Evolution, 44(4), 1990, pp. 820–831

MEASURING SELECTION ON REACTION NORMS: AN EXPLORATION OF THE *EUROSTA-SOLIDAGO* SYSTEM

ARTHUR E. WEIS[1] AND WENDY L. GORMAN[2]
Section of Ecology and Evolutionary Biology, Department of Biological Sciences, Northern Illinois University, DeKalb, IL 60115

Abstract.—The sensitivity of genotypic expression to the environment can be depicted as the reaction norm, which is defined as the array of phenotypes produced by a single genotype over a range of environments. We studied selection on reaction norms of the gall-inducing insect *Eurosta solidaginis* (Diptera; Tephritidae), which attacks tall goldenrod *Solidago altissima* (Compositae). Gall size was treated as a component of insect phenotype and attributes of the host plant as environmental influences on gall development. Genetic differences in the response of gall size to plant lag time (the number of days before a plant responds to the gall maker) were examined. Reaction norms for full-sib families of flies were quantified as linear functions; the elevation of the function denoted gall size produced by the family averaged across all plants, and the function's slope denoted family sensitivity to lag time. Expected fitness of each family was regressed over reaction norm elevation and slope to yield selection gradients on these reaction norm parameters.

Directional selection on gall size averaged across environments is four times stronger than selection on sensitivity. Yet, genetic variation for sensitivity contributes more than twice as much to gall phenotypic variance as family mean gall size. Our results suggest that selection on environmental sensitivity will be weak for populations restricted to a narrow segment of an environmental gradient, but strong for broadly distributed species.

Received October 21, 1988. Accepted November 20, 1989.

Even in the most ecologically specialized species, not all individuals experience the same environmental conditions during development. The influence of pervasive environmental variability on development is a major source of phenotypic variance. Goldschmidt (1940), Schmalhausen (1949), and others recognized that genetic contributions to the phenotype are conditional on the environmental background. For instance, genes that code for large size in fact code for large size given the input of sufficient nutrients, favorable climate, and any other environmental factors. Because of its sensitivity to the environment, a genotype can be characterized by its norm of reaction, that is, the array of phenotypes it produces under different conditions. Two components of the reaction norm can be defined: the average of the phenotypes produced over the range of environments that a population can expect to encounter, and the deviations from this average due to developmental sensitivity to local environments.

As attributes of genotypes, reaction norms are subject to natural selection (Schmalhausen, 1949; Waddington, 1960; Bradshaw, 1965; Via and Lande, 1985; Schlicting, 1986; Lynch and Gabriel, 1987; Scheiner and Lyman, 1989). Yet, measuring selection on reaction norms is difficult because they are not properties of individuals—an individual has a single phenotype and not an array of phenotypes. However, reaction norms can be measured by rearing genotypic replicates, members of an inbred strain, or siblings, across a range of values for some relevant environmental factor (e.g., Clausen et al., 1958; Gupta and Lewontin, 1982; Via, 1984*a*). Because they are not properties of individuals, the intensity of selection on reaction norms cannot be directly evaluated from measures of individual selection. Instead, selection must be evaluated by the differential survivorship and fecundity of genetically related groups (i.e., clones, inbred strains, sibships, etc.). Thus the intensity of selection on reaction norms is not measured in terms of phenotypic values, but in genotypic or breeding values (Lynch and Gabriel, 1987; Rausher and Simms, 1989). To our knowledge, no empirical study to date has quantified the structure of the selection regime on reaction norms in a natural pop-

[1] Present address: Department of Ecology and Evolutionary Biology, University of California, Irvine, CA 92717.

[2] Present address: Department of Virology and Molecular Biology, St. Jude Children's Research Hospital, Memphis, TN 38101.

ulation. We present the results of such a study on the reaction norm for gall size induced by the insect *Eurosta solidaginis* on the rhizomatous herb, tall goldenrod, *Solidago altissima.*

The Eurosta-Solidago *System*

Gall insects induce tumor-like growths on their host plant; the induced gall provides food and shelter to the insect, but confers no benefit to the plant (Weis et al., 1988). Gall characters, such as size, can be considered components of the insect's extended phenotype (Dawkins, 1982; Weis and Abrahamson, 1986) since the insect provides a physiological stimulus that elicits a developmental response in the host plant. Weis and Abrahamson (1986) showed that phenotypic variance in the goldenrod gall size can be traced to genetic variance in the *Eurosta* population. When the gall is viewed as a phenotype of the insect, properties of the host plant can be treated as elements of the environment that influence the phenotypic expression of gall insect genotypes. For instance, *Eurosta* gall size declines as a function of plant age at time of gall induction (Weis and Abrahamson, 1985); thus plant age is an environmental influence when gall development is viewed from the insect's perspective.

Natural enemy attack imposes directional and stabilizing selection on *Eurosta* for gall size (Weis and Abrahamson, 1986). When galls are small (<20 mm) gall inducers are vulnerable to attack by the parasitoid wasp *Eurytoma gigantea* (Hymenoptera; Eurytomidae); the thin walls of small galls allow the wasp's ovipositor to reach the gall inducer (Weis et al., 1985). On the other hand, large galls (>22 mm) are attractive to insectivorous birds, such as the downy woodpecker and the black-capped chickadee (Cane and Kurczewski, 1976). Since wasp attack is generally more intense than bird attack, there is a net directional selection pressure favoring increased gall size (Abrahamson et al., 1989).

Methods

Selection on Reaction Norms

Using the *Eurosta-Solidago* system, we ask, how does the strength of selection on a trait as expressed in the average environment compare to selection on sensitivity to environmental variation? In this section we outline the relationship of individual phenotypic selection to selection on reaction norms, and so outline the logic that underlies our analysis of the *Eurosta-Solidago* system.

