

Soft and hard selection on plant defence traits in *Arabidopsis thaliana*

Joanna L. Kelley,¹ John R. Stinchcombe,^{2*} Cynthia Weinig³
and Johanna Schmitt²

¹Department of Genome Sciences, University of Washington, Seattle, WA 98195,
²Ecology and Evolutionary Biology, Brown University, Box G-W, Providence, RI 02912
and ³Department of Plant Biology, University of Minnesota, 1445 Gortner Avenue,
St. Paul, MN 55108, USA

ABSTRACT

Whether population regulation is global or local determines if hard or soft selection, respectively, is operating. Differences between hard and soft selection, in turn, affect the maintenance of genetic variation and the mean strength of natural selection in heterogeneous environments. Here we investigated whether there were differences in the strength of soft and hard selection on plant defence traits in a spatially replicated field experiment with *Arabidopsis thaliana*. We estimated selection gradients for basal branch number and susceptibility to apical meristem herbivory in 10 spatial environments of an *Arabidopsis* field experiment. Net patterns of soft selection across these 10 spatial environments were calculated as weighted averages of the selection gradients, weighted by the frequency of each environment. Hard selection was estimated in a similar fashion, except that selection gradients for each spatial environment were first multiplied by a fitness ratio that assigned more or less weight to selection gradients depending on whether local mean fitness in that environment was greater or less than global mean fitness across all environments. In general, soft selection was stronger than hard selection, possibly because selection on plant defence traits is relaxed in environments that otherwise lead to high fitness. To evaluate the sensitivity of our results to the frequency of spatial environments, we repeated our analyses under several different hypothetical scenarios in which we varied the cumulative frequency of the environments that led to highest fitness. In these analyses, we found that the mean strength of hard and soft selection, and the relative magnitude of the two estimates, varied greatly depending on the frequency of high fitness environments. These analyses also revealed that estimates of the overall pattern of selection in spatially heterogeneous environments can be very sensitive to the frequency of individual, unusual environments – for example, environments with above average fitness but local patterns of selection dramatically different from other above average fitness environments. We suggest that more empirical effort be devoted to characterizing the frequency of selective environments in natural systems, although that is likely to be extremely challenging.

Keywords: apical meristem damage, frequency of selective environments, genotypic selection analysis, hard selection, herbivory, soft selection, spatial variation, tolerance.

* Author to whom all correspondence should be addressed. e-mail: john_stinchcombe@brown.edu
Consult the copyright statement on the inside front cover for non-commercial copying policies.

INTRODUCTION

Population regulation is an important factor in determining how evolution will occur in spatially heterogeneous environments. At one end of the spectrum, population regulation can be local – for example, plant populations that are regulated within individual patches. At the other end of the spectrum, population regulation can be global – for example, plant populations that are regulated at the forest level containing several suitable patches. These two alternative modes of population regulation lead to two different models of evolution: soft and hard selection (Levene, 1953; Dempster, 1955; Wallace, 1968; Wade, 1985; Gomulkiewicz *et al.*, 2000).

Soft selection is a model of selection in which population regulation is local. Levene (1953) proposed the model of soft selection to predict the maintenance of allelic polymorphisms in a population with multiple niches. Under soft selection, successfully breeding individuals are limited by factors independent of genotype or patch fitness (Levene, 1953; Wallace, 1968; Holsinger and Pacala, 1990). The proportion of gametes contributed to the next generation from each patch is directly proportional to the frequency of the patch in the entire population (Christiansen, 1975). In a model of soft selection, all patches contribute equally to the subsequent generation (Juenger *et al.*, 2000).

In contrast to soft selection, hard selection is a model in which a population is globally regulated (Dempster, 1955), and the maintenance of genetic variation depends on the fitness of individuals within a given patch. Successfully breeding individuals are determined by a strict fitness cut-off, and only the individuals that meet the fitness criterion contribute offspring to the subsequent generation (Wallace, 1968). The proportion of gametes contributed by each patch is dependent on the fitness of individuals in the patch relative to the mean fitness of the population (Christiansen, 1975). Therefore, regardless of the size of the patch, patches with higher fitness contribute substantially more individuals to the subsequent generation under a model of hard selection (Via and Lande, 1985; Gomulkiewicz *et al.*, 2000).

The mode of population regulation determines which model of selection, soft or hard, is appropriate for estimating whether genetic variation can be maintained in a spatially heterogeneous environment, and understanding the evolutionary dynamics of a trait or multiple traits in the population (Gomulkiewicz and Kirkpatrick, 1992). Although some researchers have attempted to predict whether hard or soft selection will differ depending on life-history stage (Holsinger and Pacala, 1990), there are few empirical studies of hard and soft selection on specific phenotypic traits (but see Gomulkiewicz and Kirkpatrick, 1992; Juenger *et al.*, 2000).

There is appreciable evidence to suggest that hard and soft selection are both likely to contribute to the evolution of plant traits (Kelly, 1997). For instance, natural selection on plant traits is frequently heterogeneous on a small spatial scale (Hartgerink and Bazzaz, 1984; Kalisz, 1986; Schmitt and Antonovics, 1986; Antonovics *et al.*, 1987; Stewart and Schoen, 1987; Schmitt and Gamble, 1990; Bell *et al.*, 1991; Stratton, 1994), suggesting that differences between hard and soft selection are likely to be common for many plant traits. In addition, plant populations often exist in patches of varying size and spatial scale, with pollen and seed dispersal ranges that are often larger than the patches that the populations are located in. Such spatial variation in the strength of natural selection, the distribution of plant populations, and the range of seed and pollen dispersal all suggest the importance of hard and soft selection (Kelly, 1997). However, the consequences of hard and soft selection are largely unknown for quantitative traits. In part,

this is because much of the historic literature on hard and soft selection focuses on the maintenance of genetic variation at the allelic level (Levene, 1953; Dempster, 1955; Wallace, 1968). Although the role of hard and soft selection in the evolution of quantitative traits has been well explored theoretically (Via and Lande, 1985; Holsinger and Pacala, 1990; van Tienderen, 1991), few researchers have examined it empirically (but see Gomulkiewicz and Kirkpatrick, 1992; Juenger *et al.*, 2000). Consequently, it is difficult to determine if any general patterns are evident. We therefore asked the following questions: (1) Does selection on plant defence traits in the genetic model organism *Arabidopsis thaliana* differ under models of soft and hard selection? (2) How sensitive are estimates of the strength of soft and hard selection to the frequency of selective environments?

