Evolution, 57(5), 2003, pp. 1107–1120

THE CONSTANCY OF THE **G** MATRIX THROUGH SPECIES DIVERGENCE AND THE EFFECTS OF QUANTITATIVE GENETIC CONSTRAINTS ON PHENOTYPIC EVOLUTION: A CASE STUDY IN CRICKETS

MATTIEU BÉGIN[1,2,3] AND DEREK A. ROFF[4]
[1]*Department of Biology, McGill University, 1205 Dr. Penfield Avenue, Montréal, Québec H3A 1B1, Canada*
[2]*E-mail: mbegin1@po-box.mcgill.ca*
[4]*Department of Biology, University of California, Riverside, California 92521*

Abstract.—Long-term phenotypic evolution can be modeled using the response-to-selection equation of quantitative genetics, which incorporates information about genetic constraints (the **G** matrix). However, little is known about the evolution of **G** and about its long-term importance in constraining phenotypic evolution. We first investigated the degree of conservation of the **G** matrix across three species of crickets and qualitatively compared the pattern of variation of **G** to the phylogeny of the group. Second, we investigated the effect of **G** on phenotypic evolution by comparing the direction of greatest quantitative genetic variation within species ($\mathbf{g}_{max}$) to the direction of phenotypic divergence between species ($\mathbf{\Delta\bar{z}}$). Each species, *Gryllus veletis*, *G. firmus*, and *G. pennsylvanicus*, was reared in the laboratory using a full-sib breeding design to extract quantitative genetic information. Five morphological traits related to size were measured. **G** matrices were compared using three statistical approaches: the T method, the Flury hierarchy, and the MANOVA method. Results revealed that the differences between matrices were small and mostly caused by differences in the magnitude of the genetic variation, not by differences in principal component structure. This suggested that the **G** matrix structure of this group of species was preserved, despite significant phenotypic divergence across species. The small observed differences in **G** matrices across species were qualitatively consistent with genetic distances, whereas ecological information did not provide a good prediction of **G** matrix variation. The comparison of $\mathbf{g}_{max}$ and $\mathbf{\Delta\bar{z}}$ revealed that the angle between these two vectors was small in two of three species comparisons, whereas the larger angle corresponding to the third species comparison was caused in large part by one of the five traits. This suggests that multivariate phenotypic divergence occurred mostly in a direction predicted by the direction of greatest genetic variation, although it was not possible to demonstrate the causal relationship from **G** to $\mathbf{\Delta\bar{z}}$. Overall, this study provided some support for the validity of the predictive power of quantitative genetics over evolutionary time scales.

Key words.—Common principal components, covariance matrix, genetic constraints, genetic correlation, genetic variance, heritability, line of least resistance.

Received September 5, 2002. Accepted December 23, 2002.

The evolution of quantitative traits can be modeled using the response-to-selection equation $\mathbf{\Delta\bar{z}} = \mathbf{G\beta}$, where $\mathbf{\Delta\bar{z}}$ is the vector of multivariate change in mean trait values, **G** is the matrix of additive genetic variances and covariances, and **β** is the vector of selection gradients (Lande 1979; Lande and Arnold 1983). The capacity of this equation to predict the response to natural selection in natural populations across one generation has been verified in a study of Darwin's finches by Grant and Grant (1995). However, it is still not known if the response-to-selection equation is of any value for predicting long-term evolutionary change. Investigating this question is crucial because the response to selection equation and the family of models based on it are the main theoretical links between microevolutionary processes and macroevolutionary patterns (Arnold et al. 2001). Lande (1979) postulated that long-term evolutionary trajectories can be predicted, or reconstructed, given that the **G** matrix remains approximately constant throughout evolution (i.e., changes in phenotypic mean trait values caused by directional selection are not accompanied by changes in **G**). Lande used phenotypic data from the fossil record and from selection experiments to argue that **G** may often be constant through time (references within Lande 1976, 1979) and proposed that this constancy is the result of a balance between stabilizing selection and pleiotropic mutations over time (Lande 1980).