For the purpose of illustration, start with the assumption that a species occupies a "coarse grained" environment, i.e., any individual occupies only one point along the environmental gradient during development. For simplicity, also assume that selection favors the same intermediate phenotypic value in all environments, i.e., selection is canalizing.

To understand selection on reaction norms in a population three basic functions must be considered:

(1) the population's frequency distribution of reaction norms vis a vis the relevant environmental gradient,
(2) the function that relates phenotype to fitness,
(3) the distribution of the population across the relevant environmental gradient.

As with phenotypic selection (Endler, 1986), we are asking how differential mortality and fecundity can change the within-generation distribution of traits, but in this case the traits are reaction norms.

What is meant by the distribution of reaction norms? This is easily understood if reaction norms are defined as algebraic functions, such that

$$P_i = q(e \mid g_i). \qquad (1)$$

The phenotype, P_i, produced by individuals of a given genotype, g_i, varies as a function, q, of the environment experienced during development, e. The parameters of such a function would incorporate information on the position, slope, shape, curvature, etc., of the reaction norm. For instance, if reaction norms are linear, they can be characterized by their slope and elevation. The distribution of reaction norms can then be considered as the multivariate distribution of reaction norm parameters. When reaction norms are quantified in this way, selection on the different parameters of the reaction norm can be evaluated separately (e.g., Lynch and Gabriel, 1987).

Phenotypic selection occurs when indi-

vidual differences in phenotype cause fitness differences. The relation of fitness to phenotype can be depicted as a function, $W = t(P)$, where W denotes absolute fitness. By substitution of the phenotypic reaction norm function, Eq. (1), into the fitness function, it is shown that the fitness of individuals with genotype g_i, varies as a function, t, of their phenotypes, which in turn is a function of the developmental environment, such that

$$W_i = t(q(e|g_i)). \tag{2}$$

This expression is the fitness reaction norm. Since fitness depends on the developmental environment, the expected fitness of an individual will depend on which environments it can expect to encounter such that

$$E[W_i] = \sum_{j=0}^{n} w_{ij} f_j \tag{3}$$

where $E[W_i]$ is the expected absolute fitness of individuals of the i-th genotype, w_{ij} is the fitness of the i-th genotype in the j-th environment, and f_j is the proportional frequency of the j-th level of the developmental environment. This expression denotes the absolute fitness that can be expected if an individual of genotype g_i is placed at random within the environment.

Figure 1 illustrates the relation of reaction norm parameters to expected fitness for genotypes A and B. The reaction norms (phenotype vs. environment) are linear (Fig. 1a), and thus can be characterized by their elevation and slope parameters. Elevation can be defined as the predicted value of the phenotype at an appropriate intermediate point on the environmental gradient, such as the mean or median inhabited environment. Under this definition the elevation indicates the expected phenotype under average conditions as well as the mean phenotype across all conditions. (The y-intercept also describes elevation, but has no biological interpretation.) The slope denotes the developmental sensitivity of the genotypes to variation in the environment. Absolute fitness is assumed to vary as a parabolic function of the phenotype. The corresponding fitness reaction norms (fitness vs. environment) for A and B are derived by substituting the phenotypic reaction norm equation into the fitness function (fitness vs. phenotype). Since the same intermediate phenotype is favored at all points of the environmental gradient, the steep slope of genotype A's reaction norm means that it has low fitness at extreme ends of the environmental gradient, where it produces extreme phenotypes, but higher fitness in intermediate environments. Since the reaction norm of genotype B has zero slope (i.e., no sensitivity to environmental variation) it has the same fitness at all points along the gradient.

The expected absolute fitness of each genotype is calculated by multiplying its fitness reaction norm (fitness vs. environment) by the frequency distribution of the environment (frequency vs. environment) and summing over all levels of the environment (Fig. 1b). For this example, the reaction norm with the highest expected fitness would be one with zero slope and with an elevation equal to the optimal phenotype. Of the two genotypes, A has the optimal elevation and B has the optimal slope. Figure 1b shows that the expected fitness for genotype A is greater than for genotype B. This is because reaction norm B falls below the optimal phenotype at all points along the environmental gradient. On the other hand, genotype A shows an optimal phenotype at the modal environment; the very low fitness phenotypes at the extreme points on the gradient have a small impact on fitness because these environmental conditions are unlikely to be encountered.

Although we chose a case of canalizing selection to illustrate selection on reaction norms, the same logic applies to cases where phenotypic plasticity is favored. In that case, the fitness function is extended to be a function of the environment and the phenotype. The optimal reaction norm is defined in this case as that array of phenotypes that yields the highest fitness at all points on the environmental gradient, and the fitness reaction norms become functions of the environment-specific deviations of the phenotypic reaction norm from the optimal (Weis, unpubl.).

If reaction norm parameters and expected fitness can be independently measured, then hypotheses concerning the intensity of natural selection acting on the reaction norm distribution can be tested statistically. The