METHODS

Study species

Arabidopsis thaliana (L.) (Brassicaceae) is a self-fertilizing annual plant that is a model organism for genetic and ecological studies (e.g. Mauricio *et al.*, 1997; Bergelson *et al.*, 2001; Kliebenstein *et al.*, 2001; Pigliucci, 2003). In a recent field experiment, we used the Landsberg *erecta* × Columbia recombinant inbred lines developed by Lister and Dean (1993) that are available from the *Arabidopsis* stock centre (www.arabidopsis.org). Because *Arabidopsis* undergoes many generations of selfing after rare outcrossing events, recombinant inbred lines (RILs) may be a close approximation to natural *Arabidopsis* populations (Weinig *et al.*, 2002). The data used in our analysis were originally collected and analysed by Weinig *et al.* (2002, 2003a). As full details of the field methods are given by Weinig *et al.* (2002, 2003a), only a brief summary is presented here.

Experimental design

Seeds from 98 RILs were planted in randomized flats in the Brown University greenhouse. After germination, the seedlings were transplanted to a field plot at Brown's Haffenreffer Reserve, Bristol, Rhode Island. Replicate seedlings from each RIL were sown into 30 randomized blocks in the field. Plantings occurred in the fall and spring to simulate fall and spring germinating cohorts. Replication of each RIL was one per block. To increase the number of individuals per block per RIL, we combined the 30 blocks in the field into ten superblocks for analysis; this was achieved by grouping the three nearest neighbouring blocks in the field into triangular superblocks. Each superblock constitutes a different spatial environment.

During the experiment, the plants experienced a considerable amount of apical meristem damage, which circumstantial evidence suggested was due to rabbit herbivory (e.g. Weinig *et al.*, 2003a,b). We estimated apical meristem damage (AMD) as either present (1) or absent (0) on each individual plant; taking the mean AMD for an RIL (*see below*) estimates the percentage of RIL replicates that received AMD (*see Weinig et al.*, 2003a, for more details). Because the intensity of AMD varied across the plot, we hypothesized that hard and soft selection on plant defence traits may lead to qualitatively or quantitatively different evolutionary dynamics for these traits. We therefore chose to analyse patterns of hard and soft selection on susceptibility to AMD and basal branch number (a tolerance trait). Susceptibility

to AMD increases the likelihood that an individual plant will experience damage, while the number of basal branches on a plant affects a plant's tolerance to herbivore damage (Weinig *et al.*, 2003a,b).

Statistical analysis

Experimentally observed patterns

For the first analysis, we estimated selection on susceptibility to AMD and basal branch number separately for each season and superblock. Before analysis, we calculated the average value of susceptibility to AMD, basal branch number and fruit number for each RIL within each superblock for use in selection analyses. Our goal in utilizing the RIL means within individual superblocks was to reduce the influence of environmental covariances on our estimates of selection (Rausher, 1992; Stinchcombe *et al.*, 2002). We estimated directional selection gradients in each superblock as the partial regression coefficients from a linear regression model with relativized fruit number, which is our estimate of fitness, as the dependent variable, and both susceptibility to AMD and basal branch number as independent variables. Fitness was relativized by dividing by overall mean fitness within each superblock.

Under the model of soft selection, each environment contributes propagules in proportion to the frequency of the environment. For this data set, the environments are the superblocks, which occur with equal frequency of 0.1. To estimate the overall strength of selection on susceptibility to AMD and basal branch number under a model of soft selection, we computed a weighted average of the selection gradients for that trait, weighted by the frequency of the superblock:

$$\beta_s = \sum_i f_i \beta_i \quad (1)$$

where f_i is the frequency of the i th environment and β_i is the selection gradient in the i th environment. In this formulation (e.g. Gomulkiewicz and Kirkpatrick, 1992), β_s is the overall strength of selection on the trait under soft selection. (Note that the sum of the environmental frequencies equals one, i.e. $\sum_i f_i = 1$.)

Under a model of hard selection, the relative fitness of the individuals in each superblock influences the proportion of propagules contributed to the next generation. To estimate the overall strength of selection on susceptibility to AMD and basal branch number under hard selection, we computed a weighted average of the selection gradient by multiplying the selection gradient for an individual environment by the frequency of that environment and by a fitness ratio. The fitness ratio was the mean fitness in the individual superblock divided by the mean overall fitness in the season. Thus, hard selection was estimated as:

$$\beta_h = \sum_i f_i \left(\beta_i \frac{W_i}{\bar{W}} \right) \quad (2)$$

where f_i is the frequency of the i th environment, β_i is the selection gradient in the i th environment, W_i is the mean fitness in the i th environment and \bar{W} is the grand mean of

fitness in the season (Gomulkiewicz and Kirkpatrick, 1992). For both hard and soft selection, we present 95% confidence limits of the net patterns of selection as estimated by 1000 bootstraps.

We present unstandardized selection gradients rather than standardized selection gradients for two reasons. First, our goal is to compare hard and soft selection for the same trait, rather than to compare the strength of natural selection on traits measured in different units. Second, the standard deviations for the traits vary considerably from superblock to superblock. Accordingly, standardizing separately in each superblock can create the impression that the pattern of selection on the traits differs between superblocks, even when the estimated unstandardized selection gradients are similar or even identical. To provide some intuitive understanding of the magnitude of the unstandardized selection gradients, we also present fitness elasticities of susceptibility to AMD and basal branch number – that is, the proportional change in relative fitness caused by a proportional change in the trait (see, for example, Morgan and Schoen, 1997; van Tienderen, 2000; Hansen *et al.*, 2003). Operationally, this amounts to standardizing traits before analysis by dividing them by their means, or multiplying unstandardized selection gradients by the mean of the trait (for more details, see Morgan and Schoen, 1997; van Tienderen, 2000; Hansen *et al.*, 2003). In this instance, we multiplied the selection gradient for each superblock by the mean value of the trait in that superblock.

Sensitivity to environmental frequencies

The calculations described above for soft and hard selection include the frequency superblocks (see equations 1 and 2), which, in our experiment, were all equal to 0.1. However, the natural frequency of spatial and selective environments will rarely, if ever, be evenly distributed. To study the sensitivity of the previous results to different environmental frequencies, we used the following approach. First, for each season, we ranked each superblock by its fitness ratio (W_i/\bar{W}) to determine the best superblocks in each season ('best' in this case defined by superblocks with high relative fruit production within a season); superblocks were ranked without regard to the strength of selection on either trait in that superblock (see Table 1). We then varied the cumulative frequency of the top N superblocks from zero to one systematically (in units of 0.025) and plotted how hard and soft selection would change as these superblocks changed in frequency, for the top $N=2, 3$ and 4 superblocks (as defined by their fitness ratios). In these analyses, the top N environments were constrained to have equal frequency (i.e. the top two superblocks each ranged in frequency from 0 to 50% for $N=2$, the top three superblocks each ranged in frequency from 0 to 33.3% for $N=3$, and the top four superblocks each ranged in frequency from 0 to 25% for $N=4$).