The assumption of a constant **G** could be relaxed if changes in **G** through time could be modeled, but the current lack of information on the evolution of **G** prevents this. Simulation studies have shown that the long-term constancy of **G** is unlikely, but possible under some specific assumptions (Turelli 1988; Reeve 2000). However, ignorance of the distribution of allelic effects in natural populations, a critical parameter of these models, prevents any conclusions. Empirical comparisons of **G** matrices have shown that conservation of the **G** matrix is frequent across natural populations or species (reviewed in Arnold and Phillips 1999; Roff 2000; Steppan et al. 2002), although the typically low power of these tests may sometimes incorrectly prevent the rejection of the null hypothesis of equality. The current view of the constancy of the **G** matrix is that, although **G** may stay constant for some unknown evolutionary time, it is bound to evolve at some point during the evolution of a taxonomic group. A comparative framework is therefore needed to investigate the rate and pattern of evolution of the **G** matrix (Steppan et al. 2002). One possible approach is to test the hypothesis that **G** matrix variation is phylogenetically structured. In other words, can knowledge of the phylogeny of a taxonomic group be used to predict changes in the **G** matrix? Based on this approach it is predicted that recently separated populations do not differ in their **G** matrices and that the degree of divergence of **G**

[3] Present address: Department of Biology, University of California, Riverside, California 92521.

increases with taxonomic distance. This pattern loosely corresponds to the conclusions of the recent reviews of **G** matrix comparison (Arnold and Phillips 1999; Roff 2000; Steppan et al. 2002). However, only one **G** matrix study has so far used more than two taxonomic units (Lofsvold 1986; reanalyzed in Kohn and Atchley 1988), which is the minimum requirement for testing the hypothesis of a phylogenetic structure of **G** matrix variation. Many more comparative studies that include three or more species are needed to establish generalities concerning **G** matrix evolution.

Determining an approximate upper taxonomic level under which **G** matrices tend not to change would be an important first step in validating the use of the response-to-selection equation at these levels, therefore allowing the evolutionary trajectory of a population to be modeled or past selection gradients to be reconstructed. However, little is known about the extent to which the **G** matrix, the representation of genetic constraints (Arnold 1992), determines which phenotypes can and cannot evolve. Theory predicts that if a single adaptive peak exists in the vicinity of a population and if genetic constraints are not absolute, the population will eventually reach the peak regardless of the genetic constraints, which implies that the importance of the **G** matrix on long-term evolution is very low (Lande 1979; Via and Lande 1985; Zeng 1988). By contrast, the influence of **G** is thought to be much more important and durable if the adaptive landscape is complex or changing because small deviations in the direction of evolution may lead to different adaptive peaks or even to peak shifts (Lande 1979; Bürger 1986; Price et al. 1993). Given the typical absence of information on the topography of adaptive landscapes, a possible method to investigate the influence of **G** on the outcome of phenotypic evolution is to compare the direction of population or species differentiation to the direction that is the least genetically constrained in the ancestral population (Schluter 1996). Because ancestral populations are rarely available, it is possible to substitute the ancestral **G** with the **G** matrix of a descendent population if **G** is approximately equal among several descendent populations, which can be assumed to mean that **G** was relatively unchanged through time. However, this method of investigating the importance of **G** is only correlative and cannot prove the role of the **G** matrix in determining phenotypic evolution (see Discussion section). Schluter (2000, pp. 224–231) presented evidence from several studies on recently diverged populations or species and showed that the phenotypic evolution of traits that are thought to be under selection is often biased toward the direction of least genetic constraints. He concluded that the influence of **G** on phenotypic evolution often spans evolutionary time scales. However, the studies reviewed by Schluter (2000, pp. 224–231) estimated **G** for only one of the species or populations or depended on nongenetic information. Much more information on the relation of **G** to phenotypic divergence is needed.