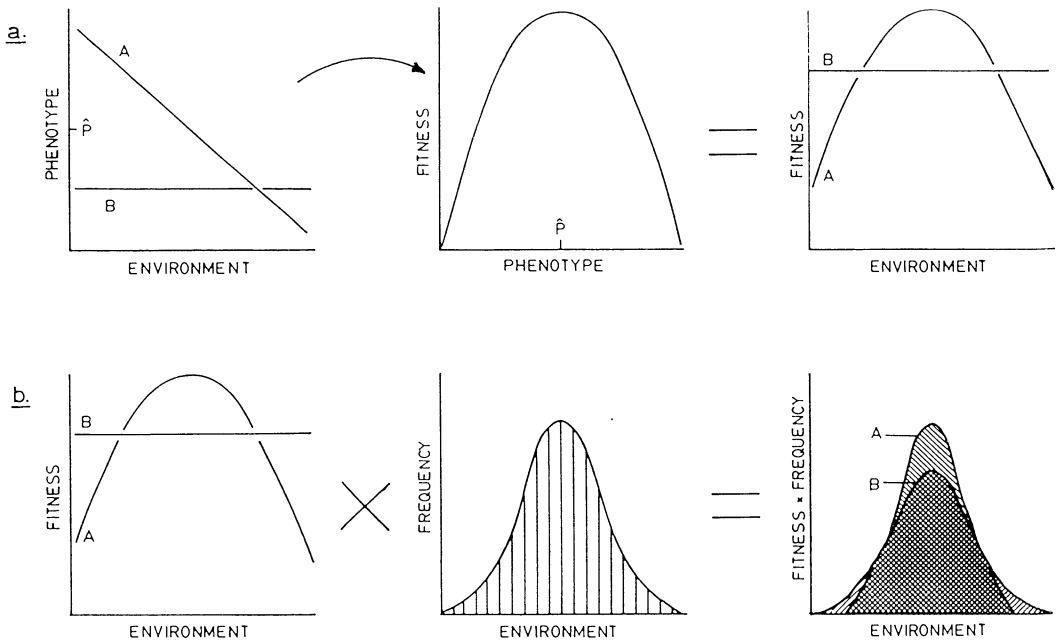

FIG. 1. Calculation of expected relative fitness for reaction norms. Genotype *A* produces extreme phenotypes at the extremes of the environmental gradient, but an optimal intermediate phenotype (*P*) at the environmental mode. Genotype *B* produces the same intermediate, but suboptimal phenotype at all points along the environmental gradient. a) Fitness reaction norms (i.e., fitness vs. environment) are calculated by substituting the phenotypic reaction norm equation into the phenotype term of the fitness function. b) Expected fitness, i.e., the mean fitness if individuals with a given genotype are placed randomly along an environmental gradient, is calculated by multiplying the fitness reaction norms by the frequency distribution of the environment, and summing over all environmental values (Eq. 2).

fitness regression method of Lande and Arnold (1983) can estimate the gradient of selection acting on reaction norm parameters by regressing relative expected fitness over the reaction norm parameters. An important distinction must be made between the phenotypic and reaction norm cases. Phenotypic selection gradients measure within-generation changes in the distribution of phenotypic values (Lande and Arnold, 1983; Endler, 1986). By contrast, coefficients for the regression of expected fitness on reaction norm parameters measure the within-generation change in the distribution of genetic components that contribute to the phenotype. Thus selection on genotypic values is measured directly (see Rausher and Simms, 1989).

Selection on Gall Size Reaction Norms

The Fitness Function.—The absolute fitness function for gall size was estimated from field data on gall size-specific survivorship of field collected *Eurosta.* In April of 1986, more than 3,500 galls were randomly sampled from 16 old-fields in Montour and Union counties, Pennsylvania. Each gall was measured to the nearest millimeter, dissected, and the fate of the gallmaker determined (see Weis and Abrahamson, 1986). The non-linear regression program from the SAS package (SAS Institute, Inc., 1985) was used to fit the Gaussian model

$$p = a + b \exp -0.5((d - D)C^{-1})^2 \quad (4)$$

where p is the probability of survival, d is the observed gall diameter, D is the diameter that gives peak survivorship, C is the width of the survivorship peak, a is the base survivorship (i.e., survivorship independent of gall size) and b is the height of the survivorship peak. The a and b parameters give this formulation greater flexibility than other formulations of Gaussian fitness functions; the a parameter accommodates non-selective survivorship (cf. Cavalli-Sforza and

Bodmer, 1971) and b allows the peak survivorship to be independent from the function's width (cf. Lynch and Gabriel, 1987).

Reaction Norms and the Distribution Along the Environmental Gradient

Reaction norms of gall diameter with respect to plant characters were estimated in a quantitative genetic experiment. Full-sib families ($N = 16$) were grown on a genetically diverse array of *S. altissima* in a greenhouse environment. The use of full-sibships has some drawbacks. The between family variance in this design consists of ½ of the additive genetic variance, ¼ of the dominance genetic variance, and ¼ of the variance due to "additive × additive" interactions, plus the variance due to maternal effects (Falconer, 1981). Thus neither the total genetic variance, which is of interest in estimating true reaction norms, nor the additive variance, which responds to a selection pressure, is cleanly estimable. The most serious problem with this design would occur if non-genetic maternal effects account for a large proportion of the phenotypic variance. Alternate designs that would allow better estimates of the genetic parameters, such as cloning flies or inseminating multiple dams in a half-sib design, are beyond the constraints of this species. However, we feel the full-sib design gives a sufficiently good approximation of reaction norms so as to allow at least a qualitative exploration of selection.

Galls were induced in the greenhouse on plants that were vegetatively propagated from individuals reported on in Weis et al. (1987). These plants were the offspring of hand pollinated crosses among the goldenrods at the Bucknell-Chillisquaque Creek Natural Area, Montour Co., Pennsylvania. The parents for these crosses were selected at random, and thus their offspring represent a random sample from the host plant gene pool. We used 72 plant genotypes (eight from each of nine crosses). Six replicates were propagated from each individual genotype, to give a total of 432 plants. Each plant was grown from a 2–4 g piece of rhizome in a 17 cm plastic pot filled with soilless mix, and was fertilized every two weeks with Hoaglands solution.