For example, for $N=2$ superblocks, we calculated hard and soft selection when the top two superblocks each ranged from 0 to 50% of the environments in which selection was acting. This procedure was repeated in an iterative fashion, gradually increasing the frequency of each of the top two superblocks from 0% to 2.5% to 5% and so on up to 50% for each superblock. For each of these analyses, the frequency of selective environments not made up by the top two superblocks was evenly distributed among the remaining superblocks (e.g. if the top two superblocks each represented 20% of the selective environments, the remaining eight each represented 7.5% of selective environments; if the top two superblocks each represented 30% of selective environments, the remaining eight each represented 5% of the selective environments, and so on).

Table 1. Estimated selection gradients from each of the ten superblocks for each trait and each season

Super-block	Season	Susceptibility, β (SE)	Basal branch number, β (SE)	Fitness ratio (W_i/\bar{W})	Rank within season (by fitness ratio)
1	F	-0.3513 (0.29)	0.2245* (0.02)	1.669	2nd
2	F	-1.033* (0.31)	-0.0595 (0.14)	0.195	10th
3	F	-0.4762 (0.38)	0.2239* (0.03)	1.577	3rd
4	F	0.00701 (0.42)	0.1259* (0.02)	2.205	1st
5	F	-0.5083* (0.25)	0.3402* (0.05)	0.427	9th
6	F	-0.1634 (0.24)	0.2763* (0.05)	0.520	7th
7	F	0.1257 (0.32)	0.1543* (0.04)	1.005	5th
8	F	-0.3291 (0.42)	0.1724* (0.04)	0.617	6th
9	F	-0.1351 (0.43)	0.3000* (0.02)	1.281	4th
10	F	-0.1819 (0.30)	0.3353* (0.05)	0.503	8th
11	S	0.0534 (0.13)	0.0975* (0.03)	0.992	5th
12	S	-0.3552 (0.32)	0.0926* (0.05)	0.590	8th
13	S	-0.3276 (0.19)	0.1744* (0.04)	0.850	6th
14	S	-0.1284 (0.12)	0.1941* (0.03)	1.606	1st
15	S	-0.4328* (0.22)	0.1464* (0.03)	0.519	9th
16	S	-0.3846 (0.24)	0.0722 (0.04)	0.326	10th
17	S	-0.03933 (0.20)	0.0538 (0.04)	1.509	2nd
18	S	-0.07171 (0.14)	0.2191* (0.03)	1.416	3rd
19	S	-0.2659 (0.22)	0.2159* (0.04)	1.404	4th
20	S	-0.3582* (0.18)	0.1447* (0.03)	0.145	7th

* Significant at the $P < 0.05$ level. SE = standard error.

RESULTS

Experimentally observed patterns

To calculate hard and soft selection for susceptibility to AMD, we used the selection gradients from each superblock (Table 1) regardless of their statistical significance. We chose to retain these selection gradients in our calculations of the net pattern of hard and soft selection (i.e. non-significant selection gradients were not set to zero when calculating net patterns of hard and soft selection), because they represent the best estimate of the actual strength of selection in that superblock (M. Kirkpatrick, personal communication) and we were not testing hypotheses about the pattern of natural selection on these traits within any individual superblock. In general, natural selection acted to reduce susceptibility to AMD and to increase basal branch number, with some variation in the magnitude of selection in each season and superblock. We did not observe any instances of significant reversals in the sign or direction of natural selection across superblocks (i.e. there were no cases of significant selection in favour of a trait in one superblock and significant selection against the same trait in another superblock).

Using equations (1) and (2), the strength of selection on susceptibility to AMD in the fall was -0.217 under a model of hard selection and -0.305 under a model of soft selection (Table 2a), although the bootstrapped 95% confidence limits for these estimates overlapped. The corresponding fitness elasticities of susceptibility to AMD under hard and soft

Table 2. Experimentally observed patterns of the overall strength of selection on susceptibility to apical meristem damage and basal branch number under models of hard and soft selection for the fall and spring cohorts

Trait	Selection	β (95% CL)	Elasticity (95% CL)
A. Fall cohort			
Suscept. AMD	Hard	-0.217 (-0.280, -0.168)	-0.183 (-0.236, -0.133)
Suscept. AMD	Soft	-0.305 (-0.352, -0.224)	-0.223 (-0.260, -0.184)
Basal branch #	Hard	0.210 (0.199, 0.233)	0.745 (0.687, 0.774)
Basal branch #	Soft	0.209 (0.195, 0.239)	0.612 (0.554, 0.684)
B. Spring cohort			
Suscept. AMD	Hard	-0.181 (-0.207, -0.166)	-0.036 (-0.041, -0.030)
Suscept. AMD	Soft	-0.231 (-0.263, -0.209)	-0.038 (-0.043, -0.034)
Basal branch #	Hard	0.151 (0.140, 0.169)	0.334 (0.304, 0.365)
Basal branch #	Soft	0.141 (0.132, 0.150)	0.288 (0.263, 0.315)

Note: The fitness elasticities of these traits are also shown for purposes of comparison. Values were calculated by multiplying the selection gradient or elasticity by the frequency of the superblocks (0.1) and then summing over all superblocks. The 95% confidence limits (CL) estimated from 1000 bootstraps are shown in parentheses after the estimates.

selection were -0.18 and -0.22, respectively, suggesting that a 10% increase in susceptibility to AMD leads to a 1.8 and 2.2% decrease in relative fitness. For the spring, the strength of selection was -0.181 and -0.231 for hard and soft selection, respectively (Table 2b), with non-overlapping 95% confidence limits. These selection gradients correspond to fitness elasticities of -0.0355 and -0.0376, respectively, indicating that a 10% increase in susceptibility to AMD in the spring would lead to a 0.35 and 0.37% decrease in relative fitness. The point estimates for hard and soft selection, as well as the 95% confidence limits for the spring cohort, both suggest that selection on susceptibility to AMD was stronger under a model of soft selection – not only are the estimated overall selection gradients larger, but proportional changes in susceptibility to AMD lead to greater proportional changes in relative fitness under the soft selection model.

