps, but all other traits may have indirect (mediated) or noncausal effects on fitness (Table 1). In this example, direct selection on inflorescence height is straightforward, $p_{54} \equiv \beta_4^*$, the regression of fruit production on inflorescence height. We use β to symbolize the selection coefficient to maintain consistency with the conventions established by Lande & Arnold (1983), but add the asterisk to indicate that the value may differ from that estimated using multiple regression (see below).

Direct selection coefficients (β_i^*) are calculated as the sum of the direct and indirect effects for a trait; while the noncausal effects estimate indirect selection. Consider

Table 1 Decomposition of the correlation between different traits and fitness under multiple regression and path analysis models (Fig. 1). Direct selection includes both direct and indirect effects, and indirect selection includes noncausal (spurious and correlational) effects. The sum of direct and indirect selection is the total selection accounted for by the model.

Trait	Total selection	Multiple regression		Path analysis	
		Direct selection	Indirect selection	Direct selection	Indirect selection
Seedling size	s_1	p_{51}	$r_{21} p_{52} + r_{32} p_{53} + r_{41} p_{54}$	$p_{21} p_{42} p_{54} + p_{31} p_{43} p_{54}$	
Bolting time	s_2	p_{52}	$r_{21} p_{51} + r_{32} p_{53} + r_{42} p_{54}$	$p_{42} p_{54}$	$p_{21} p_{31} p_{43} p_{54}$
Leaf number	s_3	p_{53}	$r_{31} p_{51} + r_{32} p_{52} + r_{43} p_{54}$	$p_{43} p_{54}$	$p_{31} p_{21} p_{42} p_{54}$
Height	s_4	p_{54}	$r_{41} p_{51} + r_{42} p_{52} + r_{42} p_{54}$	p_{54}	

selection on leaf number. In this model, leaf number has no direct effect on fitness, but the indirect effect of leaf number on fruit production through inflorescence height estimates the strength of direct selection, and is calculated as $\beta_3^* \equiv p_{43} p_{54}$. Indirect selection goes along a second pathway through a shared cause, seedling size. It is calculated by following the path backwards from leaf number to seedling size, then forward through bolting time to fitness ($p_{31} p_{21} p_{42} p_{54}$). Total selection on leaf number can be calculated by adding together direct and indirect selection: $s_3^* \equiv p_{43} p_{54} + p_{31} p_{21} p_{42} p_{54}$ (the asterisk indicates that this estimate relates to a path diagram, as above). Along a single path, coefficients are multiplied, while separate paths are summed. See Li (1975) for details on combining path coefficients. In complicated path diagrams there may be many pathways; however, the rules for calculating path coefficients are straightforward (Li, 1975; Pedhazur, 1982). Several computer programs calculate path coefficients automatically [e.g. Procedure CALIS (SAS Institute 1989a,b), LISREL (Jöreskog & Sörbom, 1988), EQS (Bentler, 1993), RAMONA (SYSTAT for Windows, SPSS, Inc.)], and the more sophisticated versions offer a variety of estimation options, including ordinary least squares and maximum likelihood. Most major software packages calculate both standardized and unstandardized coefficients, so one must take care when reading the output to make sure which coefficients are being reported where.

We denote the model-implied covariance between a trait and fitness as s^* (the 'predicted' covariance; Cohen & Cohen, 1983), to distinguish it from the observed selection differential (s) which does not depend on the causal model. Likewise, we distinguish between selection gradients estimated from a path model (β_i^* ; see also Koenig *et al.*, 1991) and those estimated from a linear regression (β_i). Structural equation modelling programs readily calculate s^* and β_i^* values as part of the model-dependent predicted covariance matrix.

The sum of direct and indirect selection estimates the selection differential (s^*) for a trait. The model-implied selection differential (s^*) need not equal the observed

selection differential (s) when the path diagram is 'over-identified' (Pedhazur, 1982; Hayduk, 1987; see below). For example, our model (Fig. 1B) implies a covariance between leaf number and fitness of s_3^* (Table 1), a value which may deviate from the observed covariance (s_3) for a variety of reasons, including misspecification of the causal model and sampling error.

In this fashion a path-analytical framework differs from a multiple regression framework, which will of necessity exactly recreate the original covariance matrix from the selection coefficients because it is 'just-identified' (Pedhazur, 1982; Hayduk, 1987), having just enough information to estimate the regression coefficients. In contrast, many path models are over-identified, in that they contain more information than is needed to estimate the path coefficients. For example, in the linear path model (Fig. 1B) there are 10 correlations among the five traits, but the path model only estimates five path coefficients.

This redundancy of information due to over-identification has two consequences. First, it means that the observed covariance between variables may differ from that implied by the model (Table 2). These deviations are an indication of the extent to which the data are consistent with the path diagram. Second, these deviations can be used to test the goodness-of-fit between the model and the data (Pedhazur, 1982; Loehlin, 1987; Hayduk, 1988; Mitchell, 1993).

