Evolution, 57(3), 2003, pp. 487–495

EXPRESSION OF ADDITIVE GENETIC VARIANCES AND COVARIANCES FOR WILD RADISH FLORAL TRAITS: COMPARISON BETWEEN FIELD AND GREENHOUSE ENVIRONMENTS

JEFFREY K. CONNER,[1] RACHAEL FRANKS, AND CHRISTY STEWART
W. K. Kellogg Biological Station and Department of Plant Biology, Michigan State University, Hickory Corners, Michigan 49060
[1]*E-mail: connerj@msu.edu*

Abstract.—Measurements of the genetic variation and covariation underlying quantitative traits are crucial to our understanding of current evolutionary change and the mechanisms causing this evolution. This fact has spurred a large number of studies estimating heritabilities and genetic correlations in a variety of organisms. Most of these studies have been done in laboratory or greenhouse settings, but it is not well known how accurately these measurements estimate genetic variance and covariance expressed in the field. We conducted a quantitative genetic half-sibling analysis on six floral traits in wild radish. Plants were grown from seed in the field and were exposed to natural environmental variation throughout their lives, including herbivory and intra- and interspecific competition. The estimates of heritabilities and the additive genetic variance-covariance matrix (**G**) obtained from this analysis were then compared to previous greenhouse estimates of the same floral traits from the same natural population. Heritabilities were much lower in the field for all traits, and this was due to both large increases in environmental variance and decreases in additive genetic variance. Additive genetic covariance expressed was also much lower in the field. These differences resulted in highly significant differences in the **G** matrix between the greenhouse and field environments using two complementary testing methods. Although the **G** matrices shared some principal components in common, they were not simply proportional to each other. Therefore, the greenhouse results did not accurately depict how the floral traits would respond to natural selection in the field.

Key words.—Expression of additive genetic variance-covariance matrix, floral morphology, **G** matrix, heritability, *Raphanus raphanistrum*, wild radish.

Received February 1, 2002. Accepted November 13, 2002.

In the past two decades, evolutionary biologists have become increasingly interested in studying the processes of phenotypic evolution by natural selection. Much of this work has used the techniques of quantitative genetics, which can be encapsulated in a pair of closely related equations (Lande 1979; Lande and Arnold 1983; Arnold and Wade 1984a,b):

$$R = h^2 S \qquad \Delta\bar{\mathbf{z}} = \mathbf{G}\boldsymbol{\beta} \tag{1}$$

The first equation is the univariate case and the second is the analogous equation for the multivariate case, in which more than one trait is considered simultaneously; the latter is more useful because individual traits do not evolve in isolation. In both equations, the response to selection (R or $\Delta\bar{z}$) is the change in the mean value of a trait across one generation. This is short-term phenotypic evolution, which is the product of genetic variance (h^2 or **G**) and the strength of selection (S or $\boldsymbol{\beta}$). Heritability (h^2) is the proportion of the total phenotypic variance that is additive genetic (V_A/V_P), and thus readily available for selection to act upon. Since the environmental and nonadditive genetic (dominance and epistatic) variance is in the denominator of heritability, differences in heritability can be due to differences in additive genetic variance, nonadditive genetic variance, environmental variance, or some combination of the three. The most important difference between these equations is that **G** also contains the genetic covariances among traits, but like the heritability is a measure of potential for trait evolution under a given strength of selection.

To apply this very powerful conceptual and mathematical framework to understanding phenotypic evolution, we need measurements of both **G** and $\boldsymbol{\beta}$ in natural populations. Thanks to the widely applicable methods of Lande and Arnold (1983), we now have many measurements of S and $\boldsymbol{\beta}$ in natural populations, especially for morphological traits (Kingsolver et al. 2001). However, heritability and the **G** matrix are much more difficult to measure in the field, because large numbers of individuals of known genetic relationship, for example, parents and offspring or siblings, are necessary. For this reason most estimates of heritability and **G** come from laboratory or greenhouse studies. These lab and greenhouse estimates are likely to differ from the heritabilities and **G** matrix in the field due to different levels of environmental variance and differences across environments in the expression of genetic variance and covariance (Riska et al. 1989; Hoffmann and Merilä 1999 and references therein), but the magnitude of this difference is not well known.

There have been many demonstrations of significant genotype-by-environment (g-e) interaction (see Schlichting 1986; Scheiner 1993; Roff 1997 for reviews), and this is evidence for environmental effects on the expression of genetic variance. However, a significant additive genotype-by-environment interaction by itself is not evidence for differences in additive variance across environments, because with crossing reaction norms there can be g-e interaction without differences in variance across environments (see Hoffmann and Merilä 1999, Box 1). In this case, the rate of response to selection would be similar in the two environments, but the alleles increasing in frequency would be different across environments. In addition, most studies of g-e interaction do not directly address changes in genetic covariance across environments.

Although a number of animal studies have addressed differences in heritabilities and/or additive variation in favorable and unfavorable environments, generalities are difficult to

discern. Hoffmann and Merilä (1999) present eight hypotheses to explain differences in heritability and additive variance across environments. They classify them in terms of whether heritability is greater, smaller, or unpredictable between favorable and unfavorable environments, but in many cases they equate unfavorable environments with rare or novel environments. This is based on the reasonable assumption that natural populations will be best adapted to the environment they experience most often. This may not always be true in nature, however, particularly due to biotic interactions, since high levels of competition, predation, parasitism, and disease may be both common and clearly unfavorable. For example, many species become invasive after being introduced into a novel environment with fewer natural enemies or competitors. Novel environments will also be favorable in comparisons between the laboratory and field environments in species for which the laboratory or greenhouse is a less stressful environment than the field. Perhaps partially because of these ambiguities, but also perhaps due to lack of strong evidence and the diversity of nature, Hoffmann and Merilä are able to cite studies consistent with each of their eight hypotheses.