Using equations (1) and (2) to estimate the strength of selection on basal branch number under models of soft and hard selection, respectively, we detected similar patterns in the strength of hard and soft selection for the fall cohort. For the fall, the strength of selection under a model of hard selection was 0.210, while under a model of soft selection the strength of selection was 0.209 (Table 2a). These selection gradients correspond to fitness elasticities for basal branch number of 0.75 and 0.61, indicating that proportional changes in basal branch number had dramatic effects on relative fitness (i.e. a 10% change in basal branch number would lead to either a 7.5 or 6.1% increase in relative fitness), and suggesting a trend for the fitness impacts of changes in basal branch number to be greater under a model of hard selection. In the spring cohort, selection on basal branch number was reduced: the strength of selection was 0.151 under a model of hard selection and 0.141 under a model of soft selection (Table 2b). In addition, in the spring cohort, these selection gradients correspond to much lower fitness elasticities – these selection gradients correspond to fitness elasticities of 0.33 and 0.29, both of which indicate lower proportional changes in relative fitness for a given percentage change in basal branch number. For both

the fall and spring cohorts, the trend for greater fitness elasticities under the hard selection model appear to be due to the fact that superblocks with high fitness ratios (i.e. $W_i/\bar{W} > 1$) also had higher means for basal branch number. In general, fitness elasticities for basal branch number were higher, regardless of selection model, than fitness elasticities for susceptibility to AMD, indicating that the proportional effects on relative fitness of changes in basal branch number are greater than changes in susceptibility to AMD. These data are consistent with previous reports of significant overlap between QTL mapped for fitness and basal branch number in this experiment (Weinig *et al.*, 2003a).

Sensitivity to environmental frequencies

Susceptibility to apical meristem damage

The strength of both hard and soft selection on susceptibility to AMD, and the relative difference between the two, depends on the frequency of selective environments (e.g. Fig. 1). For susceptibility to AMD in the fall cohort, the strength of soft selection declines as high fitness superblocks become more and more common. In other words, when superblocks that had high relative fruit production within a season represent a greater proportion of the selective environments, the net strength of soft selection is reduced. The decline in the strength of soft selection is most apparent when the top two superblocks are increasing in

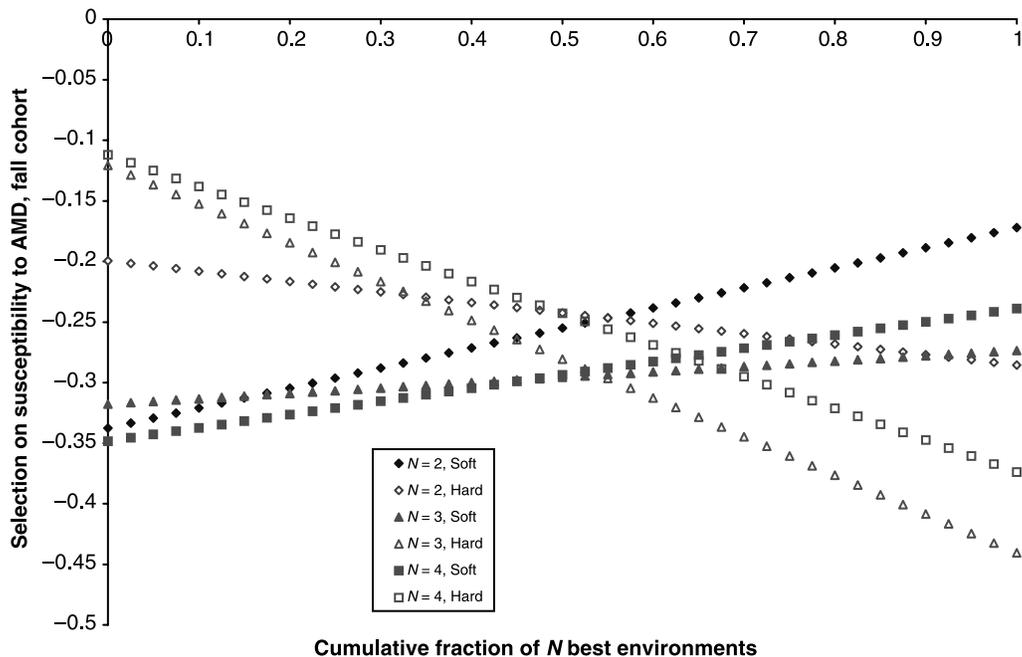


Fig. 1. Plot of hard and soft selection on susceptibility to AMD in the fall cohort, as a function of the cumulative frequency of the top two, three and four superblocks with the highest fitness ratios. Solid symbols represent analyses of soft selection, open symbols hard selection. Diamonds: cumulative frequency of top two superblocks. Triangles: cumulative frequency of top three superblocks. Squares: cumulative frequency of top four superblocks.

frequency (solid diamonds in Fig. 1), whereas when the top three or four superblocks (solid triangles and squares in Fig. 1) increase in frequency the strength of soft selection declines, but not to the same extent. These trends are driven by the fact that natural selection on susceptibility to AMD is, on average, weaker in these superblocks with high fitness ratios than in the other superblocks (see below). The patterns for hard selection on susceptibility to AMD are reversed – when the weighting ratio is added to the analysis, hard selection on susceptibility to AMD becomes appreciably more negative (this result is most apparent with the analyses of the top three and four superblocks). This apparently counter-intuitive pattern is driven entirely by two superblocks (1 and 3 of Table 1) that have the second and third highest fitness ratios and are also the superblocks in which we observed the third and fourth strongest selection gradients on susceptibility to AMD. The trend is less pronounced for the analyses of the top two superblocks, mainly because the superblock with the highest fitness ratio (superblock 4) had the weakest overall pattern of selection on susceptibility to AMD.