As a result, statistical significance testing plays a different role in a path-analytical framework. When a direct selection coefficient comes from combining direct and indirect effects, its statistical significance is not directly tested. One can only test the significance of individual path coefficients (Jöreskog & Sörbom, 1988). Instead, one tests the statistical significance of the entire model and compares alternative models (Loehlin, 1987; Hayduk, 1988; Mitchell, 1993). After a model is chosen, the selection coefficients are calculated and their relative magnitudes compared. There are no unambiguous guidelines about how to interpret such selection coefficients. Magnitudes of $<|0.2|$ are typically considered

Table 2 Linear selection analyses showing direct and indirect selection coefficients for multiple regression and path analysis models.

Coefficients in bold type are statistically significant at $P < 0.0001$. Statistical significance can be assessed for all traits for total selection and direct selection in the multiple regression model. In the path analysis model, significance can be assessed only for direct selection on height because only this trait is directly connected to fitness. Indirect selection in the path analysis model is calculated by subtracting direct selection (total effects) from s^* (the model-implied correlation between the trait and fitness). R^2 for the multiple regression model = 0.761. R^2 values for the path model are: Bolting time = 0.045, Leaf number = 0.011, Height = 0.359, Fitness = 0.749.

Trait	Multiple regression			Path analysis		
	Total selection (s)	Direct selection (β)	Indirect selection	Predicted covariance (s^*)	Direct selection (β^*)	Indirect selection
Seedling size	0.216	0.004	0.212	0.070	0.070	–
Bolting time	–0.153	–0.081	–0.072	–0.094	–0.084	–0.010
Leaf number	0.589	0.130	0.459	0.512	0.510	0.002
Height	0.865	0.780	0.085	0.865	0.865	–

small, those from |0.2| to |0.4| are considered moderate and greater values are considered large.

Nonlinear and stabilizing/disruptive selection

Above we presented a path-analytical framework for directional selection, representing selection which changes population means. Selection can also be nonlinear, representing selection which changes population variances. Such nonlinear selection coefficients are typically symbolized as γ . To incorporate selection on variances into a path analysis, one creates new traits, which are squares of the original traits (Fig. 2). Such analysis requires that the original trait be centred to a mean of 0 by subtracting the population mean from each observation. In the analyses below we used z-scores, calculated by subtracting the mean and then dividing by the standard deviation, in order to simplify interpretation (Pedhazur, 1982; Lande & Arnold, 1983). Our use of squared traits to indicate quadratic paths is actually equivalent to what is done in a multiple regression analysis. Such an analysis, however, is seldom illustrated, only shown in equation form. Thus, the relationship is not immediately apparent.

The path diagram indicates that these squared traits are correlated with (rather than caused by) the original trait from which they are calculated, since neither is conceptually nor biologically antecedent to the other. The paths from and to the squared traits are designated as q , to differentiate them from linear paths (p) and correlations (r). To distinguish paths through traits from those through squared traits, we indicate the latter with a prime. For example, the path from z_1 to z_3^2 is $q_{3'1}$. While we do not deal with the calculation of the path coefficients here, we note that the linear and nonlinear models must be estimated separately (Cohen & Cohen, 1983; Lande & Arnold, 1983; Brodie *et al.*, 1995). In a nonlinear (quadratic) model, it is the combination of the variable and its square that is of interest, not either one in

isolation. The linear coefficient in a quadratic analysis does not necessarily equal the coefficient in a linear analysis because of the shared variance between the trait and its squared trait. The linear analysis provides the best estimate of directional selection, so in a nonlinear analysis the linear coefficients are typically ignored.

These new path coefficients through the squared traits measure nonlinear selection, selection on the variance (Lande & Arnold, 1983; Mitchell-Olds & Shaw, 1987). If there is an internal maximum (minimum), then they measure stabilizing (disruptive) selection in the classic sense. Otherwise they represent selection which decreases (increases) the phenotypic variance. In standard linear path analysis these terms would not exist. If there were no internal maximum or minimum, a suitable transformation would be found to linearize the relationship. The case of an internal maximum or minimum is simply ignored in most cases. Linearization is typically justified on the grounds that the researcher is simply trying to create the most useful, predictive model with no epistemological import being given to the exact form of the equations. However in the case of natural selection, the nonlinear terms have real meaning. Thus, we need a method for incorporating them into the analysis.

Path and selection coefficients in this model are calculated as before, but involve a slightly more complicated decomposition of the selection differential because of the more complicated model. Selection directly on quadratic components is assessed simply by the direct and indirect effect coefficients (direct selection) for the squared traits as described above. Thus, direct nonlinear selection on leaf number is $\gamma_3^* = q_{4'3'} + q_{54'} + q_{43'} + p_{54}$. Because there are no nonlinear (or linear) terms connecting leaf number with bolting time, this latter trait does not contribute to direct nonlinear selection on leaf number. Direct nonlinear selection is quantified by the total effect of each of the squared traits. Indirect selection involves the noncausal effect and can be complicated and tedious to calculate. For example, as part of the