The full-sib insect families that induced

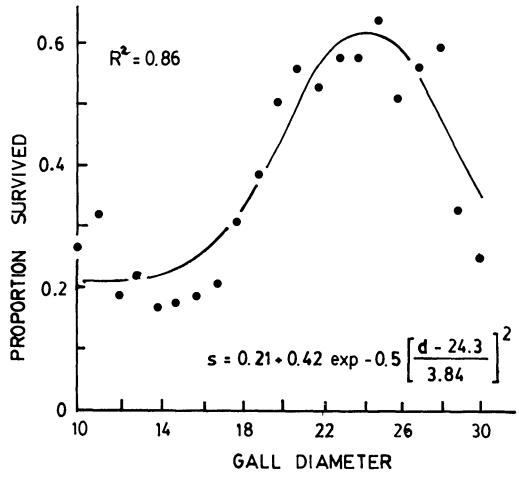

FIG. 2. Fitness function for *Eurosta* gall diameter. Each point represents the proportion of gall inducers in a size class to survive under field conditions.

the galls were the offspring of parental flies reared from overwintered galls collected in the same field as the plants. Singly mated females were supplied with a randomly chosen plant. A fresh plant was substituted as soon as an oviposition scar was seen, or after one hour. All plants were subsequently arranged in a random array on the greenhouse benches. Oviposition spanned three days. Not all plants with oviposition scars produced mature galls; some plants were probably rejected; in other cases early larval death caused failure of gall induction or early cessation of gall growth. Of the 417 plants punctured, 172 produced galls with mature larvae. Only families with five or more mature larvae were analyzed. This gave a total of 135 galls distributed among 16 families with a range of 5 to 16 ($\bar{x} = 8.3$) offspring per family.

Plants were censused daily for gall initiation. Gall diameters were measured 27 days after initiation, when they had achieved full size. Because greenhouse galls are smaller than galls growing in the field, a 2-mm correction factor was added to the day 27 diameter to facilitate subsequent analyses involving the field derived fitness function. Stem height was measured weekly and stem diameter (at 6 cm above soil line) monthly.

As stated above, when gall size is viewed as an insect trait, continuous plant characters can be considered as environmental

gradients that may influence gall development. Forty plant measurements, which included stem dimensions and growth rates at several points of development, initial rhizome weight, and the lag time between egg deposition and gall initiation, were available for analysis. We used two methods to find plant characters that influenced gall size. First, forward selection in a stepwise regression routine showed that one character, plant lag time (days between oviposition and gall initiation), significantly influenced gall size. Second, Principal Component Analysis was used to reduce the data to underlying components; four principal components (PCs) with eigen values >1 were found. Gall size showed a significant regression over the fourth PC only; lag time had a very high loading on this PC (0.93) while all other plant characters had very low loadings ($<|0.15|$). These analyses indicated two additional weak influences on gall size. Stepwise regression showed stem height growth rate during the week of fastest gall growth had a marginally significant effect ($P = 0.09$) on final gall size. Multiple regression of gall size over the PCs showed that the third PC, which defined stem diameter, also had a marginally significant effect ($P = 0.07$). These were not included in the final analysis because they showed no significant influence on gall size in the analysis of covariance comparing the insect sibships (see below).

Intuitively, lag time would seem an insect trait. This character is of necessity the result of a physiological interaction of the insect's stimulus and the plant's reactivity. Plant lag time can be considered an element of the insect's environment if it can be demonstrated that insect genotypic variation makes no contribution to lag time variance. One-way ANOVA showed no significant differences among the insect full-sibships for lag time ($F_{15, 120} = 0.90$). Conversely, we found a plant genetic component for lag time; a nested analysis of variance showed significant differences among plant sibships ($F_{8, 9} = 7.51$; $P = 0.008$). Although there is no evidence of an additive genetic effect on lag time, it is possible that lag time variance has a significant component due to "insect genotype × plant phenotype" interaction effects. The available data cannot eliminate this possibility, but we can show that insect genetic contributions to gall diameter are uncorrelated with potential contributions to such an interaction effect; the correlation between family mean lag time and family mean gall size, which can be used as an estimate of the genetic correlation (Via, 1984*b*), was -0.09 ($P = 0.73$). If there is an "insect genotype–plant phenotype" interaction effect on lag time, the lack of a genetic correlation of lag time with gall diameter means that lag time is no more than part of the "genetic background" in which the loci contributing directly to gall size are selected. We feel these initial analyses justified the use of lag time as an environmental gradient and not an insect trait.

Reaction norms were estimated for each family as the regression of gall diameter of lag time. The analysis assumes that the reaction norms are linear over the range of the environmental gradient encountered. To test this assumption, we performed parabolic regressions of gall diameter over plant lag time, but using the Bonferoni method for simultaneous tests of regression coefficients (Morrison, 1983) the quadratic term was not significant in any of the families.

We used analysis of covariance (ANCOVA) to determine if reaction norms of gall diameter with respect to lag time differed among insect families. Data were tested with the GLM procedure of the SAS package (SAS Institute, Inc., 1985) against the mixed model ANCOVA described by Henderson (1982)

$$d_{ij} = \mu + F_i + \beta + \beta_i + \epsilon_{ij} \qquad (5)$$

where μ is the mean diameter, F_i is the deviation due to the random effect of the i-th insect family, β is the fixed effect of the regression of diameter over plant lag time (i.e., the mean slope), β_i is the deviation due to the random effect of the regression of diameter over plant lag time for the i-th family (i.e, the family-specific slope), and ϵ_{ij} is deviation due to random error in the j-th member of the i-th family. Type IV sums of squares were used. A significant family term would indicate genetic differences in gall size across lag times, i.e., differences in reaction norm elevation. A significant family-specific slope term would indicate that families differ in the linear

component of their response to lag time; this term is a major component of the "genotype × environment" interaction term (Falconer, 1981). We did not find higher order components for the response to lag time (see above). We computed the variance components caused by the terms in the ANCOVA model by maximum likelihood procedures in the P3V program of the BMDP package (Jennrich and Sampson, 1983). The heritability of reaction norm elevation and slope are equal to twice the proportion of phenotypic variance explained by the family variance and the family-specific slope variance components respectively (Scheiner and Lyman, 1989).

The design reflects a more or less natural situation in which female flies encounter a random array of plant genotypes for oviposition. Since many of the plant characters that were likely to influence gall growth could not be measured until after gall initiation, a priori cross classification was not possible. Most importantly, the design we used allowed simultaneous estimation of the frequency distribution of these plant characters, and without these frequency distributions, selection on reaction norms parameters could not be evaluated.