For the spring cohort, the strength of soft selection on susceptibility to AMD also declined as high fitness superblocks increased in cumulative frequency (Fig. 2). Similar to the fall cohort, this appears to be due to the strength of selection on AMD being reduced in high fitness superblocks. Under a model of hard selection, these trends also appear to hold – at least for the cases in which the top two and top three superblocks increase in cumulative frequency. However, when the top four superblocks increase in frequency, hard selection

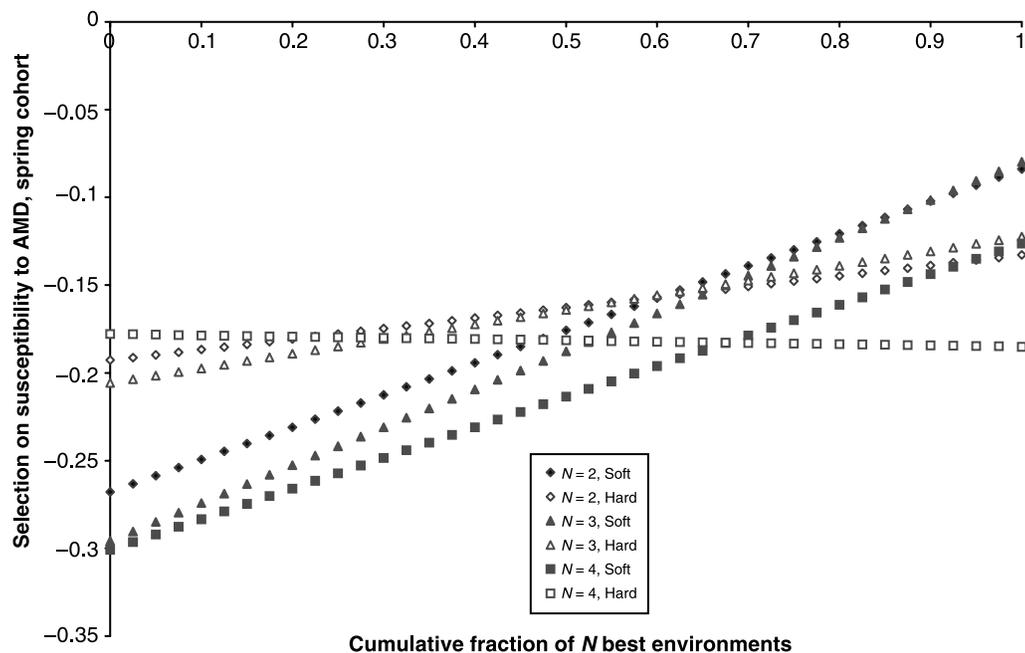


Fig. 2. Plot of hard and soft selection on susceptibility to AMD in the spring cohort, as a function of the cumulative frequency of the top two, three and four superblocks with the highest fitness ratios. Solid symbols represent analyses of soft selection, open symbols hard selection. Diamonds: cumulative frequency of top two superblocks. Triangles: cumulative frequency of top three superblocks. Squares: cumulative frequency of top four superblocks.

on susceptibility to AMD begins to decline slightly, mainly because the superblock with the fourth highest fitness ratio was an environment (superblock 19) in which selection on susceptibility to AMD was appreciably stronger than in the other top three superblocks.

Basal branch number

The patterns we observed for hard and soft selection on basal branch number under variable environmental frequencies were similar in both the fall and spring cohorts (Figs. 3 and 4). In general, under hard selection, the net strength of selection on basal branch number increases in magnitude when high fitness environments increase in frequency (solid symbols in Figs. 3 and 4). This pattern is driven by environments in which selection on basal branch number was moderately strong and positive, receiving added weight due to the fitness ratios that are greater than 1.

The strength of soft selection in the fall cohort appears to be largely insensitive to the cumulative frequency of the top four superblocks (solid squares in Fig. 3), although the strength of soft selection actually declines when only the top two or three superblocks increase in cumulative frequency. The trend for soft selection to decline when the top two or three superblocks increase in frequency is driven by the top superblock (#4), in which selection on basal branch number was approximately half as strong as selection in the second and third best superblocks. A similar pattern is seen for the strength of soft selection

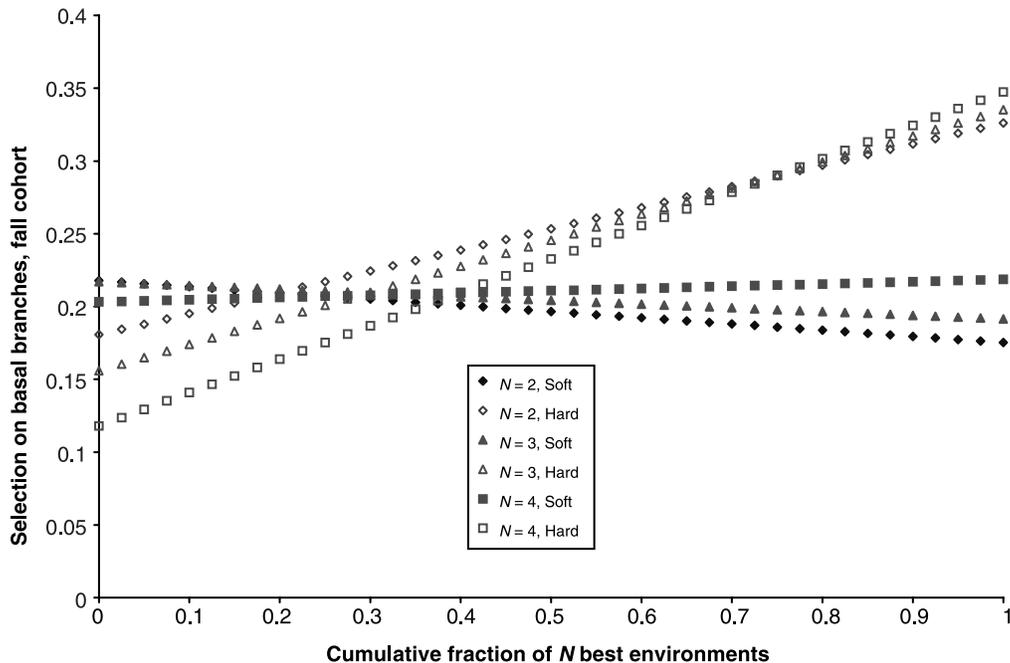


Fig. 3. Plot of hard and soft selection on basal branch number in the fall cohort, as a function of the cumulative frequency of the top two, three and four superblocks with the highest fitness ratios. Solid symbols represent analyses of soft selection, open symbols hard selection. Diamonds: cumulative frequency of top two superblocks. Triangles: cumulative frequency of top three superblocks. Squares: cumulative frequency of top four superblocks.