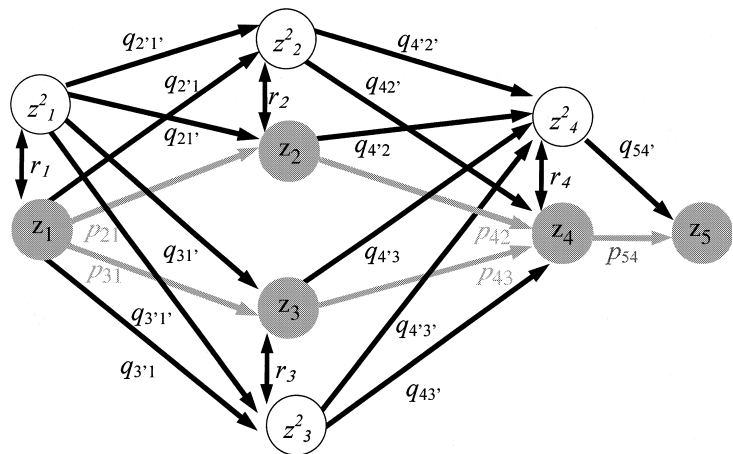


Fig. 2 The same path diagram as in Fig. 1(B), but with nonlinear path coefficients (q) now shown. These nonlinear coefficients are regressions on squares of the trait values and are indicated by squared traits, z^2 terms. To distinguish paths through traits and squared traits, the latter are indicated with a prime. For example, the path from z_1 to z_3^2 is $q_{3'1}$. Variation due to error (U) is not included for simplicity. Conventions follow Fig. 1. Grey is used to denote elements that are also in the model in Fig. 1(B).

calculation of indirect nonlinear selection on seedling size, you would include the path from z_1^2 through z_1 , z_2 and z_4 to z_5 ($r_1 p_{21} p_{42} p_{54}$). In practice, it is easiest to calculate the indirect selection coefficient by subtracting the direct selection coefficient of the trait (β^*) from the model-implied covariance between the trait and fitness (s^*).

In a similar fashion, correlational selection can be incorporated into a path-analytical framework. Estimating the interaction terms for such an analysis requires construction of new traits by multiplying together any pair of traits that have a direct effect on a third trait. The analysis is analogous to that described above, but is beyond the scope of this paper.

Selection on *Arabidopsis thaliana*

Materials and methods

To demonstrate our method, we use data from a study of *A. thaliana* conducted at the University of Tennessee Agricultural Experiment Station in Knoxville, Tennessee. These data are a portion of a 2-year, two-site reciprocal transplant study (Callahan & Pigliucci, unpublished data). The example presented here is not meant to be a definitive analysis of selection in this system. Rather, we examined a deliberately simplified version in order to make the methodology easy to understand. A full analysis would result in the methodological forest being lost among the ecological trees.

In May 1996, seeds were collected from two populations, from 60 plants in a very large population (>500 plants) at the UT Agricultural Experiment Station Plant Sciences Unit in Knox County (83°57'W, 35°56'N), and from 40 in a smaller population (≈150 plants) in Sharps Ridge Park, Knox County (83°56'W, 36°N). Bulk collections of these seeds are available from the Arabidopsis Information Management System stock centre at Ohio State University.

Seed families were maintained in separate envelopes and stored dry until 16 October 1997. We then imbibed the seeds in the dark at 4 °C for 1 week on moist filter paper. From 29 October to 2 November, we randomly planted four seeds per family in each of five 200-cell germination trays (20 seeds per family total), 25 families from each population. Prior to transplanting into the field, seeds were pregerminated on moist vermiculite under artificial light, two 4-foot 40-W fluorescent tubes combined with two 25-W incandescent bulbs, placed 10 cm from the seedlings. We used an 11:13 h photoperiod to approximate field conditions. We added 1 mL of standard Hoagland's solution to each seedling on 10 November, and transferred growth racks from room temperature (20–22 °C) to a 4 °C cold room on 16 November.

Before transplanting to the field we thinned to a single seedling per cell and measured the distance between the apical edges of the cotyledons to estimate seedling size.

On 19–20 November, we transplanted the seedlings into extant vegetation, spaced about 10 cm apart in a randomized array, individually marking each plant. Approximately 80% of the seedlings survived or could be tracked successfully until the following March. Beginning on 27 February 1998 we monitored the plants every other day for time of inflorescence initiation (bolting time) and counted the number of rosette leaves. We monitored all inflorescences until flowering had ceased, harvested inflorescences, and measured inflorescence height and the number of matured fruits. Seed number per plant is strongly correlated with the number of fruits (Westerman & Lawrence, 1970; Mauricio & Mojonnier, 1997).

Estimation of linear and nonlinear selection coefficients

We analysed the data using both linear and quadratic formulations. In each case, we analysed selection using both a multiple regression analysis and a path analysis. We performed the multiple regressions with Procedure REG and the path analyses with Procedure CALIS (SAS Institute, 1989a,b; see also Mitchell, 1993). In the analyses we used data from only the 289 plants that survived to harvest and for which we had records of all five traits in our model (listwise deletion). Because a preliminary analysis showed a minimal effect of the source population, the populations were pooled in the analyses shown here.