The objectives of this study were to, first, investigate the constancy of the **G** matrix through species divergence and qualitatively compare the pattern of **G** matrix variation to the phylogeny of the group. Second, we investigated the importance of **G** in determining the direction of phenotypic evolution, as measured by the angle between the direction of least genetic constraints ($\mathbf{g}_{max}$) and the direction of species differentiation ($\Delta\bar{\mathbf{z}}$). These issues were addressed using three congeneric species of field crickets, *Gryllus veletis*, *Gryllus firmus*, and *Gryllus pennsylvanicus*, for which quantitative genetic variation in traits related to size was measured.

TABLE 1. Average genetic distances within and among species, based on a mitochondrial sequence that includes the whole cytochrome *b* gene and a 16S rRNA fragment. The units are average numbers of substitutions per site between a pair of individuals or species. Sequences were taken from one individual of *Gryllus pennsylvanicus* and two individuals of the two other species.

	G. veletis	*G. firmus*	*G. pennsylvanicus*
G. veletis	0.005		
G. firmus	0.126	0.003	
G. pennsylvanicus	0.103	0.021	—

MATERIALS AND METHODS

Study Organism

Crickets of the genus *Gryllus* are wing dimorphic orthopterans. *Gryllus veletis* and *G. pennsylvanicus* are found mainly in the northeastern United States and southern Canada, whereas *G. firmus* is distributed mainly along the southeastern coast of the United States (Alexander 1968). In this study, *G. veletis* and *G. pennsylvanicus* were both sampled in a field near Montreal, Quebec, Canada, and *G. firmus* was sampled in Florida. In addition to sharing a large part of their geographical distribution, *G. veletis* and *G. pennsylvanicus* live in the same microhabitats, are morphologically similar, and have virtually identical calling songs (Alexander and Bigelow 1960). However, these two species do not interbreed because of differences in the timing of reproduction and in developmental processes (Alexander and Bigelow 1960; Bigelow 1960), and they are in fact relatively distantly related (Table 1; Huang et al. 2000). *Gryllus firmus* and *G. pennsylvanicus* are known to hybridize along a narrow hybrid zone located in the Blue Ridge and Appalachian Mountains, approximately from Virginia to New York (Harrison and Arnold 1982). The populations sampled for the present study were separated by approximately 2500 km, however and hence gene flow between them is highly improbable. *Gryllus veletis* and *G. firmus* do not produce hybrids (Alexander 1968).

The phylogeny of these three cricket species is based on a mitochondrial sequence of 1536 base pairs including the whole cytochrome *b* gene and a 16S rRNA fragment (Huang et al. 2000). Genetic distances were estimated using maximum likelihood and the general time reversible model with among-site rate heterogeneity (Huang et al. 2000). Distances are calculated as the average number of substitutions per site between a pair of species or individuals (Table 1). Sequences were taken from two individuals of each species, except in the case of *G. pennsylvanicus*, where only one individual was used. Table 1 shows that *G. firmus* and *G. pennsylvanicus* are closely related, and that *G. veletis* is the most distant relative. Genetic distances between individuals of the same species are also given in Table 1 and confirm that interspecific divergence is greater than intraspecific variation.

TABLE 2. Summary of the rearing protocol for each species. The photoperiod is given in number of hours of light followed by the hours of darkness; ''mac'' refers to the macropterous (long-winged) morph and ''mic'' to the micropterous (short-winged) morph.

Species	Number of families	Number of individuals	Temperature	Photoperiod	Density per cage	Wing morph
Gryllus veletis	67	1096	28°C	15:9	40	mac
G. firmus	62	862	28°C	15:9	40	mac
G. pennsylvanicus	39	505	24°C	17:7	25	mic
G. veletis field	—	57	field	field	—	mic