Studies of crop plants have generally found lower heritabilities under stressful conditions (Blum 1988; but see Ceccarelli 1994), but this has mainly been attributed to increased V_E rather than decreased V_A. Surprisingly few studies have explicitly reported changes in V_A across environments for the same natural plant population (Hoffmann and Merilä 1999; but see Mazer and Schick 1991; Mazer and Wolfe 1992; Montalvo and Shaw 1994; Schoen et al. 1994; Bennington and McGraw 1996; Thiede 1998). In addition, most of the estimates of V_A in plants are plagued by small sample sizes, that is, less than 25 half-sibling families, which means that the estimates have large standard errors (but see Bennington and McGraw 1996). Even fewer studies of natural plant populations have reported changes in additive correlations or covariances across environments (but see Mazer and Schick 1991; Bennington and McGraw 1996; Thiede 1998). Therefore, there are major gaps in our knowledge of changes in **G** across environments, especially in plants. In this paper we address two related questions: (1) are estimates of heritability and the **G** matrix from laboratory or greenhouse studies likely to be good indicators of the evolutionary potential of traits in the field? and (2) are there differences in the expression of additive variance and covariance across these two environments? To answer these questions, we estimated the **G** matrix for six floral traits in both the field and the greenhouse, using samples from the same natural population and similar designs.

Methods

We studied six floral traits in wild radish, *Raphanus raphanistrum*, which has become a model system for evolutionary studies, particularly of reproduction and floral biology (e.g. Stanton et al. 1986; Mazer 1987; Snow 1990; Conner 1997). Most of the greenhouse estimates were published previously (Conner and Via 1993).

Experimental Design

Generally the most efficient and least biased way to estimate additive genetic variances and covariances in different environments is with a nested half-sibling design. Hoffmann and Merilä (1999, p. 98) discuss several problems in comparing heritability estimates across environments; almost all of these are solved by using a paternal half-sibling design in species in which fathers contribute little or no resources to the offspring (true for plants and many animals).

For both the previous greenhouse and the current field study, the parental generation was raised in the greenhouse from seeds collected from an alfalfa field near Binghamton, New York, in 1989 (Conner and Via 1993). The greenhouse study had 50 sires each used to pollinate six unique randomly chosen dams for a total of ~~350~~ 300 dams. Four offspring from each dam were planted for a total of 1200 offspring in the design. The final sample size was 1133, because some did not germinate. The floral traits were measured on the parents and the offspring of known genetic relationship, and the genetic variances and covariances were estimated from these measurements. See Conner and Via (1993) and below for details.

The field study reported here was similar, except the design was 76 sires, each mated to between 5 and 9 randomly chosen dams (mean = 7.8), for a total of 593 dams. Seeds were planted in a tilled field at Kellogg Biological Station in Hickory Corners, Michigan. This high level of disturbance is typical of where wild radish is found naturally. The field was divided into two blocks, with 12 seeds from each dam planted in a randomly chosen location within each block (1186 locations). The 12 seeds were planted in a 8.5 × 8.5 cm grid in each location, and locations were in a grid with 1-m spacing. If multiple seeds germinated, we thinned to one randomly chosen sprout. Thus, we attempted to grow two offspring from each full-sibling family at 1-m spacing, one in each of the two blocks. At least one seed germinated in 1050 (88%) of the locations. Initially, we did some weeding of other plant species to ensure that sprouts became established, but in later developmental stages the plants experienced natural, and in many cases intense, levels of competition. The main competitor was lamb's quarters, *Chenopodium album*. The field was fenced to exclude rabbits, but otherwise plants were exposed to natural levels of herbivory, water, and nutrients throughout the study. Because all analyses are based on measurements of these offspring raised from seed in the field (see below), this provides a true field estimate of the **G** matrix.

Plants were monitored daily for flowering. Of the locations with seed germination, 888 (85%; 75% of all planted) produced a plant with a measurable flower; most of the other plants died before flowering. All sires had at least five offspring with a measured flower, and 70 of the 76 sires had nine or more. We attempted to measure the third flower on the central stalk, but in most cases (64%) this bud was aborted or there were gross abnormalities, in which case an alternate flower was measured. The vast majority of flowers measured were in the first 10 produced on the central stalk. Most plants that flowered (713) did so long enough for a second flower to be measured, and on five plants a third flower was measured as well. Results are from the average of all flowers

measured on a plant; analyses using just the first or second measurements alone produced identical qualitative results.

Each flower was dissected, placed on glass slides, and measured with digital calipers. Petal length and width, corolla tube length, short filament length, long filament length, and pistil length were measured (see Conner and Via 1993). These traits have been shown to be important in pollination success in a variety of studies and species, including wild radish (e.g. Galen 1989; Wolfe and Barrett 1989; Murcia 1990; Young and Stanton 1990; Campbell et al. 1991; Harder and Barrett 1993; Conner et al. 1995; Conner and Rush 1996; Cresswell 2000). We also analyzed genetic variance in two composite anther position traits: anther exsertion, defined as the long filament length minus the corolla tube length, and filament dimorphism, defined as the difference between the long and short filament lengths. Exsertion is important in pollination success and is a determinant of male fitness (Conner et al. 1995; Morgan and Conner 2001), and dimorphism is of interest because it is diagnostic to the family Brassicaceae (Karoly and Conner 2000; note that the measure of dimorphism in that paper is slightly different). Exsertion and dimorphism were not included in the matrix comparisons, because they are composites of other traits in the analysis.