Standard linear selection analyses indicate strong selection differentials for all traits, since all four morphological and phenological traits covaried significantly with our measure of fitness, number of fruits (Table 2). The multiple regression indicates that direct selection was strongest on inflorescence height, and was significant for all traits except cotyledon size (Table 2). Direct selection on leaf number was small, while indirect selection was substantially larger. The conclusions from this analysis are that height is the primary trait directly influencing fitness, and that most of the effect of leaf number is due to its correlation with height.

The path analysis generates different conclusions (Fig. 3, Table 2). Now height is the only trait allowed to

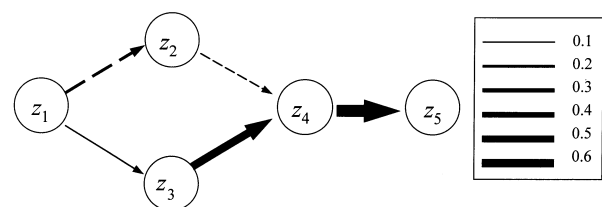


Fig. 3 Solved path diagram for the linear analysis. Dashed lines indicate negative coefficients, and the width of the arrow indicates the strength of the effect. Variation due to error (U) is not included for simplicity.

directly affect fitness, and that effect is very strong. Yet the path analysis indicates that direct selection of leaf number is also considerable, because it strongly affects height (Fig. 3). Thus, though leaf number has no direct effect on fitness, it has a strong indirect effect resulting in direct selection mediated by height. What was a correlated effect in the multiple regression is now part of direct selection because of the change in causal structure. This may seem to be only a nuance, but it may be a better way of examining how the strength of selection on leaf number differed between the two sites and the two years of the study. This is especially true because the correlation between height and leaf number varied considerably, especially in the study's second season, when it was 0.37 at the Agricultural Experiment station site and 0.74 at the other site (Callahan & Pigliucci, unpublished data). In contrast, direct selection of bolting time and seedling size is about equally small in both analyses. Thus, the conclusion from this analysis is that both height and leaf number experience substantial direct directional selection.

Note that s^* for seedling size (the model-implied covariance for seedling size and fitness, calculated as $p_{21} p_{42} p_{54} + p_{31} p_{43} p_{54}$) differs markedly from that observed, an indication that the model fails to capture some of the important ecology for this part of the life cycle. This difference is perhaps unsurprising, since seedling size was both expressed and measured in a laboratory environment, while all of the later traits were expressed and measured in the field. When environmental change is important, it can be taken into account using a modified version of path analysis, as discussed in Scheiner & Callahan (1999). For the other traits, s^* is reasonably close to the observed value (s). As seedling size is the earliest trait expressed, this difference suggests that unmeasured early growth traits might have important fitness effects.

Statistical significance of the selection coefficients cannot be assessed for direct selection, except for traits like height that directly affect fitness. Instead we can

assess the total fit of the model. In this instance, the agreement between model and data is weak (the goodness-of-fit test shows a significant difference between the model and the data: $\chi^2_5 = 30.5$, $P < 0.0001$), indicating that further work is necessary to understand the underlying biology and ecology of this system.

In the quadratic analysis, we are only concerned with the quadratic coefficients and ignore the linear coefficients, as explained above. The selection differentials and the multiple regression show a significant positive quadratic coefficient for height (Table 3), although the multiple regression indicates that much of it is indirect. Because there is no local minimum (determined by inspection, figure not shown), this is selection to increase the variance; classical disruptive selection requires a local minimum (Lande & Arnold, 1983; Mitchell-Olds & Shaw, 1987). Leaf number shows substantial positive selection on its variance when looking at total selection, but this relationship is also due to a large indirect selection component through height. The conclusion from this analysis is that both height and leaf number are under quadratic selection, and most of that is indirect.

The quadratic path analysis shows a similar overall picture, but with some important differences (Fig. 4, Table 3). Height still experiences more indirect than direct quadratic selection, with no evidence for a local minimum for the direct quadratic effect of height (but see Discussion), and seedling size still has little quadratic effect on fitness. However, the quadratic path analysis shows moderate quadratic selection on bolting time and leaf number. Furthermore, leaf number also shows much weaker indirect selection than in the multiple regression. Unlike the linear path analysis, in the quadratic model seedling size and height can experience indirect effects that trace through their respective linear terms. The predicted covariances between fitness and each trait (s^*) deviate only a little from the observed value (s), though the overall agreement of the model and data is poor ($\chi^2_{14} = 54.6$, $P < 0.0001$). Thus, the conclusions from this analysis differ from the multiple regression analysis;

Table 3 Nonlinear selection analyses showing direct and indirect selection coefficients for multiple regression and path analysis models. Effects in bold type are significant at $P < 0.0001$. Statistical significance can be assessed for all traits for total selection and direct selection in the multiple regression model. In the path analysis model, significance can be assessed only for direct selection on height because only this trait is directly connected to fitness. Because this analysis is only concerned with the quadratic coefficients, linear coefficients are not shown. Indirect selection in the path analysis is calculated by subtracting direct selection (total effect) from s^* (the model-implied correlation between the trait and fitness). R^2 for the multiple regression model = 0.792. R^2 values for the path model are: Bolting time = 0.052, Leaf number = 0.033, Height = 0.381, Bolting time² = 0.029, Leaf Number² = 0.016, Height² = 0.212, Fitness = 0.776.