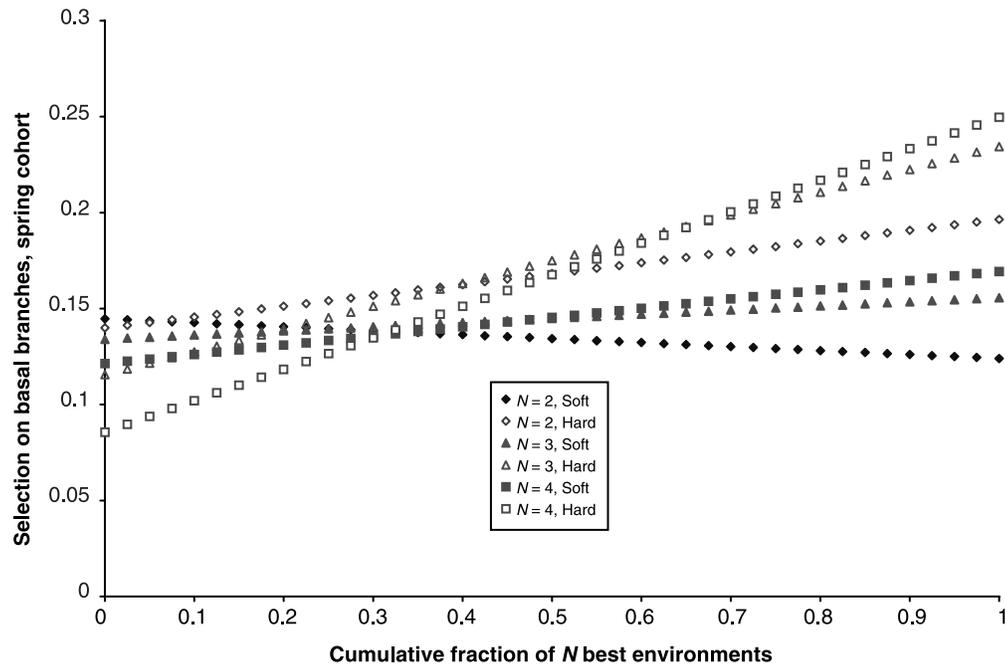


Fig. 4. Plot of hard and soft selection on basal branch number in the spring cohort, as a function of the cumulative frequency of the top two, three and four superblocks with the highest fitness ratios. Solid symbols represent analyses of soft selection, open symbols hard selection. Diamonds: cumulative frequency of top two superblocks. Triangles: cumulative frequency of top three superblocks. Squares: cumulative frequency of top four superblocks.

in the spring cohort – as the cumulative frequency of the top three or four superblocks increases, the strength of soft selection is relatively stable or mildly increasing. However, the strength of soft selection actually declines when the top two superblocks increase in cumulative frequency – in this case, driven by the second ranked superblock (#17) in which selection on basal branch number was approximately one-quarter as strong as it was in the first, third and fourth ranked superblocks.

DISCUSSION

Two major results emerge from our analysis of hard and soft selection on plant defence traits in *Arabidopsis thaliana*: first, in general, it appears that the mean strength of soft selection is greater than the mean strength of hard selection; and second, that the frequency of selective environments has dramatic effects on the strength of hard and soft selection, and the relative difference between the two.

Previous empirical studies of hard and soft selection on quantitative traits have failed to reveal a clear consensus on how commonly these two models of selection result in differences in the mean strength of selection, or evolutionary trajectory of traits. For example, in simulation studies, Gomulkiewicz and Kirkpatrick (1992) determined that differences between hard and soft selection could lead to the evolution of strikingly different

reaction norms of node number in response to light availability in the agricultural weed *Abutilon theophrasti*. In contrast, Juenger *et al.* (2000) failed to detect significant differences between hard and soft selection on node production, an important tolerance trait, in the field gentian (*Gentianella campestris*). Our results for susceptibility to AMD in both the spring and fall cohorts, and basal branch production in the fall cohort, are similar to those of Gomulkiewicz and Kirkpatrick (1992), in that we detected consistent differences between hard and soft selection (in this case, finding that soft selection was of a greater magnitude than hard selection). However, our results for basal branch production in the spring cohort are similar to those of Juenger *et al.* (2000), in that we detected minimal differences between hard and soft selection. Although we did detect a stronger pattern of soft selection than hard selection on basal branch production, these results were only detectable at the thousandth decimal place, a level of precision that is unlikely to be obtainable in most field estimates of selection (see, for example, the standard errors of the selection gradients in Table 1). Further evidence in support of this interpretation comes from our analysis of the fitness elasticities for basal branch number under models of hard and soft selection in the spring cohort: the percent change in relative fitness caused by a fixed percentage change in basal branch production differed by only 0.4% between models of soft and hard selection, also suggesting minimal differences between hard and soft selection.

If soft selection was in fact acting on *Arabidopsis* in our experiment, our finding that soft selection was of greater magnitude than hard selection stands in contrast to the theoretical results presented by Holsinger and Pacala (1990). It is important to note, however, that our results only bear on the consequences of what would occur *if* hard or soft selection were acting, rather than which type of selection was in fact occurring in our experiment. In their study, Holsinger and Pacala (1990) attempted to predict whether soft or hard selection was likely to predominate for certain types of traits, or alleles for those traits. Their findings suggested that soft selection appears to be the appropriate model of selection for traits influencing early life-cycle attributes such as germination and establishment, because germination and establishment are often space-limited life stages and not dependent on overall population fitness. Holsinger and Pacala (1990) also suggested that hard selection controls the evolution of late life-stage traits that depend on the genotypes found within subpopulations, including survival and reproduction. Because both of the traits we studied are expressed later in the life cycle of *Arabidopsis* – for instance, susceptibility to AMD can only be expressed after the differentiation to the apical meristem and the transition to flowering, and basal branches are often only initiated after flowering has occurred on the apical meristem – we were unable to compare the consequences of hard and soft selection on early versus late life-stage traits.

Our finding that the net pattern of selection on susceptibility to AMD was of greater magnitude under a model of soft selection than under a model of hard selection suggests that in good environments that would otherwise lead to higher fitness, resisting herbivory is under weaker selection (Table 2, β column) and has lower proportional effects on fitness (Table 2, elasticity column). Under a model of hard selection, good environments with local mean fitness that is greater than the global mean fitness (i.e. $W_i/\bar{W} > 1$) are assigned more weight in estimating the net pattern of selection. If, in these environments that led to above average fitness, selection was relaxed on plant defence traits, the net pattern of selection on these traits would be reduced under a model of hard versus soft selection because of the additional weighting factor given to selection gradients that are smaller in absolute

magnitude and the decreased weighting given to selection gradients that are larger in absolute magnitude. Our data provide support for this hypothesis. For instance, in the fall cohort, the average selection gradient for susceptibility to AMD in the superblocks with the five highest fitness ratios was -0.17 , while the average selection gradient for susceptibility to AMD in the superblocks with the five lowest fitness ratios was -0.44 . Results from the spring cohort show a similar pattern: the average selection gradient for susceptibility to AMD in the superblocks with the five highest fitness ratios was -0.09 , while the corresponding average of the selection gradients in the superblocks with the five lowest fitness ratios was -0.37 . Thus for both the fall and spring cohorts, it appeared that susceptibility to AMD was under weaker selection in environments that otherwise led to high fitness.