Evaluating Selection on Reaction Norm Parameters

To estimate selection, fitness reaction norms were derived by substituting the gall size reaction norm functions into the d term on the fitness function (Eq. 4). From the fitness reaction norms, environment-wide expected fitnesses were calculated by multiplying specific fitness for each lag time by the observed frequency of that lag time, and summing across all lag times (see Eq. 3). This yields the fitness that insect family members would have realized on average if distributed at random among the plants.

Multiple regression was used to evaluate the intensity of natural selection acting on reaction norm parameters. The relative expected fitness for the family reaction norms was the dependent variable, and the parameters for the family reaction norms were the independent variables, i.e., as the characters being selected. This method is adapted from the approach used by Lande and Arnold (1983; see also Endler, 1986; Mitchell-Olds and Shaw, 1987; Rausher and Simms, 1989).

Directional selection on reaction norm parameters was evaluated by linear multiple regression on reaction norm elevation and slope. We used the model

$$w_i = a + \beta_l L_i + \beta_s S_i \qquad (6)$$

where w_i is the relative expected fitness for members of the i-th family, a is a constant, L_i is the elevation, S_i is the slope, and β_l and β_s are the partial regression coefficients of reaction norm elevation and slope respectively. Stabilizing selection was evaluated by quadratic regression, using the model

$$w_i = a' + \beta'_l L_i + \beta'_s S_i + \gamma_l (L^2_i/2) + \gamma_s (S^2_i/2) + \gamma_{ls} L_i S_i \qquad (7)$$

where w_i is the relative expected fitness, γ_l and γ_s are the quadratic partial regression coefficients of elevation and slope. When the independent variables are skewed, the linear coefficients in this model are biased (Mitchell-Olds and Shaw, 1987), so that directional selection is evaluated through the linear regression only. In addition, the quadratic regression included the term γ_{ls} for the cross products of the elevation and slope, which measured selection on the elevation-slope combination. A significant quadratic term for a trait does not necessarily indicate stabilizing selection, i.e., that an intermediate phenotype is optimal (Mitchell-Olds and Shaw, 1987; Schluter, 1988). To confirm stabilizing selection, partial regression plots were drawn to determine if fitness peaked at intermediate values of the reaction norm parameters. Regression coefficients were standardized to facilitate comparisons between reaction norm parameters with different variances (Lande and Arnold, 1983). To standardize, the linear regression coefficients for elevation and slope were multiplied by the respective standard deviations of each parameter (Lande and Arnold, 1983). Similarly, the quadratic regression coefficients were standardized by multiplying by their respective variances, and the correlational coefficient was standardized by multiplying by the product of the two standard deviations.

The regression analysis provided a measure of the intensity of natural selection, but violation of an important regression as-

TABLE 1. Analysis of covariance of gall size, by insect family and plant lag time.

Source	Mean squares	Sum of squares	*d.f.*	*F*-ratio	*F*	Variance component (std. error)	Proportion variance
Family	(M4)	344.95	15	M4/M1	2.14*	1.688 (1.460)	0.103
Lag time	(M3)	258.70	1	M3/M2	10.02***	1.082 (0.191)	0.066
Fam-lag	(M2)	387.02	15	M2/M1	2.40**	4.395 (2.048)	0.269
Error	(M1)	1,107.77	103	—	—	9.156 (1.585)	0.561

* $P < 0.05$.
** $P < 0.01$.
*** $P < 0.001$.

sumption prevented us from testing the significance of the coefficients. The only way to estimate the relative expected fitness in this case was to substitute the reaction norm regressions into the fitness function, then sum over environments. Thus errors made in estimating reaction norm parameters (the independent variables) are also included in the errors in estimating the expected fitness (the dependent variable). This non-independence means our results cannot be interpreted as a test of the hypothesis that the selection gradients acting on slope and elevation are different from zero. Nonetheless, the contributions of reaction norm parameters to fitness can be compared informally.

RESULTS

The Fitness Function

Field data indicate that from *Eurosta*'s perspective, gall size is under directional and stabilizing selection (Fig. 2). Eighty-six percent of the variance in size-specific survivorship was explained by the Gaussian fitness function. When this fitness function was applied to the distribution of gall sizes from the quantitative genetic experiment, it yielded a directional selection differential (Falconer, 1981) of 1.15 mm, which is 27.4% of the phenotypic standard deviation. The stabilizing selection differential (Endler, 1986) is equal to −3.88, or a 22% reduction in the phenotypic variance. If gall size were highly heritable to the insect, the directional selection differential would lead to rapid evolution.

Reaction Norms and the Distribution Along the Environmental Gradient

In the quantitative genetic experiment, we found significant variance among families for gall size across all lag times (i.e., reaction norm elevation) and for reaction norm slope (Table 1). As stated above, the 16 families occupied plants that did not differ significantly in lag time, and so these results are not artifacts due to the families occupying different segments of the lag time gradient. Correlations between family mean lag time and family-specific elevation and slope were performed as an additional check against artifactual results; the correlation coefficients were not significant, with values of −0.089 ($P = 0.73$) and −0.034 ($P = 0.91$) respectively.

The variance components for the family-specific elevations and slopes accounted for 10.34% and 26.93% respectively of the phenotypic variance in gall size. The heritability estimate for reaction norm elevation is thus 0.21 (SE = 0.179) and for slope is 0.54 (SE = 0.251), assuming negligible effects of dominance, epistasis and maternal environment. It should be noted that although the ANCOVA indicated significant among-family variance for elevation (Table 1), the maximum likelihood estimates do not confirm a heritability greater than zero. A significant heritable component for reaction norm slope was confirmed in both analyses. Although these results do not conclusively demonstrate a contribution of the purely additive genetic variance to gall size variance, they do demonstrate a significant influence of "genotype × environment" interaction effects.