Experimental Protocol

Gryllus veletis and *G. firmus* were reared in the laboratory under identical conditions, and data for *G. pennsylvanicus* were taken from another experiment (Simons and Roff 1994) in which the rearing protocol differed (Table 2). The breeding protocols for *G. veletis* and *G. pennsylvanicus* were initiated immediately after sampling from the wild, whereas the breeding protocol for *G. firmus* was initiated after five generations in the laboratory. This procedure thus minimizes the probability of adaptation to the laboratory environment. For each species, the parental generation was reared in a growth chamber under constant temperature and photoperiod (Table 2). Shortly before the final molt, the sexes were separated to avoid uncontrolled matings. After emergence of most adults, male-female pairs were randomly formed and put into plastic containers with an unlimited amount of water and food (rabbit chow) and a dish of moist earth for oviposition. Twelve to 14 days after pairing, the earth dish was removed and put into a separate plastic container to facilitate the collection of nymphs. A total of 80 (*G. veletis* and *G. firmus*) or 50 (*G. pennsylvanicus*) newly hatched offspring were collected from each family. The emerging nymphs were collected within 3 or 4 days to avoid large differences in size between siblings. The nymphs of each family were stored in two 4-L plastic buckets, at a density of 40 (*G. veletis* and *G. firmus*) or 25 (*G. pennsylvanicus*) per bucket, with constant environmental conditions (Table 2). Throughout this experiment, individuals were killed and preserved within 3 days after reaching adulthood. The number of families and the number of individuals per species are given in Table 2. In addition to these laboratory-reared populations, adult *G. veletis* were sampled directly from the field (Table 2) and preserved immediately.

The different rearing protocol used for *G. pennsylvanicus* (see above) may potentially bias comparisons with the two other species, along with the fact that most *G. pennsylvanicus* individuals were micropterous (short winged), as opposed to the other two species, which produced mostly macropterous individuals (long winged, Table 2). However, Bégin and Roff (2001) showed that rearing *G. pennsylvanicus* in two very different environments (laboratory vs. cages in the field) did not cause a significant difference in the **G** matrix of this species. In addition, a comparison of **G** across wing morphs in *G. firmus* suggested that the **G** matrix of the two morphs do not differ significantly in their principal component structure (M. Bégin and D. A. Roff unpubl. data).

Measurements

Five morphological measurements were taken from each female: femur length (FEMUR), head width (HEAD), prothorax length (PTHL), prothorax width (PTHW), and ovipositor length (OVIP). These five traits are all related to overall size. An analysis of measurement error using a small subsample of all five traits revealed that the repeatability of the measurements was close to 98% for each trait. All measurements were ln-transformed to remove the correlation between trait mean and variance. The univariate normality of the residuals of these variables was tested using the one-sample Kolmogorov-Smirnov test, Lilliefors option in SYSTAT (Systat Software Inc., Richmond, CA). Results (not shown) indicated that approximately a third of the trait distributions deviated significantly from normality. However, every distribution was unimodal and apparently symmetric. These deviations from normality were therefore small and insufficient to invalidate the results of statistical tests as long as no outliers are present (Tabachnick and Fidell 2001, pp. 72, 329). We used Mahalanobis' distance to test for multivariate outliers and found outliers only in the variable OVIP of *G. pennsylvanicus*. This problem could not be solved by any transformation, especially because the distribution of this trait in the other species was satisfactory. We removed some of the most extreme outliers and compared the resulting **G** matrices, but results were unchanged for two of the three statistical approaches (see description below). Because no simple solution was apparent, we chose not to further transform the measurements of the trait OVIP in *G. pennsylvanicus*.

Quantitative Genetic Methodology

The estimation of quantitative genetic parameters was based on a nested ANOVA/ANCOVA, with family and cage nested within family as the two independent variables and the five morphological traits as dependent variables. The genetic (co)variances were estimated as twice the among-family component of variance (Roff 1997, pp. 41–43). A jackknife procedure (see description in the MANOVA method section) was then implemented to estimate variances and covariances and their standard errors. The (co)variances were therefore estimated by the mean of all jackknife pseudovalues and their standard error estimated as the standard error of the pseudovalues. The number of jackknife iterations was equal to the number of families in the case of genetic (co)variances and to the number of individuals in the case of phenotypic (co)variances. The jackknife has been shown through simulations to produce accurate estimates and standard errors of heritabilities (Simons and Roff 1994) and genetic correlations (Roff and Preziosi 1994).