Trait	Multiple regression			Path analysis		
	Total selection (s)	Direct selection (γ)	Indirect selection	Predicted selection (s^*)	Direct selection (γ^*)	Indirect selection
Seedling size ²	-0.051	-0.033	-0.018	-0.028	-0.082	0.054
Bolting time ²	-0.062	0.064	-0.126	-0.017	0.132	-0.149
Leaf number ²	0.346	0.031	0.315	0.296	0.119	0.177
Height ²	0.576	0.194	0.382	0.579	0.185	0.394

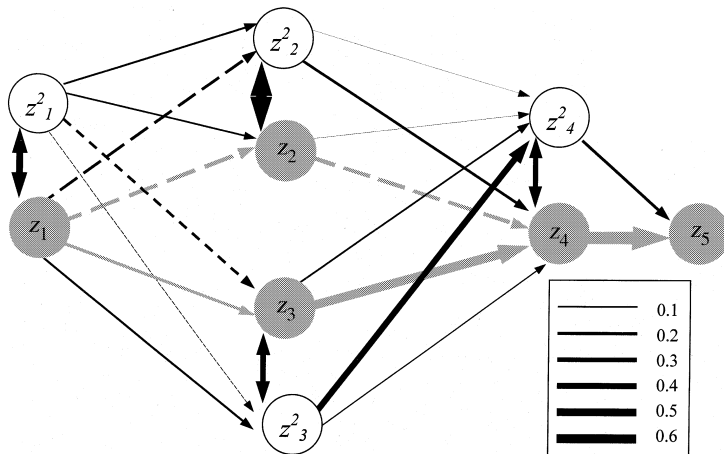


Fig. 4 Solved path diagram for the quadratic analysis. Dashed lines indicate negative coefficients, and the width of the arrow indicates the strength of the effect. Variation due to error (U) is not included for simplicity.

height, bolting time and leaf number are all under quadratic selection and the mix of direct and indirect selection differs for the latter two traits.

Discussion

In this paper we accomplish three goals. First, we show how to estimate selection on traits that are indirectly connected to fitness through intermediate traits. We thus reiterate points made previously (e.g. Arnold, 1983; Crespi & Bookstein, 1989; Mitchell-Olds & Bergelson, 1990; Kingsolver & Schemske, 1991; Conner, 1996; Pugsek & Tomer, 1996). On first glance this estimation appears easy to implement, but actually involves a number of practical complications which we explicate. Second, we demonstrate how a discrepancy between estimates of total selection (s) and the predicted covariance (s^*) provide clues about the possible importance of unmeasured traits. This issue has dogged the use of multiple regression to measure selection since it was first proposed by Lande & Arnold. Our method provides one way to address these concerns. Third, we show how to incorporate nonlinear path coefficients in such an analysis so as to estimate nonlinear selection. To our knowledge, relatively little attention has been paid to such interpretations of the shape of nonlinear components in other fields that make use of path analysis (e.g. psychology, sociology, econometrics), perhaps because only in the context of natural selection do such nonlinear components have ready or informative interpretations.

Analysing selection in a path-analytical framework promises to provide a better understanding of natural selection by more accurately indicating when during the life cycle selection is operating (Arnold, 1983). In our example, the multiple regression analysis indicates that selection was operating only on height, late in the life cycle. However, the path analysis found that leaf number was also under substantial selection. In addition, the discrepancy between s and s^* for seedling size suggests

that other early life cycle traits may be important. Because environments change during an organism's lifetime, such insights may more clearly indicate the ecological factors responsible for selection (Scheiner & Callahan, 1999).

Answers depend on the causal structure

One important difference between path analysis and multiple regression is that the former assumes a specific causal structure among the variables, while the latter assumes that all traits affect fitness directly. The calculated selection coefficients strongly depend on this structure. For example, if we had instead used a model based on a sequential ordering of events (seedling size \rightarrow leaf number \rightarrow bolting time \rightarrow height \rightarrow fitness), direct selection on leaf number would be estimated as -0.0006 instead of 0.510 . However, the fit of such a sequential model to the data is much worse than the model which we used ($\chi^2_6 = 168.2$ vs. $\chi^2_5 = 30.5$), an indication that the sequential model is probably incorrect. Such a model comparison, if decided on *a priori*, would be acceptable, while any such comparison done *a posteriori* would be considered data dredging (see below).