Analysis

G matrices were estimated and compared using two methods with complementary strengths and weaknesses, both using restricted maximum likelihood (REML). The two sets of estimates were very similar, so only one set of estimates for each parameter will be presented as described below. The first method used the program pcrf1, part of the Quercus package (Shaw 1991; available at http://biosci.cbs.umn.edu/eeb/quercus.html). Restricted maximum likelihood estimates were made with and without the pair of matrices constrained to be the same, and the fit of the models compared using two times the difference in log-likelihood, which has a chi-square distribution. The matrices were constrained to be positive definite, so that variances were non-negative (see Shaw et al. 1995, appendix, for details). The advantage of this method is that it incorporates all the information in the mating design in making the comparison and estimating standard errors, so that the test is very robust. This includes information from parents and offspring in a half-sibling mating design, which was available for the greenhouse data only. Parents of the field experiment were raised in the greenhouse, so their flowers were not measured. A disadvantage of this method is that it is not very statistically powerful (Shaw 1991).

The second method of estimating and comparing **G** matrices relied on sire breeding values estimated using best linear unbiased prediction (BLUP; Littell et al. 1996, ch. 6) in Proc Mixed (SAS Institute 1999) combined with common principal components analysis (CPC; Flury 1988; Phillips and Arnold 1999; Steppan et al. 2002) on the breeding value variance/covariance matrices. The advantages of this method are that, rather than just test whether the matrices are equal, CPC can be used to test a variety of hypotheses about the relationship between two matrices (see below). Common principal components analysis is also complementary to the Quercus approach because it is very sensitive to matrix differences (Houle et al. 2002). The BLUP estimates of sire breeding values are more accurate than sire family means because they use all available information, and thus are not biased by dominance and environmental effects, as are family means (Shaw et al. 1995). A disadvantage to this approach is that, by using breeding value matrices as input, the analysis does not incorporate error in estimation of these breeding values in the test, and therefore may overestimate the significance of matrix differences. Programs to perform CPC on half-sibling data directly have not been developed (P. Phillips, pers. comm.).

The BLUP breeding value **G** matrices were used to perform the set of hierarchical comparisons that is the hallmark of the CPC approach. The simplest comparison is whether the two matrices are equal, which is what the Quercus approach tests as well. If they are not equal, one can test whether the two matrices differ by similar proportions across all their elements. If the matrices are not proportional, then one can determine how many principal components they share in common. We used three different statistical methods within the CPC analysis with different strengths and weaknesses, all implemented using the CPC and CPCrand programs provided by P. Phillips (http://darkwing.uoregon.edu/~pphil/software.html). The Step-up and Jump-up (Phillips and Arnold 1999) are hypothesis-testing approaches. In the former, each model is tested against the model immediately below it, starting from the bottom. For example, the model in which the matrices share one principal component (PC) in common is tested against the model that assumes completely unrelated matrices; then two PCs in common is tested against one PC in common, and so on. Parametric chi-squared tests of log-likelihood ratios were used for the Step-up approach. In the Jump-up tests, each model is tested against the lowest-level model; that is, the two matrices are completely unrelated to each other. We used both parametric and randomization tests with 10,000 iterations for the Jump-up approach. The final method is a model-building approach, in which the best-fitting model was determined using Akaike's Information Criterion (AIC).

Additive genetic variances were also estimated as four times the sire variance component, which in turn was estimated using Proc Mixed models that included block (fixed effect), sire, dam nested within sire, and the sire-by-block interaction (random effects). These were the same models used to estimate the BLUPs. Block effects were statistically significant for all traits except anther exsertion, but always explained $<4\%$ of the variance. Sire-by-block interactions were never close to significance. Equality of variance across treatments was tested in all cases, with no significant deviations from equal variances found. Additive genetic variances were also estimated using Type I parametric methods (Proc Nested; SAS Institute 1999), and were virtually identical to the REML estimates so are not presented here. Statistical significance of the sire variance component (and thus significance of additive genetic variance) was assessed using the difference in $-2X$ log likelihood between the full model and a model without the sire effect. This difference is distributed as chi-square with one df; tests are one-tailed because variance components cannot be negative (Littell et al. 1996, p. 44). These estimates of additive variance were then used to

TABLE 1. Descriptive statistics; for each trait the greenhouse values are given in the top row and the field values below. V_P is total phenotypic variance, CV_A is the coefficient of additive genetic variation (Houle 1992), and h^2 is narrow sense heritability (V_A/V_P). Additive variance values used in CV_A and h^2 estimated as four times the sire variance component (restricted maximum likelihood estimate from Proc Mixed). Significance tests of individual heritabilities are from likelihood-ratio tests of the sire variance component (see Methods) and are denoted with asterisks: * $P < 0.05$, ** $P < 0.01$, **** $P < 0.0001$.

	Mean	SE	V_P	CV_A	h^2
Petal length	10.69	0.04	1.36	8.65	0.63****
	7.71	0.04	1.54	5.47	0.12*
Petal width	8.78	0.04	1.42	14.67	1.17****
	6.89	0.04	1.28	9.44	0.33****
Corolla tube	11.58	0.03	0.85	6.50	0.67****
	9.20	0.04	1.64	3.22	0.05
Short filament	9.58	0.03	0.71	6.25	0.51****
	6.69	0.05	2.47	0	0
Long filament	11.71	0.03	0.89	6.08	0.57****
	8.86	0.05	2.54	0	0
Pistil	14.79	0.05	3.17	11.44	0.90****
	11.29	0.06	3.02	9.06	0.35****
Anther exsertion	0.13	0.02	0.35	307.08	0.46****
	−0.34	0.03	0.69	−107.90	0.19**
Filament dimorphism	2.13	0.02	0.26	13.99	0.35****
	2.18	0.02	0.50	5.01	0.02

calculate the narrow-sense heritability and the coefficient of additive genetic variation, CV_A (Houle 1992) that are presented in Table 1.