The frequency of selective environments is important for the evolution of plant traits (Via and Lande, 1985; van Tienderen, 1991; Gomulkiewicz and Kirkpatrick, 1992). Yet, the natural distribution of selective environments is largely unknown (but see Weis and Gorman, 1990; Kingsolver *et al.*, 2001; Arnold and Peterson, 2002; Huber *et al.*, 2004), and selective environments are almost certainly not uniformly and evenly distributed like our experimental superblocks. To evaluate the robustness of experimentally observed results, we explored the strength of hard and soft selection under different scenarios in which the cumulative frequency of high fitness environments was altered. The results of these analyses speak to the importance of the frequency of selective environments. For instance, changes in the cumulative frequency of high fitness environments often led to dramatic changes in the magnitude of the net pattern of selection (e.g. Fig. 1, open triangles). Depending on the cumulative frequency of the high fitness environments, we either observed all the possible configurations of the relative strength of the point estimates for hard or soft selection: hard selection being of greater, approximately equal or less magnitude than soft selection.

The sensitivity of our selection estimates to the frequency of environments also illustrates the potential effect of one or two unrepresentative environments on the overall pattern of hard and soft selection. In general, we found that selection on susceptibility to AMD was weaker in high fitness environments (see above), contributing to our overall finding of weaker patterns of hard selection. However, in varying the cumulative frequency of the high fitness environments for the fall cohort, we found a counter-intuitive trend for natural selection on susceptibility to AMD to become stronger (Fig. 1, open symbols). This trend is entirely driven by two superblocks with strong selection on susceptibility to AMD and high fitness ratios, and exists despite the overall pattern in our data for selection on susceptibility to AMD to be *weaker* in environments with high fitness ratios. These results illustrate the imperfect correlation between an environment's quality (i.e. its fitness ratio) and the strength of selection in that environment. In addition, if by chance we had only sampled these regions of the field for this experiment (or a smaller experiment, for example), we might have obtained a dramatically different picture of the nature of selection on susceptibility to AMD. Taken together, our analyses of how selection on susceptibility to AMD and basal branch number changes with the cumulative frequency of high fitness environments suggest that greater empirical effort should be devoted to measuring the frequency of selective environments – despite the considerable logistical, experimental and statistical challenge of doing so.

The spatial variability of the environment will also affect the magnitude of the differences between hard and soft selection. For instance, there will be little or no difference in the calculated strength of selection between models of hard and soft selection in spatially

homogeneous environments in which either the strength of natural selection or mean fitness is constant or similar across environments. However, as the spatial environment becomes more varied and either the strength of selection within each environment or the mean fitness in that environment becomes more heterogeneous, differences between hard and soft selection become more pronounced. For example, the selection gradients for susceptibility to AMD varied from -1.033 to 0.1257 in the fall, whereas the selection gradients all fell within the range of -0.0595 to 0.3402 for basal branch number. Thus our finding that differences between soft and hard selection are more pronounced for susceptibility to AMD than for basal branch number is due in large part to the greater heterogeneity among selection gradients for that trait.

Distinguishing between hard and soft selection statistically is likely to be a challenging endeavour, whether done by bootstrapping (as we did) or by analytical methods. Given spatial variation in estimates of selection, spatial variability in local fitness estimates is the only factor that can produce differences in the point estimates of hard and soft selection (i.e. equations 1 and 2 differ only by the weighting ratio). While such spatial variability in local fitness estimates leads to greater differences in point estimates of hard and soft selection, such variability might paradoxically *reduce* our ability to distinguish between hard and soft selection statistically in many cases. When there is more spatial variability in local fitness estimates, there will be more variability among individual bootstrap estimates of hard and soft selection. One consequence of this is that the 95% confidence limits of hard and soft selection estimates are likely to be larger and more likely to include each other. Nevertheless, difficulty in ascribing a *P*-value to the difference between hard and soft selection statistically does not imply that the distinction is not useful, nor that there is no heuristic value to be gained by examining point estimates of hard and soft selection, and how these change with different frequencies of selective environments.

Despite the theoretical importance of the distinction between hard and soft selection, relatively few empirical studies have sought to evaluate whether qualitatively or quantitatively different outcomes are likely to be produced under a model of hard versus soft selection. Our results suggest that soft selection typically has a greater absolute magnitude than hard selection, which we hypothesize is due to selection on plant defence traits being relaxed in favourable environments. It would be of interest to establish if this pattern is consistent in studies of other traits, and in other systems in which spatial environments are known to vary in a consistent manner (e.g. density of inter-specific competitors or predators, nutrient or water availability, and so on). By measuring natural selection on quantitative traits in multiple spatial environments and in multiple study systems, an assessment of how frequently hard versus soft selection leads to different evolutionary trajectories should be possible.

ACKNOWLEDGEMENTS

We thank Lisa Dorn, Zac German, Anthony Giunta, Marah Gotschik, Kelly Gravuer, Mason Israel and Nolan Kane for extensive field help, and Rod Pacheko at the Haffenreffer Grant in Bristol, RI. We are grateful to the University of Texas Section of Integrative Biology, Tom Juenger and Mark Kirkpatrick for housing and their hospitality during a visit of J.R.S. and J.L.K. to Austin, TX, and to Gordon Fox and an anonymous reviewer for insightful comments and discussion on a previous version of this manuscript. Our work has been supported by NSF grants DEB-0129018 to J. Schmitt and DEB-9976997 to J. Schmitt, T.F.C. Mackay and M.D. Purugganan. J.L.K. was supported by a Brown University UTRA fellowship/ REU fellowship.