Because of the strong dependence of results on the hypothesized causal structure of the system, it is important that one has confidence in the path model. The model can be built from a combination of logic, biological knowledge, and experiments [but see Shipley (1997) for another approach]. For example, on logical and physiological grounds we hypothesized that the effect of seedling size on height is mediated through effects on leaf number, not through a direct connection. Likewise, experimental manipulations of bolting time or leaf number might suggest similar hypotheses and could be incorporated into the model. Many different models might be feasible for any given system, and path analysis can help in choosing among them. Once one arrives at one or more causal hypotheses, their usefulness can be

assessed in at least three ways: (1) by the logical and biological grounds used to construct the model, (2) by the goodness-of-fit between the model and the data and (3) by experimental tests of the predictions of the model. The utility of these different methods is discussed in detail elsewhere (e.g. Bollen, 1989), but we emphasize the importance of experimental manipulations as a means to confirm the causal linkages embodied in the path diagram. For example, our model predicts that manipulation of leaf number should have strong effects on height and, through those effects, should alter fitness. If one altered leaf number and did not affect height but did affect fitness, it would be worthwhile to re-think the path model.

Several methods can be used to revise the model even when additional data are not available. Such revision, however, is controversial. One school of thought holds that any revision should either be based on *a priori* reasoning or experimental tests of the model assumptions (i.e. causal relationships). An alternative approach involves splitting the data in half (Loehlin, 1987). The first half is used to build the model, modifying as needed to improve fit; then the second half is used to confirm that the good fit is not simply the result of capitalization on chance. This approach requires huge sample sizes, but is usually considered the least biased. A third approach is exploratory analysis (Shipley, 1997) which involves a systematic search for those models that best fit the data. The fourth approach is often referred to as data dredging. The original model is modified using various criteria such as modification indices to improve fit. This process continues until either the best possible fit or an acceptable fit is obtained. This approach is generally frowned upon unless the final model is used only as a hypothesis for future data gathering and experimentation.

Unmeasured traits

Path analysis can provide guidelines about possible unmeasured traits. Unmeasured traits have two meanings in this context. In the criticisms of the Lande & Arnold approach, unmeasured traits are other aspects of the morphology, physiology, behaviour and so forth that were not measured, but which might be under direct selection. Estimates of selection on measured traits might therefore attribute direct selection to the measured traits when in fact the selection was indirect through the unmeasured traits. The second meaning is of traits that are inherently unmeasurable because they are not a trait *per se* but rather a linear combination of measured traits such as an overall size or shape factor. Such traits are typically referred to as latent variables. Here we deal with the former issue; for the latter see Crespi & Bookstein (1989) and Pugesek & Tomer (1996).

In the previous section we discuss the revision of models in the light of a lack of goodness-of-fit. The issue here is similar except that the lack of fit involves the

discrepancy of s and s^* for a particular trait. Now any model revisions can be more focused because a particular part of the model has been identified, while a new study could be sure to measure additional traits associated with that part of the model.

For example, our analysis found a substantial discrepancy for seedling size, but not the other traits (Table 2). Model changes should therefore involve this trait. We could add a path from seedling size directly to height and see if the fit was substantially improved. In fact, structural equation modelling programs provide 'modification indices' that indicate the extent to which adding a new path would improve the fit of model to data. In our case, including a path from seedling size directly to height would decrease the χ^2 value from 30.5 (5 d.f.) to 19.5 (4 d.f.), a significant improvement in the fit of the model. However, this approach suffers from the same drawbacks as the data dredging mentioned above. All of the strategies for model improvement and associated caveats hold as before.

As a further example of how missing traits may be identified, consider the following. Our analyses indicate that number of leaves had an important, but indirect, influence on fitness. Imagine that we had neglected to measure that trait and analysed the data using a path diagram like that in Fig. 1(B), but omitting leaf number. Such an analysis results in a poor fit of model to data ($\chi^2_3 = 20.4$, $P < 0.0001$), and low R^2 values (<4%) for bolting date and height, indicating substantial room for model improvement. The largest modification index involves seedling size and height, indicating that the addition of a path between those traits would improve the χ^2 by 13.8 (1 d.f.), a highly significant improvement in the fit of the model to the data. This might therefore lead investigators to suspect that some intermediary trait may connect seedling size and height. Leaf number is one such trait, but our analyses (Table 2; previous paragraph) indicate that there are probably others. Such data exploration provides a general idea of where to look for missing traits, but gives no indication of which specific trait(s) is(are) missing. That requires biological knowledge and intuition about the functional relationships among traits, and can only be tested with more data.

Path analysis challenges

This change to a path-analytical framework also opens up a new set of challenges. It is currently unclear how to apply selection coefficients estimated from a path analysis to equations for the response to selection. The standard multivariate equations (e.g. $\Delta \mathbf{z} = \mathbf{G}\beta$, Lande & Arnold, 1983) are based on the assumption that total selection on a trait can be decomposed into direct and indirect components and then combined with a genetic covariance matrix. However, the direct and indirect selection coefficients estimated from a path analysis do not necessarily equal total selection for over-identified

models, that is, s and s^* may not be equal. It may be that the coefficients from the path analysis are more informative in this context, but that is an open question. Janzen & Stern (1998) discuss a similar issue with regard to the use of logistic regression to estimate selection coefficients, instead of linear regression as is used here and in typical multiple regression analyses. They recommend using the experimental response to selection as a test of the relative predictabilities of the two regression methods. A similar procedure could be used here.