Quercus was also used to estimate the **E** matrices, which represent the residual variances and covariances from all nonadditive genetic sources. The **E** matrices are likely to mainly reflect environmental variance and covariance in this study, because common environment effects were minimized by our design and our focus on floral traits (Roach and Wulff 1987; Thiede 1998 and references therein), and dominance variance is likely to be small because dam variance components were rarely larger than sire variance components (Conner and Via 1993; Falconer and Mackay 1996; J. K. Conner, unpubl. data).

RESULTS

The field and greenhouse were clearly very different environments with respect to the expression of the floral traits (Table 1). Mean trait values were significantly less in the field for all traits except filament dimorphism. Mean filament dimorphism did not differ between the greenhouse and field because each filament differed the same amount across environments. Expression of total phenotypic variance was similar in the greenhouse and field for petal length and width and pistil length, but much higher in the field for the other traits, particularly the filament lengths. Coefficients of additive genetic variance and heritabilities were substantially lower in the field for all traits. All traits were highly significantly heritable (i.e., significant sire variance component) in the greenhouse, but only four of the eight were in the field. In fact, the two filament lengths had zero heritability in the field based on the Proc Mixed analysis. Interestingly, in spite of the lack of significant heritability in the field for either corolla tube or long filament length, their difference (anther exsertion) was significantly heritable.

Taken together, these results suggest lower additive var-

TABLE 2. **G** matrices. The top numbers in each cell are the greenhouse values, with the values for the field below. Additive genetic variances are on the diagonal and covariances below the diagonal. Standard errors (in parentheses) were derived from the information matrix; asterisks mark pairs of values for which 95% confidence intervals do not overlap. Estimated with restricted maximum likelihood using the program pcrf1 in Quercus (Shaw 1991).

	Petal length	Petal width	Corolla tube	Short filament	Long filament	Pistil
Petal length	0.919 (0.096)*					
	0.198 (0.106)					
Petal width	0.405 (0.074)	0.792 (0.086)				
	0.165 (0.095)	0.475 (0.116)				
Corolla tube	0.284 (0.055)	0.193 (0.052)	0.520 (0.058)*			
	0.030 (0.086)	0.082 (0.088)	0.132 (0.104)			
Short Filament	0.275 (0.052)*	0.158 (0.048)	0.365 (0.047)	0.445 (0.050)		
	−0.038 (0.098)	0.054 (0.101)	0.089 (0.110)	0.091 (0.142)		
Long filament	0.306 (0.060)*	0.147 (0.054)	0.480 (0.055)*	0.490 (0.053)*	0.659 (0.064)*	
	−0.056 (0.101)	−0.001 (0.103)	0.085 (0.115)	0.103 (0.140)	0.151 (0.152)	
Pistil	0.517 (0.108)	0.291 (0.099)	0.426 (0.083)	0.339 (0.077)	0.461 (0.087)	2.417 (0.208)*
	0.251 (0.142)	0.270 (0.142)	0.274 (0.144)	0.171 (0.159)	0.248 (0.170)	1.082 (0.279)

TABLE 3. **E** matrices, which represent the residual variances and covariances after additive genetic effects are removed. These are likely to be mainly caused by the environment rather than nonadditive genetic effects (see Methods). The top values in each cell are from the greenhouse and the bottom from the field, with standard errors in parentheses. None of the 95% confidence intervals overlapped, as denoted by the asterisks. For further details see Table 2.

	Petal length	Petal width	Corolla tube	Short filament	Long filament	Pistil
Petal length	0.557 (0.056)* 1.404 (0.118)					
Petal width	0.405 (0.044)* 0.933 (0.096)	0.515 (0.050)* 0.827 (0.103)				
Corolla tube	0.037 (0.032)* 1.114 (0.102)	0.124 (0.031)* 0.842 (0.092)	0.386 (0.036)* 1.566 (0.123)			
Short filament	0.106 (0.031)* 1.413 (0.123)	0.120 (0.029)* 0.981 (0.109)	0.248 (0.029)* 1.598 (0.136)	0.336 (0.030)* 2.453 (0.181)		
Long filament	0.158 (0.033)* 1.488 (0.126)	0.175 (0.031)* 1.097 (0.111)	0.260 (0.032)* 1.723 (0.141)	0.247 (0.030)* 2.221 (0.174)	0.296 (0.034)* 2.455 (0.186)	
Pistil	0.352 (0.059)* 1.154 (0.141)	0.301 (0.054)* 0.860 (0.128)	0.142 (0.045)* 1.241 (0.146)	0.196 (0.042)* 1.399 (0.168)	0.185 (0.045)* 1.513 (0.176)	0.731 (0.103)* 1.958 (0.247)

iance and greater environmental variance in the field compared to the greenhouse; the diagonals of the **G** and **E** matrices confirm this (Tables 2 and 3). The additive genetic variances in the field were only 20–60% of the values in the greenhouse, whereas, conversely, the environmental variances in the greenhouse were only 12–62% of the field values. All pairs of field and greenhouse variances had nonoverlapping 95% confidence intervals with the exception of the additive genetic variances for petal width and short filament length. The covariances showed similar patterns, with the additive genetic covariances always less, and the environmental covariances always greater, in the field compared to the greenhouse. However, most of the additive covariance pairs had overlapping 95% confidence intervals (none of the environmental covariance pairs did).