A visual examination of reaction norms (Fig. 3a) shows small family differences in elevation but greater differences in slope. Some families were insensitive to plant lag time, whereas others were highly sensitive, i.e., galls grew larger on fast reacting plants but smaller on slow plants. The correlation between elevation and slope was not significant ($r = -0.27$, $P = 0.30$). Plant lag

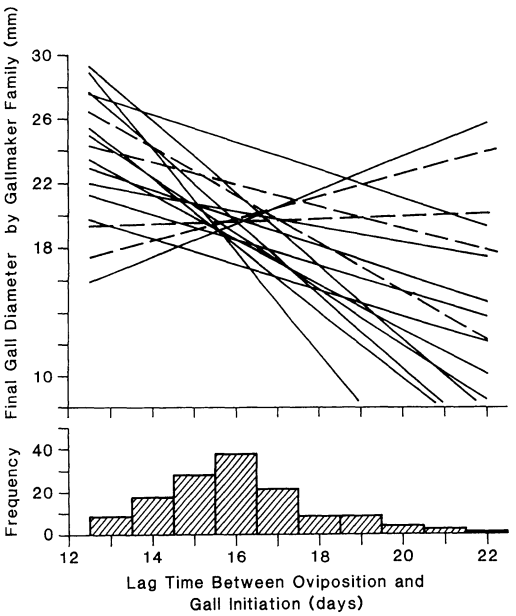

FIG. 3. Gall size reaction norms, with respect to plant lag time. a) Family-specific regressions of gall diameter over plant lag time; solid lines have slopes significantly different from zero, dashed lines do not. b) The frequency distribution of lag times among the host plants as observed in the quantitative genetic experiment.

times showed a unimodal distribution with a slight skew to the right (Fig. 3b); the mean lag time was 16.21 days. Many of the reaction norms intersect at or near the modal lag time (Fig. 3a and 3b). This indicates that the 16 families were most similar to one another in the segment of the environmental gradient most frequently occupied. There is no mathematical reason why they must converge in the average environment; it simply indicates that the variance in elevation is less than the variance in slope.

Evaluation of Selection on Reaction Norms

Given the observed fitness function and distribution along the plant lag time gradient, we found the following patterns of selection on gall size reaction norms. First, there was moderate selection favoring an increase in reaction norm elevation (Table 2). This agreed with the previous observation that natural enemy attack favors insects in larger galls (Weis and Abrahamson, 1986; Abrahamson et al., 1989). In addition, Table 2 shows that directional selection favors higher reaction norm slopes; since the observed mean slope was negative (−1.03 mm/day) the mean slope would evolve toward zero. With this episode of selection, the mean of the distribution for genetic values for elevation would shift by 12.2% of a standard deviation; mean genetic value for slope would shift by only 2.8% of a standard deviation. Stabilizing selection on both parameters was weak, with genetic variances in elevation and slope being reduced by 7.2% and 5.2% respectively. Partial regression plots confirmed the existence of an optimal intermediate elevation, but were ambiguous with respect to the position of the optimal slope. The main conclusion is that selection acts more strongly on the genetic component contributing to reaction norm elevation than on genetic contributions to slope. Although genetically influenced sensitivity to plant lag time explains more than half of the phenotypic variance in gall size, selection to change sensitivity is surprisingly weak.

TABLE 2. Selection on reaction norm parameters. The intensities of selection on reaction norm parameters are quantified as the standardized coefficients of the partial regression of expected fitness on elevation and slope. Standard errors on the coefficients are inaccurate because of violation of the independent errors assumption (see text), but are included to indicate the variability around each parameter.

	Regression coefficient	Std. error
Directional		
Elevation	0.122	0.014
Slope	0.028	0.014
Stabilizing		
Elevation	−0.072	0.016
Slope	−0.052	0.016
Correlational		
Elevation-Slope	0.062	0.014

DISCUSSION

Schmalhausen (1949) cited a number of studies that showed that genotypes are not necessarily expressed in the same way under all environmental conditions. He went on to argue that natural selection will cause a population to evolve toward a pattern of

genetic expression that fits the environments it encounters. When different phenotypes are favored under different environmental conditions, the evolution of phenotypic plasticity is expected (Bradshaw, 1965; Slobodkin and Rapoport, 1974). On the other hand, when a single phenotype is favored under all environmental conditions, selection will lead to canalization (Waddington, 1960) of development. In either case, alleles that contribute toward an optimal reaction norm will on average be favored over those that do not.

Bradshaw (1965) noted that phenotypic plasticity at one level of organization may be favorable for maintaining homeostasis at another level. The *Eurosta-Solidago* system illustrates how this may be so. Insect genotypes that show no plasticity in gall size (reaction norms with zero slope) may in fact be plastic for their stimulus phenotypes, which enable them to adjust their gall-inducing stimulus to suit the reactivity of the plant stem tissue.

There have been two recent trends in the study of the evolution of environmental sensitivity. The first trend consists of theoretical studies of the dynamics and stability of genetic variation for environmental sensitivity under various scenarios for environmental variation (e.g., Slatkin and Lande, 1976; Via and Lande, 1985, 1987; Lynch and Gabriel, 1987) The second trend has been the measurement of genetic variation for environmental sensitivity in natural populations (e.g., Via, 1984*a*, 1984*b*; Scheiner and Goodnight, 1984; Scheiner and Lyman, 1989; Service and Lenski, 1982; Schlicting and Levin, 1984; Taylor and Aarssen, 1988). Unfortunately the link between theory and observation is usually weak. One problem that hampers the reconciliation of the empirical studies with theory is that environmental sensitivity (or phenotypic plasticity) has been evaluated in ways that sacrifice information on phenotypes per se; this in turn makes it difficult to establish a causal link between sensitivity and fitness. For instance, performance regression (Taylor and Aarssen, 1988; Eberhart and Russell, 1966), (where sensitivity of a particular trait in a particular genotype is measured as the regression of its environment-specific deviation from the population phenotypic mean over the environment-specific mean of all genotypes) graphically displays the degree of sensitivity, but does not provide a link between sensitivity and fitness. Methods by Scheiner and Goodnight (1983) and Scheiner and Lyman (1989), which measure plasticity in a way analogous to heritability, reveal the potential for response to selection but not the selection intensity. The procedure used in this paper makes the link between sensitivity and fitness. Unfortunately, violations and statistical assumptions preclude hypothesis testing in this case. We would like to point out however, that in some systems it should be possible to overcome the non-independence of reaction norms and expected fitness estimates that hamper this study. This can be done if reaction norms for both the character of interest and for fitness can be measured on the same experimental units simultaneously. For instance, in cloned replicates of perennial plants both morphological traits and reproductive output could be independently measured as functions of an environmental gradient. An approach similar to this was taken by Rausher and Simms (1989) to evaluate the cost of resistance to herbivory.