A further complication is that selection coefficients from a multiple regression and those from a path analysis differ in how they are standardized. The former are standardized for fitness values only, while the latter are also standardized for trait values. The doubly standardized coefficients provide information on responses in standard deviation units, rather than trait-value units. Because the distribution of trait values (standard deviations) can change from one generation to the next, these coefficients are best for comparing selection among traits within a generation, rather than for predicting long-term evolutionary responses.

Additional challenges exist for nonlinear coefficients. First, the Lande–Arnold multiple regression framework of selection requires that quadratic coefficients (e.g. the coefficient for z_1^2 but not z_1z_2) should be divided by two. This is commonly achieved by dividing z_1^2 by two before analysis. How to accommodate this within a path analysis framework is not clear, since standardization to unit variance renders this transformation superfluous, and unstandardized results are difficult to interpret.

Second, it is difficult to determine whether the nonlinear selection coefficients represent stabilizing or disruptive selection (a local maximum or minimum of the response surface). While we can determine if there is a local maximum or minimum for a single direct effect (or for a pair of direct effects) by inspection (Phillips & Arnold, 1989), doing so for a selection coefficient that is a combination of path coefficients involves partial derivatives of total selection equations that can involve many variables (Tate, 1998) and more than three dimensions (see also Brodie *et al.*, 1995). For example, the equation for direct selection on seedling size² is the combination of all direct pathways between seedling size² and fitness:

$$\begin{aligned} \gamma_1^* = & q_{2'1'} q_{4'2'} q_{54'} + q_{2'1'} q_{4'2'} p_{54} + q_{2'1'} q_{4'2} q_{54'} \\ & + q_{2'1'} p_{42} p_{54} + q_{3'1'} q_{4'3} q_{54'} + q_{3'1'} p_{43} p_{54} \\ & + q_{3'1'} q_{4'3'} q_{54'} + q_{3'1'} q_{4'3'} p_{54}. \end{aligned}$$

The indirect selection pathways involve all pathways from the correlation of seedling size² with seedling size (r_1). The partial derivative of this equation is not simple to calculate or interpret. How can we proceed in the absence of a straightforward approach to determining the local slope of these fitness functions? If none of the direct effects shows a local maximum or minimum we might safely conclude that a particular causal pathway does not.

But if multiple local maxima or minima exist, they could either cancel or enhance each other. Even worse, one coefficient might have a local minimum while another has a local maximum. The resultant shape of the curve is likely to be unpredictable and undeterminable. We are not aware of any substantial discussion of this issue in the path analysis literature (Tate, 1998), and this problem therefore remains to be explored and solved.

Conclusions

Natural selection may often act on traits that have sequential or structured causal relationships with one another, and many biological processes have multiple pathways through which they may affect fitness. Path analysis allows one to explicitly incorporate such developmental and temporal complexity in estimates of selection, and therefore should improve our ability to understand natural selection and adaptive evolution. For example, in the analyses shown here conclusions concerning leaf number differed substantially with the inclusion of a direct causal relationship from leaf number to inflorescence height. More complex scenarios are likely to result in even more dramatic shifts in selection coefficients. The now classic example used in the development of methods to estimate selection on multiple traits is that of Bumpus' sparrows, which involves only morphological traits (Lande & Arnold, 1983; Crespi & Bookstein, 1989; Pugesek & Tomer, 1996; Janzen & Stern, 1998). For those data a multiple regression analysis is sufficient because none of the traits is obviously causal on any other. As a result nearly all selection analyses to this point have been done as multiple regressions. We encourage others who have analysed multivariate selection to revisit their data and reconsider its causal structure. Our understanding of patterns of selection may be substantially altered by that endeavour.

Acknowledgments

We thank the staff at Knoxville/Knox County Parks and at the University of Tennessee Agricultural Experiment Station for assistance with preparing and maintaining field sites and B. Stawicki, R. Szmajterowicz, M. Camara and N. Byrd for help with field work and data collection and thank S. Arnold, C. Boake, A. Moore and two anonymous reviewers for comments on the manuscript. We thank R. Hall for patient advice on structural equation modelling and quadratic models. H.S.C. was supported by the University of Tennessee Department of Botany and National Science Foundation grants DEB 9527551 to M. Pigliucci and IBN 9707552 to M.P. and H.S.C.

References

- Arnold, S.J. 1983. Morphology, performance and fitness. *Amer. Zool.* **23**: 347–361.