The strong differences between the **G** matrices are confirmed by the formal tests. The Quercus test showed a highly significant difference between the **G** matrices ($\chi^2 = 49.3$, 21 df, $P = 0.0004$). This difference is still highly significant when the two filament lengths (which had no additive variance in the Proc Mixed analysis) were eliminated from the analysis ($\chi^2 = 41.1$, 10 df, $P = 0.00001$). The CPC analysis on BLUP breeding values also shows that the **G** matrices differ strongly between the greenhouse and field, but that they did share principal components in common (Table 4). The Jump-up test using randomization suggests that the matrices shared three principal components in common. Because the parametric Step-up test for one principal component in common versus unrelated matrices was significant, both the Step-up and parametric Jump-up tests indicate that the matrices are completely unrelated, although the marginal *P*-value for this result provides some support for the model with two principal components in common. The model-fitting approach using Akaike's Information Criterion (AIC) agrees with this latter conclusion, that the matrices share two principal components in common, but note that all models from unrelated to three PCs in common had similar AIC values. Taken together, these results suggest quite strongly that the matrices share only two or three of their six principal components.

Because the two filament traits had zero additive variance in the field in the Proc Mixed analysis (Table 1), the CPC analysis was repeated with the other four traits only (Table 5). Here all three approaches agree completely, suggesting that these reduced field and greenhouse **G** matrices share all four principal components in common, but that they are not merely proportional to one another.

TABLE 4. Results of the common principal components (CPC) analysis with all six floral traits included. Proportional means that all elements of the two matrices differ by a similar proportion, CPC means that the matrices share all principal components in common, whereas 1 PC, 2 PC, 3 PC, and 4 PC means the matrices share 1, 2, 3, or 4 principal components in common, respectively. The chi-square values are for the Step-up test (Phillips and Arnold 1999); starting at the bottom and moving up, the first significant *P*-value indicates that the higher-level model is a significantly worse fit than the lower model, so that the latter is supported. The Jump-up *P*-values in the table are based on 10,000 randomizations, and test each higher model against the model that assumes the matrices are completely unrelated to each other. AIC is Akaike's Information Criterion, in which the smallest value indicates that the higher model is the best fit.

Higher model	Lower model	Chi-square	df	Step-up *P*	Jump-up *P*	AIC
Equality	Proportional	620.67	1	<0.0001	<0.0001	825.84
Proportional	CPC	134.91	5	<0.0001	<0.0001	207.18
CPC	4 PC	0.003	1	0.96	0.0014	82.26
4 PC	3 PC	44.69	2	<0.0001	0.0002	84.26
3 PC	2 PC	12.28	3	0.006	0.36	43.57
2 PC	1 PC	1.31	4	0.86	0.47	37.28
1 PC	Unrelated	11.97	5	0.04	0.11	43.97
Unrelated	—					42.00

TABLE 5. Common principal components analysis for the four traits with nonzero additive variance in the Proc Mixed analysis (i.e., eliminating the two filament lengths). See text and Table 4 for details.

Higher model	Lower model	Chi-square	df	Step-up *P*	Jump-up *P*	AIC
Equality	Proportional	348.06	1	<0.0001	<0.0001	400.54
Proportional	CPC	48.13	3	<0.0001	0.0013	54.47
CPC	2 PC	0.32	1	0.57	0.84	12.34
2 PC	1 PC	0.96	2	0.62	0.72	14.02
1 PC	Unrelated	3.06	3	0.38	0.67	17.06
Unrelated	—					20.00

DISCUSSION

We found large differences between the **G** matrices for the same population of wild radish raised in the field versus the greenhouse. Both the additive genetic variances and covariances were much smaller in the field compared to the greenhouse, and although the two matrices shared principal components in common, they were not simply proportional. Therefore, responses to selection, both direct and correlated, would be much less than would have been expected based on the greenhouse results. However, given the shared structure in the matrices, and the fact that the covariances seemed to differ less than the variances, it is not clear that the direction of the evolutionary trajectories would differ greatly between the two environments.

The differences between field and greenhouse in the additive coefficient of variation (Table 1) are not as dramatic as the differences in additive variance, partly because the means are lower in the field but also because taking the square root of the variance (as is done in the CV_A) tends to reduce the relative differences. Still, the CV_A values are considerably larger in the greenhouse, and it is the variances that determine the absolute response to selection. Also note that due to the much greater environmental variance in the field, total phenotypic variance does not scale with the mean, but rather is about the same or greater in field. Therefore, the lower additive variance in the field is not simply due to a difference in scale, and it is clear that the evolutionary response to selection acting on the phenotypic variance would be greatly reduced in the field compared to the greenhouse environments.

We will next discuss possible reasons for the large differences in additive genetic variance we found, and then attempt to assess the generality of our results by reviewing other studies comparing quantitative genetic estimates made in the laboratory or greenhouse versus field environments.

Hypotheses for Changes in Expression of Additive Variance and Covariance

The large differences in additive variance for the same population suggests differences in gene expression between the greenhouse and field environments. Larger flowers are more attractive to pollinators and there is direct selection for increased flower size in wild radish (Conner and Rush 1996; Conner et al. 1996). Therefore, the significantly larger measurements for all traits in the greenhouse show that the novel greenhouse environment is more favorable than the field. Although this is opposite to the situation considered by Hoffmann and Merilä in their review (1999), it is valuable to review their hypotheses that predict greater additive variance in novel or less stressful environments.