A more difficult problem in empirical studies on the adaptive significance of environmental sensitivity is determining the distribution of the population along the environmental gradient, then recreating that gradient in an experimental setting. This can be especially troubling when the "selectional" environment is not easily disentangled from the "developmental" environment (Lewontin, 1983; Antonovics et al., 1988). Yet, without an accurate measure of the distribution across the environment, measuring selection on phenotypic plasticity is simply not possible (Lewontin, 1974; Scheiner and Goodnight, 1984; Via and Lande, 1985). Our success with the *Eurosta* system depended on the fact that another organism, the host plant, which was easily measured, represented the environmental gradient. In any case, dealing with an environmental factor that the population experiences in a "coarse grain" fashion will be more amenable to analysis than a factor such as temperature, which in nature fluctuates by the hour (cf. Arnold, 1988).

Our analysis shows that selection on gall size in the average environment (elevation) is stronger than selection on the sensitivity of gall size to environmental variation (slope). However, the evidence indicates that the genetic variance available for response to selection is less for elevation than for slope. This result corresponds with the prediction that characters under stronger selection will show less genetic variation than those under weaker selection (Fisher, 1958; Falconer, 1981). Yet theoretical studies (Lynch and Gabriel, 1987; Slatkin and Lande, 1976) would predict that when the same phenotype is favored over all environments, selection will eventually eliminate genetic variance for sensitivity to environmental variation. Sensitivity variance could be maintained if correlational selection was strong, (i.e., if genotypes with "low elevation, steep negative slope" reaction norms had nearly equal fitness to those with "high elevation, zero slope"), but this was not the case. One possible mechanism to maintain genetic variance in sensitivity would be a selection-mutation balance (Lande, 1976; Turelli, 1984; Via and Lande, 1987); the weak selection against sensitivity means it could have potentially high equilibrium genetic variance.

Would selection against sensitivity be weak under all conditions? We explored this question with a reanalysis of the data. Expected fitnesses for the 16 families were recalculated assuming that plant lag times were evenly distributed. If lag time had been distributed in this way, a family member would have been just as likely to develop on a very slow or fast plant as on an average plant. Had this been the case, families that were sensitive, (steep reaction norm slopes) would have produced proportionately more very large and very small galls, and fewer of intermediate size. With this environmental distribution, selection on slope became as strong as that on elevation. It then follows that the rate at which genetic variation for environmental sensitivity is eroded by selection depends on how broadly a population is distributed on the developmental environmental gradient, relative to the magnitude of sensitivity. The implications that this holds for phenomena such as genetic assimilation (Waddington, 1961) deserve further investigation.

Acknowledgments

We thank W. Abrahamson, R. Lande, R. Lenski, T. Meagher, M. Parker, P. Price, S. Scheiner, D. Schemske, D. Snow, S. Via, R. Walton, and two anonymous reviewers for their critical comments on presentations of this material or earlier versions of the manuscript. We were assisted in the greenhouse by V. Gorski, M. Piro and C. Wolfe. Funding was provided by NSF grants DEB82-05856 and BSR86-14896.

Literature Cited

Abrahamson, W. G., J. F. Sattler, K. D. McCrea, and A. E. Weis. 1989. Variation in selection pressures on the goldenrod gall fly and the competitive interactions of its natural enemies. Oecologia 79: 15–22.

Antonovics, J., N. C. Elstrand, and R. N. Brandon. 1988. Genetic variation and environmental variation: Expectations and experiments, pp. 275–303. *In* L. D. Gottleib and S. K. Jain (eds.), Plant Evolutionary Biology. Chapman and Hall, N.Y.

Arnold, S. J. 1988. Quantitative genetics and selection in natural populations: Microevolution of vertebral numbers in the garter snake *Thamnophis elegans,* pp. 619–636. *In* B. S. Weir, E. J. Eisen, M. M. Goodman, and G. Namkoong (eds.), Proceedings of the Second International Congress on Quantitative Genetics. Sinauer, Sunderland, MA.

Bradshaw, A. D. 1965. Evolutionary significance of phenotypic plasticity in plants. Adv. Genet. 13:115–155.

Cane, J. H., and F. E. Kurczewski. 1976. Mortality factors affecting *Eurosta solidaginis* (Diptera: Tephritidae). J. N. Y. Entomol. Soc. 84:275–282.

Cavalli-Sforza, L. L., and W. F. Bodmer. 1971. The Genetics of Human Populations. W. H. Freeman and Co., San Francisco, CA.

Clausen, J., D. D. Keck, and W. M. Heisey. 1958. Experimental studies on the nature of species, III: Environmental responses of climatic races of *Achillea.* Carnegie Institution of Washington Publication 581.

Dawkins, R. 1982. The Extended Phenotype. Oxford Univ. Press, Oxford.

Eberhart, S. A., and W. A. Russell. 1966. Stability parameters for comparing varieties. Crop. Sci. 6: 36–40.