- Arnold, S.J. & Wade, M.J. 1984. On the measurement of natural and sexual selection: theory. *Evolution* **38**: 709–719.
- Bentler, P.M. 1993. *EQS/Windows: User's guide*. BMDP Statistical Software, Los Angeles.
- Bollen, K.A. 1989. *Structural Equations with Latent Variables*. Wiley, New York.
- Brodie, E.D. III, Moore, A.J. & Janzen, F.J. 1995. Visualizing and quantifying natural selection. *Trends Ecol. Evol.* **10**: 313–318.
- Cohen, J. & Cohen, P. 1983. *Applied Multiple Regression/Correlation Analysis for the Behavioral Sciences*. Lawrence Erlbaum, Hillsdale, California.
- Conner, J.K. 1996. Understanding natural selection: an approach integrating selection gradients, multiplicative fitness components, and path analysis. *Ethol. Ecol. Evol.* **8**: 387–397.
- Crespi, B.J. & Bookstein, F.L. 1989. A path-analytic model for the measurement of selection on morphology. *Evolution* **43**: 18–28.
- Fisher, R.A. 1930. *The Genetical Theory of Natural Selection*. Oxford University Press, Oxford.
- Hayduk, L.A. 1987. *Structural Equation Modeling with LISREL. Essential and Advances*. Johns Hopkins University Press, Baltimore.
- Janzen, F.J. & Stern, H.S. 1998. Logistic regression for empirical studies of multivariate selection. *Evolution* **52**: 1564–1571.
- Jöreskog, K.G. & Sörbom, D. 1988. *LISREL 7. A Guide to the Program and Applications*. Scientific Software, Morrisville, Illinois.
- Kingsolver, J.G. & Schemske, D.G. 1991. Analyzing selection with path analysis. *Trends Ecol. Evol.* **5**: 276–280.
- Koenig, W.D., Albano, S.S. & Dickinson, J.L. 1991. A comparison of methods to partition selection acting via components of fitness: do larger male bullfrogs have greater hatching success. *J. Evol. Biol.* **4**: 309–320.
- Lande, R. & Arnold, S.J. 1983. The measurement of selection on correlated characters. *Evolution* **37**: 1210–1226.
- Li, C.C. 1975. *Path Analysis. A Primer*. The Boxwood Press, Pacific Grove, California.
- Loehlin, J.C. 1987. *Latent variable models*. Lawrence Erlbaum Associates, Hillsdale, New Jersey.
- Mauricio, R. & Mojonner, L.E. 1997. Reducing bias in the measurement of selection. *Trends Ecol. Evol.* **12**: 433–436.
- McCall, C., Waller, D.M. & Mitchell-Olds, T. 1994. Effects of serial inbreeding on fitness components in *Impatiens capensis*. *Evolution* **48**: 818–827.
- Mitchell, R.J. 1993. Path analysis: Pollination. In: *Design and Analysis of Ecological Experiments*. (S. M. Scheiner & J. Gurevitch, eds), pp. 211–231. Chapman & Hall, New York.
- Mitchell-Olds, T. & Shaw, R.G. 1987. Regression analysis of natural selection: Statistical inference and biological interpretation. *Evolution* **41**: 1149–1161.
- Mitchell-Olds, T. & Bergelson, J. 1990. Statistical genetics of an annual plant, *Impatiens capensis*. II. Natural selection. *Genetics* **124**: 416–421.
- Pedhazur, E.J. 1982. *Multiple Regression in Behavioral Research*, 2nd edn. Holt, Rinehart & Winston, New York.
- Phillips, P.C. & Arnold, S.J. 1989. Visualizing multivariate selection. *Evolution* **43**: 1209–1222.
- Pugesek, B.H. & Tomer, A. 1996. The Bumpus house sparrow data: a reanalysis using structural equation models. *Evol. Ecol.* **10**: 387–404.
- Rausher, M.D. 1992. The measurement of selection on quantitative traits: biases due to environmental covariances between traits and fitness. *Evolution* **46**: 616–626.
- SAS Institute Inc. 1989a. *SAS/STAT user's guide, Version 6*, 4th edn., Vol. 1. SAS Institute Inc., Cary, North Carolina.
- SAS Institute Inc. 1989b. *SAS/STAT user's guide, Version 6*, 4th edn., Vol. 2. SAS Institute Inc., Cary, North Carolina.
- Scheiner, S.M. & Callahan, H.S. 1999. Measuring natural selection on phenotypic plasticity. *Evolution* **53**: 1704–1713.
- Schemske, D.W. & Horvitz, C.C. 1988. Plant–animal interactions and fruit production in a neotropical herb: a path analysis. *Ecology* **69**: 1128–1137.
- Shipley, B. 1997. Exploratory path analysis with applications in ecology and evolution. *Am. Nat.* **149**: 1113–1138.
- Tate, R.L. 1998. Effect decomposition in interaction and non-linear models. In: *Interaction and Nonlinear Effects in Structural Equation Modeling*. (R. E. Schumacker & G. A. Marcoulides, eds), pp. 167–181. Lawrence Erlbaum Associates Inc., New Jersey.
- Wade, M.J. & Kalisz, S. 1989. The additive partitioning of selection gradients. *Evolution* **43**: 1567–1569.
- Westerman, J.M. & Lawrence, M.J. 1970. Genotype–environment interaction and developmental regulation in *Arabidopsis thaliana*. I. Inbred lines; Description. *Heredity* **25**: 609–627.
- Wright, S. 1920. The relative importance of heredity and environment in determining the piebald pattern of guinea pigs. *Proc. Nat. Acad. Sci. (U.S.A.)* **6**: 320–332.

Received 9 November 1999; revised 21 December 1999; accepted 17 January 2000