One hypothesis is that alleles that are only unfavorable in a rare or novel environment will not be removed by selection, and therefore expressed genetic variation and covariation will be greater in the novel environment (Service and Rose 1985; Holloway et al. 1990). The greater variation expressed in the novel greenhouse environment in our study lends support to this hypothesis. The larger means of all the traits in the greenhouse could be evidence against unfavorable alleles for floral traits being expressed in this novel environment, but these increased means are likely due to the favorable environment, which might overwhelm any deleterious genetic effects. Our field conditions might be considered a novel environment as well, since we planted the seeds in a tilled garden, the population was from New York but planted in Michigan, and wild radish is not native to North America. However, wild radish normally grows in highly disturbed environments such as agricultural fields, no difference in these floral traits between New York and Michigan plants was found in a common-garden study at this same site (Williams and Conner 2001), and wild radish has been in North America for at least 150 generations (based on herbarium specimens; Panetsos and Baker 1967; J. K. Conner, unpubl. ms.). Clearly, the greenhouse is a far more novel environment than the field in our study.

A related hypothesis states that stabilizing selection in the common environment creates canalization that will reduce phenotypic differences among genotypes. On first glance it seems that this hypothesis could explain our finding of reduced additive variation in the field, but our finding of increased environmental variance in the field and equal or greater phenotypic variance in the field is evidence against canalization explaining our differences in additive variance. In particular, note that the traits with the greatest increase in phenotypic variance in the field (corolla tube and filament lengths) were those that had the least additive variance.

The hypothesis that fits our data best is that poor growth conditions prevent superior genotypes from reaching their potential, therefore reducing additive variance. Although the increased phenotypic variance seen for corolla tube and filament lengths in the field seem to argue against this hypothesis, this was due entirely to increases in V_E, and V_A was much lower in the field. Further study is necessary to confirm or refute this hypothesis.

There is a possible methodological problem that could explain the reduced variance and covariance in the field compared to the greenhouse; that is, if there was stronger selection at the level of half-sib family in the field environment. It is

certainly true that a much smaller proportion of plants survived to flower in the field (72% vs. 94%), but there was not strong selection at the half-sib family level—all sires had at least five surviving offspring, and 70 of the 76 had nine or more. In addition, for selection to have altered variance for floral traits there would have to be genetic correlations between mortality and floral traits, but most of the mortality was lack of germination, and much of this was due to partial seed abortion. In the greenhouse there were negative but nonsignificant genetic correlations between germination time and the six floral traits (Conner and Via 1993). Therefore, selection seems an unlikely explanation for our differences in **G**.

Laboratory versus Field Estimates of ***G***

One message from our study is that laboratory or greenhouse studies will sometimes grossly overestimate heritability in the field, due to increased environmental variance and decreased expression of additive variance. This is somewhat discouraging, because it is always difficult to obtain field estimates of heritability, and for some organisms it is impossible with traditional methods. There may be more hope with newer methods that do not rely on controlled crosses (Ritland 1996; Lynch 1999) but these methods have not been widely tested and adopted to date (Lynch and Walsh 1998; but see Reale and Roff 2001). Traditional quantitative genetic experiments in the field are less difficult for short-lived plants and birds than for most other organisms, and a number of these kinds of studies have been done (Weigensberg and Roff 1996 review bird studies; Mitchell-Olds 1986; Shaw 1986; Campbell 1996; Galen 1996 and others cited in the introduction are examples of field plant studies).

It is hard to assess the generality of the conclusion that laboratory heritabilities will overestimate field heritabilities, because very few other studies have measured heritability of the same traits in the same population in both the laboratory and field environments. The studies that have been done suggest that our conclusions may be general; that is, laboratory measures will overestimate field heritabilities due to both decreased V_E and increased V_A in the laboratory. Weigensberg and Roff (1996) analyzed the eight studies that had done so by 1996, and found that the lab heritabilities were 0.11 higher on average than the field estimates. This difference was marginally significant ($P = 0.07$) even with the very small sample size. To our knowledge only six previous studies have also reported the variance components underlying heritability in the field and lab for the same traits in the same populations. Five of these (Coyne and Beecham 1987; Montalvo and Shaw 1994; Schoen et al. 1994; Simons and Roff 1994; Thiede 1998) found very similar results to ours; that is, generally lower V_A and h^2, and higher V_E and V_P in the field compared to the laboratory or greenhouse. The other study (Bryant and Meffert 1998) also found increased V_E and V_P in the field, but in contrast to the other studies, V_A was also higher in the field on average, resulting in no significant change in h^2.

A challenge to the conclusion that lab studies will overestimate heritabilities comes from the broader survey of Weigensberg and Roff (1996), in which they compared 165 field estimates of heritabilities from 45 animal studies to 189 laboratory heritability estimates from Mousseau and Roff (1987). They found that field studies reported slightly higher heritabilities on average than did field studies. However, due to likely publication bias in field heritability studies (Palmer 2000) we believe that studies of the same traits and populations are necessary. Weigensberg and Roff report that 84% of the field heritability estimates were significant; this very high percentage suggests that nonsignificant estimates were not published. If heritabilities are really lower in the field, then it is likely that a lower proportion of them will be statistically significant in the field compared to the lab, as we found in our study. This problem will be exacerbated if the more difficult field studies also have lower sample sizes.

Conclusions

To understand phenotypic evolution, we need to know how common the large differences in expression of variance we found are over space and time, and in a wider variety of organisms. Our field estimate was from one species in one year at one location, which limits our ability to generalize because natural environments are so variable. Therefore, more studies that measure expression of additive variance and covariance in different environments are clearly needed. These studies will also help settle the question of the usefulness of laboratory or greenhouse measures of the **G** matrix and heritability. In addition, the time may be ripe for studies of the mechanisms of differential expression of additive variance across environments. Promising methods include both quantitative trait loci mapping within species across multiple environments (Dean 1995; Stratton 1998; Wu 1998) and DNA microarray analysis (Jain 2001).