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

Falconer, D. S. 1981. An Introduction to Quantitative Genetics, 2nd ed. Longmann, London, U.K.

Fisher, R. A. 1958. The Genetical Theory of Natural Selection, 2nd ed. Dover Publications, Inc., N.Y.

Goldschmidt, R. 1940. The Material Basis of Evolution. Yale Univ. Press, New Haven, CT.

GUPTA, A. P., AND R. C. LEWONTIN. 1982. A study of reaction norms in natural populations of *Drosophila pseudoobscura*. Evolution 36:934–948.

HENDERSON, C. R., JR. 1982. Analysis of Covariance in the mixed model: Higher-level, nonhomogeneous and random regressions. Biometrics 38:623–640.

JENNRICH, R., AND P. SAMPSON. 1983. General mixed model analysis of variance, pp. 413–425. *In* W. J. Dixon (ed.), BMDP Statistical Software. Univ. California Press, Berkeley, CA.

LANDE, R. 1976. The maintenance of genetic variability by mutation in polygenic characters with linked loci. Genet. Res. 26:221–235.

LANDE, R., AND S. J. ARNOLD. 1983. The measurement of selection on correlated characters. Evolution 37:1210–1226.

LEWONTIN, R. C. 1974. The Genetic Basis of Evolutionary Change. Columbia Univ. Press, N.Y.

———. 1983. Genes, organisms and environment, pp. 273–285. *In* D. S. Bendal (ed.), Evolution from Molecules to Men. Cambridge Univ. Press, Cambridge.

LYNCH, M., AND W. GABRIEL. 1987. Environmental tolerance. Am. Nat. 129:283–303.

MITCHELL-OLDS, T., AND R. G. SHAW. 1987. Regression analysis of natural selection: Statistical inference and biological interaction. Evolution 41:1149–1161.

MORRISON, D. F. 1983. Applied Linear Statistical Methods. Prentice-Hall, Englewood Cliffs, NJ.

RAUSHER, M. D., AND E. L. SIMMS. 1989. The evolution of resistance to herbivory in *Ipomoea purpurea*. I. Attempts to detect selection. Evolution 43: 563–572.

SAS INSTITUTE, INC. 1985. SAS User's Guide: Statistics, Version 5. SAS Institute, Inc., Cary, NC.

SCHEINER, S. M., AND C. J. GOODNIGHT. 1984. The comparison of phenotypic plasticity and genetic variation in populations of *Danthonia spicata*. Evolution 38:845–855.

SCHEINER, S. M., AND R. F. LYMAN. 1989. The genetics of phenotypic plasticity, I: Heritability. J. Evol. Biol. 2:95–107.

SCHLICTING, C. D. 1986. The evolution of phenotypic plasticity in plants. Ann. Rev. Ecol. Syst. 17:667–693.

SCHLICTING, C. D., AND D. A. LEVIN. 1984. Phenotypic plasticity of annual *Phlox*: Tests of some hypotheses. Amer. J. Bot. 71:252–260.

SCHLUTER, D. 1988. Estimating the form of natural selection on a quantitative trait. Evolution 42:849–861.

SCHMALHAUSEN, I. I. 1949. Factors of Evolution. Univ. of Chicago Press, Chicago, IL.

SERVICE, P. M., AND R. E. LENSKI. 1982. Aphid genotypes, plant phenotypes and genetic diversity; A demographic analysis of experimental data. Evolution 36:1276–1282.

SLATKIN, M., AND R. LANDE. 1976. Niche width in a fluctuating environment—density independent model. Am. Nat. 110:31–55.

SLOBODKIN, L. B., AND A. RAPOPORT. 1974. An optimal strategy of evolution. Q. Rev. Biol. 49:181–200.

TAYLOR, D. R., AND L. W. AARSSEN. 1988. An interpretation of phenotypic plasticity in *Agropyron repens* (Graminae). Am. J. Bot. 75:401–413.

TURELLI, M. 1984. Heritable genetic variation via mutation-selection balance: Lerch's zeta meets the abdominal bristle. Theor. Pop. Biol. 25:138–193.

VIA, S. 1984*a*. The quantitative genetics of polyphagy in an insect herbivore, I: Genotype-environment interactions in larval performance on different host plant species. Evolution 38:881–895.

———. 1984*b*. The quantitative genetics of polyphagy in an insect herbivore, II. Genetic correlations in larval performance within and among host plants. Evolution 38:896–905.

VIA, S., AND R. LANDE. 1985. Genotype-environment interaction and the evolution of phenotypic plasticity. Evolution 39:505–522.

———. 1987. Evolution of genetic variability in a spatially heterogeneous environment: Effects of genotype-environment interaction. Genet. Res., Camb. 49:147–156.

WADDINGTON, C. H. 1960. Experiments on canalizing selection. Genet. Res., Camb. 1:140–150.

———. 1961. Genetic assimilation. Adv. Genet. 10: 257–290.

WEIS, A. E., AND W. G. ABRAHAMSON. 1985. Potential selection pressures by parasitoids on a plant-herbivore interaction. Ecology 66:1261–1269.

———. 1986. Evolution of host-plant manipulation by gall makers: Ecological and genetic factors in the *Solidago-Eurosta* system. Am. Nat. 127:681–695.

WEIS, A. E., H. G. HOLLENBACH, AND W. G. ABRAHAMSON. 1987. Genetic and maternal effects on seedling characters of *Solidago altissima* (Compositae). Amer. J. Bot. 74:1476–1486.

WEIS, A. E., W. G. ABRAHAMSON, AND K. D. MCCREA. 1985. Host gall size and oviposition success by the parasitoid *Eurytoma gigantea*. Ecol. Entomol. 10: 341–348.

WEIS, A. E., R. WALTON, AND C. L. CREGO. 1988. Reactive plant tissue sites and the population biology of gall makers. Annu. Rev. Entomol. 33:467–486.

Corresponding Editor: T. R. Meagher