By definition, the (co)variances estimated using a full-sib family design include additive and nonadditive genetic com-

ponents and may be contaminated by maternal effects and common family environmental effects (Falconer and Mackay 1996; Roff 1997). We assumed that potential inbreeding had a negligible effect. The common family environmental effects were corrected for by our use of two cages per family but all other effects could potentially inflate the elements of **G**. However, a general review of the heritability of morphological traits indicated that, for a given trait, heritabilities estimated from a full-sib design are generally similar to those estimated from a parent-offspring regression (Mousseau and Roff 1987), which implies that dominance variance is generally not important in this type of trait. Moreover, the importance of dominance and maternal effects in femur length was investigated in *G. firmus* by comparing estimates from a full-sib, a half-sib, and a parent-offspring regression (Roff 1998). The results showed that these two sources of variation are not important in this trait in this species (Roff 1998). In the present study, it is therefore probable that genetic (co)variances estimated using a full-sib design are good estimates of additive genetic (co)variances.

To obtain as much information as possible about differences between **G** matrices, we used three matrix comparison methods: the T method, the Flury hierarchy, and the MANOVA method.

The T Method

The T method was developed by Roff et al. (1999) and is similar to the matrix disparity method suggested by Willis et al. (1991). The T method is based on the sum of element-by-element absolute differences between two matrices and tests the hypothesis that two matrices are equal, by calculating *T*:

$$T = \sum_{i=1}^{c} |\boldsymbol{M}_{i1} - \boldsymbol{M}_{i2}|, \tag{1}$$

where $\boldsymbol{M}_{i1}$ and $\boldsymbol{M}_{i2}$ are the estimates of the *i*th element of each of the two matrices and *c* is the number of nonredundant elements in the matrix (sum of the number of diagonal elements plus the number of elements above the diagonal). Note that squared differences can be used instead of absolute differences for the calculation of the *T* statistic without changing the outcome of this test. The probability that the two matrices come from the same statistical population is estimated by randomization (4999 iterations), where families are randomly assigned to one species or the other and quantitative genetic parameters are estimated for each iteration. The probability that two matrices are equal is based on the number of iterations in which the randomized dataset produced a *T* statistic larger than the observed *T*. The randomization procedure sets the mean and standard deviation to zero and one, respectively, for each trait in each randomized dataset.

Additionally the T method can be used to quantify the difference between two matrices using the *T*% statistic:

$$T\% = \frac{T/c}{(\bar{\boldsymbol{M}}_1 + \bar{\boldsymbol{M}}_2)/2} \times 100, \tag{2}$$

where $\bar{\boldsymbol{M}}_1$ and $\bar{\boldsymbol{M}}_2$ are the averages of the elements of the two matrices. This statistic measures the absolute difference between the elements of two matrices as a percentage of the overall average size of the elements in these matrices.

The Flury Hierarchy

The Flury hierarchy is a principal components approach to the comparison of matrices that has been applied to **G** matrix comparisons by Phillips and Arnold (1999). This method, based on maximum likelihood, determines which model is the best descriptor of the structural differences between two or more matrices. The hierarchically nested models are: (1) unrelated structure: matrices have no eigenvector in common; (2) partial common principal components: matrices share some eigenvectors; (3) common principal components: matrices share all eigenvectors; (4) proportionality: matrices share all eigenvectors, and eigenvalues all differ by the same constant between matrices; and (5) equality: matrices share eigenvectors and eigenvalues. For each model, the Flury hierarchy calculates a log-likelihood statistic to quantify the fit of that model to the observed matrices. A likelihood ratio is then calculated for each model against the model of unrelated structure C(''jump-up'' procedure, Phillips and Arnold 1999). To avoid the assumption of multivariate normality in hypothesis testing and because the degrees of freedom are unknown under the null hypothesis, randomization is used to determine the probability that a model fits the data significantly better than the unrelated structure model. In this analysis, 4999 randomized datasets were created, each iteration randomly assigning whole families to one species or the other. The best fitting model (referred to as the verdict in the Results section) is determined as the model immediately under the first significant probability, going from the bottom (unrelated structure model) to the top (equality model) of the hierarchy (jump-up procedure, Phillips and Arnold 1999). For the purpose of comparison with the other methods, only the probability of the equality model is given in the Results section. The randomization procedure sets the mean and standard deviation to zero and one, respectively, for each trait in each randomized dataset. This analysis was performed using the program CPCrand (Phillips 1998a).