Acknowledgments

We thank E. Powers, V. Prihoda, and J. Reed for assistance with data collection; P. Phillips and R. Shaw for advice on data analysis; F. Shaw for the Quercus analyses; and J. Fant, F. Knapczyk, H. Sahli, J. Willis, and an anonymous reviewer for comments on previous versions of this manuscript. This research was supported by the National Science Foundation under grant nos. DEB 9796183, DEB 9796185, and DBI 9605168. This is Kellogg Biological Station contribution no. 983.

Literature Cited

Arnold, S. J., and M. J. Wade. 1984a. On the measurement of natural and sexual selection: applications. Evolution 38:720–734.

———. 1984b. On the measurement of natural and sexual selection: theory. Evolution 38:709–719.

Bennington, C. C., and J. B. McGraw. 1996. Environment-dependence of quantitative genetic parameters in *Impatiens pallida*. Evolution 50:1083–1097.

Blum, A. 1988. Plant breeding for stress environments. CRC Press, Boca Raton, FL.

Bryant, E. H., and L. M. Meffert. 1998. Quantitative genetic estimates of morphometric variation in wild-caught and laboratory-reared houseflies. Evolution 52:626–630.

Campbell, D. R. 1996. Evolution of floral traits in a hermaphroditic plant: field measurements of heritabilities and genetic correlations. Evolution 50:1442–1453.

Campbell, D. R., N. M. Waser, M. V. Price, E. A. Lynch, and R. J. Mitchell. 1991. Components of phenotypic selection: Pollen export and flower corolla width in *Ipomopsis aggregata*. Evolution 45:1458–1467.