REFERENCES

- Antonovics, J., Clay, K. and Schmitt, J. 1987. The measurement of small-scale environmental heterogeneity using clonal transplants of *Anthoxanthum odoratum* and *Danthonia spicata*. *Oecologia*, **71**: 601–607.
- Arnold, S.J. and Peterson, C.R. 2002. A model for optimal reaction norms: the case of the pregnant garter snake and her temperature-sensitive embryos. *Am. Nat.*, **160**: 306–316.
- Bell, G., Lechowicz, M.J. and Schoen, D.J. 1991. The ecology and genetics of fitness in forest plants. III. Environmental variance in natural populations of *Impatiens pallida*. *J. Ecol.*, **79**: 697–713.
- Bergelson, J., Kreitman, M., Stahl, E.A. and Tian, D.C. 2001. Evolutionary dynamics of plant R-genes. *Science*, **292**: 2281–2285.
- Christiansen, F.B. 1975. Hard and soft selection in a subdivided population. *Am. Nat.*, **109**: 11–16.
- Dempster, E.R. 1955. Maintenance of genetic heterogeneity. *Cold Spring Harbor Symp. Quant. Biol.*, **20**: 25–32.
- Gomulkiewicz, R. and Kirkpatrick, M. 1992. Quantitative genetics and the evolution of reaction norms. *Evolution*, **46**: 390–411.
- Gomulkiewicz, R., Thompson, J., Holt, R., Nuismer, S. and Hochberg, M. 2000. Hot spots, cold spots, and the geographic mosaic theory of coevolution. *Am. Nat.*, **156**: 156–174.
- Hansen, T.F., Pelabon, C., Armbruster, W.S. and Carlson, M.L. 2003. Evolvability and genetic constraint in *Dalechampia* blossoms: components of variance and measures of evolvability. *J. Evol. Biol.*, **16**: 754–766.
- Hartgerink, A.P. and Bazzaz, F.A. 1984. Seedling scale environmental heterogeneity influences individual fitness and population structure. *Ecology*, **65**: 198–206.
- Holsinger, K.E. and Pacala, S.W. 1990. Multiple-niche polymorphism in plant populations. *Am. Nat.*, **135**: 301–309.
- Huber, H., Kane, N.C., Heschel, M.S. *et al.* 2004. Frequency and microenvironmental pattern of selection on plastic shade-avoidance traits in a natural population of *Impatiens capensis*. *Am. Nat.*, **163**: 548–563.
- Juenger, T., Lennartsson, T. and Tuomi, J. 2000. The evolution of tolerance to damage in *Gentianella campestris*: natural selection and the quantitative genetics of tolerance. *Evol. Ecol.*, **14**: 393–419.
- Kalisz, S. 1986. Variable selection on the timing of germination in *Collinsia verna* (Scrophulariaceae). *Evolution*, **40**: 479–491.
- Kelly, J.K. 1997. Fitness variation across a subdivided population of the annual plant *Impatiens capensis*. *Evolution*, **51**: 1100–1111.
- Kingsolver, J.G., Gomulkiewicz, R. and Carter, P.A. 2001. Variation, selection and evolution of function-valued traits. *Genetica*, **112**: 87–104.
- Kliebenstein, D.J., Kroymann, J., Brown, P. *et al.* 2001. Genetic control of natural variation in *Arabidopsis* glucosinolate accumulation. *Plant Physiol.*, **126**: 811–825.
- Levene, H. 1953. Genetic equilibrium when more than one ecological niche is available. *Am. Nat.*, **87**: 331–333.
- Lister, C. and Dean, C. 1993. Recombinant inbred lines for mapping RFLP and phenotypic markers in *Arabidopsis thaliana*. *Plant J.*, **4**: 745–750.
- Mauricio, R., Rausher, M.D. and Burdick, D.S. 1997. Variation in the defense strategies of plants: are resistance and tolerance mutually exclusive? *Ecology*, **78**: 1301–1311.
- Morgan, M.T. and Schoen, D.J. 1997. Selection on reproductive characters: floral morphology in *Asclepias syriaca*. *Heredity*, **79**: 433–441.
- Pigliucci, M. 2003. Studying selection in a model system: evolutionary and ecological genetics of flowering time in *Arabidopsis thaliana*. *Ecology*, **84**: 1700–1712.
- Rausher, M.D. 1992. The measurement of selection on quantitative traits: biases due to environmental covariances between traits and fitness. *Evolution*, **46**: 616–620.

- Schmitt, J. and Antonovics, J. 1986. Experimental studies of the evolutionary significance of sexual reproduction IV. Effect of neighbor relatedness and aphid infestation on seedling performance. *Evolution*, **40**: 830–836.
- Schmitt, J. and Gamble, S.E. 1990. The effect of distance from the parental site on offspring performance and inbreeding depression in *Impatiens capensis*: a test of the local adaptation hypothesis. *Evolution*, **44**: 2022–2030.
- Stewart, S.C. and Schoen, D.J. 1987. Pattern of phenotypic viability and fecundity selection in a natural population of *Impatiens pallida*. *Evolution*, **41**: 1290–1301.
- Stinchcombe, J.R., Rutter, M.T., Burdick, D.S. *et al.* 2002. Testing for environmentally induced bias in phenotypic estimates of natural selection: theory and practice. *Am. Nat.*, **160**: 511–523.
- Stratton, D.A. 1994. Genotype-by-environment interactions for fitness of *Erigeron annuus* show fine-scale selective heterogeneity. *Evolution*, **48**: 1607–1618.
- Van Tienderen, P.H. 1991. Evolution of generalists and specialists in spatially heterogeneous environments. *Evolution*, **45**: 1317–1331.
- Van Tienderen, P.H. 2000. Elasticities and the link between demographic and evolutionary dynamics. *Ecology*, **81**: 666–679.
- Via, S. and Lande, R. 1985. Genotype–environment interaction and the evolution of phenotypic plasticity. *Evolution*, **39**: 505–522.
- Wade, M.J. 1985. Soft selection, hard selection, kin selection, and group selection. *Am. Nat.*, **125**: 61–73.
- Wallace, B. 1968. *Topics in Population Genetics*. New York: W.W. Norton.
- Weinig, C., Ungerer, M.C., Dorn, L.A. *et al.* 2002. Novel loci control variation in reproductive timing in *Arabidopsis thaliana* in natural environments. *Genetics*, **162**: 1875–1884.
- Weinig, C., Stinchcombe, J.R. and Schmitt, J. 2003a. QTL architecture of resistance and tolerance traits in *Arabidopsis thaliana* in natural environments. *Mol. Ecol.*, **12**: 1153–1163.
- Weinig, C., Stinchcombe, J.R. and Schmitt, J. 2003b. Evolutionary genetics of resistance and tolerance to natural herbivory in *Arabidopsis thaliana*. *Evolution*, **57**: 1270–1280.
- Weis, A.E. and Gorman, W.L. 1990. Measuring selection on reaction norms: an exploration of the *Eurosta–Solidago* system. *Evolution*, **44**: 820–831.