Note that because the CPCrand program does not include the option of nesting cages within families, **G** matrix estimations and comparisons by the Flury hierarchy were performed by pooling the individuals of the two cages of a family. The results may therefore be biased by common family environmental effects. Because we have no way of directly testing for the effect of nesting on the results of the Flury hierarchy, we instead analyzed nonnested data with the T and MANOVA methods (see below) to determine if the results remained the same. Results (not shown) indicated that not correcting for common environmental effects produced no qualitative changes in the probabilities of the **G** matrix comparisons, although it inflated genetic (co)variances on average by 1% for *G. veletis*, 20% for *G. firmus*, and 13% for *G. pennsylvanicus*.

In the special case of the comparison of the field and laboratory samples of *G. veletis*, the parametric version of the Flury hierarchy (jump-up procedure, Phillips and Arnold 1999) was used to compare **P** matrices (program CPC, Phillips 1998b). This version tests for the significance of each

TABLE 3. Mean trait values (phenotypic standard deviations) are given for each species. Measurements are in millimeters and are not ln-transformed. Traits are defined in the text.

Species	FEMUR	HEAD	PTHL	PTHW	OVIP
Gryllus veletis	10.48 (0.41)	5.14 (0.20)	3.69 (0.19)	5.80 (0.23)	14.07 (0.81)
G. firmus	13.46 (0.69)	6.06 (0.34)	4.40 (0.27)	6.79 (0.39)	18.08 (1.34)
G. pennsylvanicus	11.55 (0.59)	5.27 (0.29)	3.64 (0.27)	6.14 (0.35)	15.83 (1.65)
G. veletis field	9.42 (0.44)	4.76 (0.24)	3.23 (0.19)	5.11 (0.25)	12.27 (0.78)

model by using a likelihood-ratio test instead of a randomization procedure. This assumes multivariate normality of the data and uses the number of individuals as the sample size. The reason for the change of method is that, when we tried the nonparametric version in this particular **P** matrix comparison, the likelihood algorithm very often did not converge, presumably because of the small sample size of the field population compared to the laboratory population.

The MANOVA Method

This method, recently developed by Roff (2002), makes use of the jackknife procedure (Manly 1997, pp. 24–33) and MANOVA approach. The jackknife is first used to produce a distribution of pseudovalues of matrix elements within each species. A pseudovalue is calculated by estimating a matrix element after deleting all individuals of one family, using the formula:

$$\phi_{ij} = n\boldsymbol{M}_i - (n - 1)\,\boldsymbol{M}_{i-j}, \tag{3}$$

where ϕ_{ij} is the pseudovalue of the ith matrix element corresponding to the deletion of family j, n is the number of families, $\boldsymbol{M}_i$ is the ith matrix element estimated for the whole dataset, and $\boldsymbol{M}_{i-j}$ is the matrix element estimated for the whole dataset minus the jth family. The number of pseudovalues calculated for a species is equal to the number of families in the case of a **G** matrix and to the number of individuals in the case of a **P** matrix. For a given family that has been removed, the pseudovalues corresponding to each matrix elements (15 elements in this case: five variances and 10 covariances) can be arranged in a row that will then constitute the pseudovalue of the whole matrix. Two or more matrices can then be compared by using the pseudovalues as data in a multivariate analysis of variance (MANOVA). The jackknife has been shown through simulations to produce accurate estimates and standard errors of heritabilities (Simons and Roff 1994) and genetic correlations (Roff and Preziosi 1994).