Ceccarelli, S. 1994. Specific adaptation and breeding for marginal conditions. Euphytica 77:205–219.
Conner, J. K. 1997. Floral evolution in wild radish: the roles of pollinators, natural selection, and genetic correlations among traits. Int. J. Plant Sci. 158:S108–S120.
Conner, J. K., and S. Rush. 1996. Effects of flower size and number on pollinator visitation to wild radish, *Raphanus raphanistrum*. Oecologia 105:509–516.
Conner, J. K., and S. Via. 1993. Patterns of phenotypic and genetic correlations among morphological and life-history traits in wild radish, *Raphanus raphanistrum*. Evolution 47:704–711.
Conner, J. K., R. Davis, and S. Rush. 1995. The effect of wild radish floral morphology on pollination efficiency by four taxa of pollinators. Oecologia 104:234–245.
Conner, J. K., S. Rush, and P. Jennetten. 1996. Measurements of natural selection on floral traits in wild radish (*Raphanus raphanistrum*). I. Selection through lifetime female fitness. Evolution 50:1127–1136.
Coyne, J. A., and E. Beecham. 1987. Heritability of two morphological characters within and among natural populations of *Drosophila melanogaster*. Genetics 117:727–737.
Cresswell, J. E. 2000. Manipulation of female architecture in flowers reveals a narrow optimum for pollen deposition. Ecology 81: 3244–3249.
Dean, A. M. 1995. A molecular investigation of genotype by environment interactions. Genetics 139:19–33.
Falconer, D. S., and T. F. C. Mackay. 1996. Introduction to quantitative genetics. Longman, Harlow, U.K.
Flury, B. D. 1988. Common principal components and related multivariate models. Wiley, New York.
Galen, C. 1989. Measuring pollinator-mediated selection on morphometric traits: bumblebees and the alpine sky pilot, *Polemonium viscosum*. Evolution 43:882–890.
———. 1996. Rates of floral evolution: adaptation to bumblebee pollination in an alpine wildflower, *Polemonium viscosum*. Evolution 50:120–125.
Harder, L. D., and S. C. H. Barrett. 1993. Pollen removal from tristylous *Pontederia cordata*: effects of anther position and pollinator specialization. Ecology 74:1059–1072.
Hoffmann, A. A., and J. Merilä. 1999. Heritable variation and evolution under favourable and unfavourable conditions. Trends Ecol. Evol. 14:96–101.
Holloway, G. J., S. R. Povey, and R. M. Sibly. 1990. The effect of new environment on adapted genetic architecture. Heredity 64: 323–330.
Houle, D. 1992. Comparing evolvability and variability of quantitative traits. Genetics 130:195–204.
Houle, D., J. Mezey, and P. Galpern. 2002. Interpretation of the results of common principal components analysis. Evolution 56: 433–440.
Jain, K. K. 2001. Biochips for gene spotting. Science 294:621–625.
Karoly, K., and J. K. Conner. 2000. Heritable variation in a family-diagnostic trait. Evolution 54:1433–1438.
Kingsolver, J. G., H. E. Hoekstra, J. M. Hoekstra, D. Berrigan, S. N. Vignieri, C. E. Hill, A. Hoang, P. Gibert, and P. Beerli. 2001. The strength of phenotypic selection in natural populations. Am. Nat. 157:245–261.
Lande, R. 1979. Quantitative genetic analysis of multivariate evolution, applied to brain:body size allometry. Evolution 33: 402–416.
Lande, R., and S. J. Arnold. 1983. The measurement of selection on correlated characters. Evolution 37:1210–1226.
Littell, R. C., G. A. Milliken, W. W. Stroup, and R. D. Wolfinger. 1996. SAS system for mixed models. SAS Institute, Cary, NC.
Lynch, M. 1999. Estimating genetic correlations in natural populations. Genet. Res. Camb. 74:255–264.
Lynch, M., and B. Walsh. 1998. Genetics and analysis of quantitative traits. Sinauer Associates, Sunderland, MA.
Mazer, S. J. 1987. The quantitative genetics of life history and fitness components in *Raphanus raphanistrum* L. (Brassicaceae): ecological and evolutionary consequences of seed-weight variation. Am. Nat. 130:891–914.
Mazer, S. J., and C. T. Schick. 1991. Constancy of population parameters for life-history and floral traits in *Raphanus sativus* L. II. Effects of planting density on phenotype and heritability estimates. Evolution 45:1888–1907.
Mazer, S. J., and L. M. Wolfe. 1992. Planting density influences the expression of genetic variation in seed mass in wild radish (*Raphanus sativus* L., Brassicaceae). Am. J. Bot. 79:1185–1193.
Mitchell-Olds, T. 1986. Quantitative genetics of survival and growth in *Impatiens capensis*. Evolution 40:107–116.
Montalvo, A. M., and R. G. Shaw. 1994. Quantitative genetics of sequential life-history and juvenile traits in the partially selfing perennial, *Aquilegia caerulea*. Evolution 48:828–841.
Morgan, M. T., and J. K. Conner. 2001. Using genetic markers to directly estimate male selection gradients. Evolution 55: 272–281.
Mousseau, T. A., and D. A. Roff. 1987. Natural selection and the heritability of fitness components. Heredity 59:181–197.
Murcia, C. 1990. Effect of floral morphology and temperature on pollen receipt and removal in *Ipomoea trichocarpa*. Ecology 71: 1098–1109.
Palmer, A. R. 2000. Quasireplication and the contract of error: Lessons from sex ratios, heritabilities and fluctuating asymmetry. Annu. Rev. Ecol. Syst. 31:441–480.
Panetsos, C. A., and H. G. Baker. 1967. The origin of variation in ''wild'' *Raphanus sativus* (Cruciferae) in California. Genetics 38:243–274.
Phillips, P. C., and S. J. Arnold. 1999. Hierarchical comparison of genetic variance-covariance matrices. I. Using the Flury hierarchy. Evolution 53:1506–1515.
Reale, D., and D. A. Roff. 2001. Estimating genetic correlations in natural populations in the absence of pedigree information: Accuracy and precision of the Lynch Method. Evolution 55: 1249–1255.
Riska, B., T. Prout, and M. Turelli. 1989. Laboratory estimates of heritabilities and genetic correlations in nature. Genetics 123: 865–871.
Ritland, K. 1996. A marker-based method for inferences about quantitative inheritance in natural populations. Evolution 50: 1062–1073.
Roach, D. A., and R. D. Wulff. 1987. Maternal effects in plants. Annu. Rev. Ecol. Syst. 18:209–235.
Roff, D. A. 1997. Evolutionary quantitative genetics. Chapman and Hall, New York.
SAS Institute. 1999. SAS/STAT user's guide. Ver 8. SAS Institute, Cary, NC.
Scheiner, S. 1993. Genetics and evolution of phenotypic plasticity. Annu. Rev. Ecol. Syst. 24:35–68.
Schlichting, C. D. 1986. The evolution of phenotypic plasticity in plants. Annu. Rev. Ecol. Syst. 17:667–693.
Schoen, D. J., G. Bell, and M. J. Lechowicz. 1994. The ecology and genetics of fitness in forest plants. IV. Quantitative genetics of fitness components in *Impatiens pallida* (Balsaminaceae). Am. J. Bot. 81:232–239.
Service, P. M., and M. R. Rose. 1985. Genetic covariation among life-history components: The effect of novel environments. Evolution 39:943–945.
Shaw, R. G. 1986. Response to density in a wild population of the perennial herb *Salvia lyrata*: variation among families. Evolution 40:492–505.
———. 1991. The comparison of quantitative genetic parameters between populations. Evolution 45:143–151.
Shaw, R. G., G. A. J. Platenkamp, F. H. Shaw, and R. H. Podolsky. 1995. Quantitative genetics of response to competitors in *Nemophila menziesii*—a field experiment. Genetics 139:397–406.
Simons, A. M., and D. A. Roff. 1994. The effect of environmental variability on the heritabilities of traits of a field cricket. Evolution 48:1637–1649.
Snow, A. A. 1990. Effects of pollen-load size and number of donors on sporophyte fitness in wild radish (*Raphanus raphanistrum*). Am. Nat. 136:742–758.
Stanton, M. L., A. A. Snow, and S. N. Handel. 1986. Floral evolution: attractiveness to pollinators increases male fitness. Science 232:1625–1627.
Steppan, S. J., P. C. Phillips, and D. Houle. 2002. Comparative

quantitative genetics: evolution of the **G** matrix. Trends Ecol. Evol. 17:320–327.
Stratton, D. 1998. Reaction norm functions and QTL-environment interactions for flowering time in *Arabidopsis thaliana*. Heredity 81:144–155.
Thiede, D. A. 1998. Maternal inheritance and its effect on adaptive evolution: a quantitative genetic analysis of maternal effects in a natural plant population. Evolution 52:998–1015.
Weigensberg, I., and D. A. Roff. 1996. Natural heritabilities: Can they be reliably estimated in the laboratory? Evolution 50: 2149–2157.
Williams, J. L., and J. K. Conner. 2001. Sources of phenotypic variation in floral traits in wild radish, *Raphanus raphanistrum* (Brassicaceae). Am. J. Bot. 88:1577–1581.
Wolfe, L. M., and S. C. H. Barrett. 1989. Patterns of pollen removal and deposition in tristylous *Pontederia cordata* L. (Pontederiaceae). Biol. J. Linn. Soc. 36:317–329.
Wu, R. 1998. The detection of plasticity genes in heterogeneous environments. Evolution 52:967–977.
Young, H. J., and M. L. Stanton. 1990. Influences of floral variation on pollen removal and seed production in wild radish. Ecology 71:536–547.

Corresponding Editor: J. Willis