Calculation of $\boldsymbol{g}_{max}$, $\boldsymbol{\Delta\bar{z}}$, and the Angle between These Two Vectors

The methodology used to compare the direction of greatest genetic variation ($\mathbf{g}_{\max}$) to the direction of phenotypic divergence ($\boldsymbol{\Delta\bar{z}}$) was developed by Schluter (1996). The vector $\mathbf{g}_{\max}$ is obtained by estimating the first eigenvector of a **G** matrix through a principal component analysis and therefore represents the multivariate direction that accounts for the most genetic variation within a species. The vector $\boldsymbol{\Delta\bar{z}}$ is obtained using the equation:

$$\boldsymbol{\Delta\bar{z}} = \frac{\bar{\boldsymbol{z}}_1 - \bar{\boldsymbol{z}}_2}{[(\bar{\boldsymbol{z}}_1 - \bar{\boldsymbol{z}}_2)^{\mathrm{T}}(\bar{\boldsymbol{z}}_1 - \bar{\boldsymbol{z}}_2)]^{1/2}}, \tag{4}$$

where $\bar{\boldsymbol{z}}_1$ and $\bar{\boldsymbol{z}}_2$ are vectors of mean trait values of the two species, and T is the symbol for matrix transposition. The vector $\boldsymbol{\Delta\bar{z}}$ is the standardized multivariate direction of phenotypic differences between the two species. The angle θ between $\mathbf{g}_{\max}$ and $\boldsymbol{\Delta\bar{z}}$ is calculated as:

$$\theta = \cos^{-1}[(\mathbf{g}_{\max})^{\mathrm{T}}\,\boldsymbol{\Delta\bar{z}}]. \tag{5}$$

The angle θ can also be used to calculate the correlation between the two vectors (Cheverud and Leamy 1985): $r_{\mathrm{v}} = \cos\theta$.

It is possible to test the null hypothesis that the two vectors are parallel by using a test developed by Schluter (1996). However, this test was incorrectly described in the original paper (D. Schluter, University of British Columbia, pers. comm.) and the correct methodology is as follows. A bootstrap procedure (Manly 1997, pp. 34–68), which samples whole families with replacement to create new populations, is first used to obtain a distribution of $\mathbf{g}_{\max}$ vectors for each species. For each bootstrap run (4999), the angle θ between the bootstrapped $\mathbf{g}_{\max}$ and the observed $\mathbf{g}_{\max}$ is estimated (the original paper incorrectly stated that the angle between $\mathbf{g}_{\max}$ and $\boldsymbol{\Delta\bar{z}}$ was repeatedly estimated). The probability corresponding to the null hypothesis of parallelism of the two vectors is calculated based on the number of times the bootstrapped angle is larger than the observed angle between $\mathbf{g}_{\max}$ and $\boldsymbol{\Delta\bar{z}}$. This effectively tests whether the vector $\boldsymbol{\Delta\bar{z}}$ is contained within the bootstrapped distribution of $\mathbf{g}_{\max}$, and therefore assumes that $\boldsymbol{\Delta\bar{z}}$ is estimated without error.

RESULTS

Overview of the Data

The mean value of each trait was significantly different among the three laboratory-reared species (five ANOVAs, $P < 0.001$ in each case), and the five corresponding Tukey post hoc tests yielded a $P < 0.001$ for all pairwise combinations of species. Similarly, a MANOVA showed a highly significant multivariate difference between the three species (Wilks' $\lambda = 0.07$; approximate $F = 1336$; df = 10, 4912; $P < 0.001$). Despite these significant results, the difference in mean trait values between *G. pennsylvanicus* and *G. veletis* were small, as illustrated by the substantial overlap of four of their five trait distributions (Table 3). By contrast, the trait distribution of *G. firmus* generally did not overlap much with those of the two other species. The comparison of the laboratory-reared and field-caught populations of *G. veletis* revealed that wild individuals were significantly smaller than laboratory

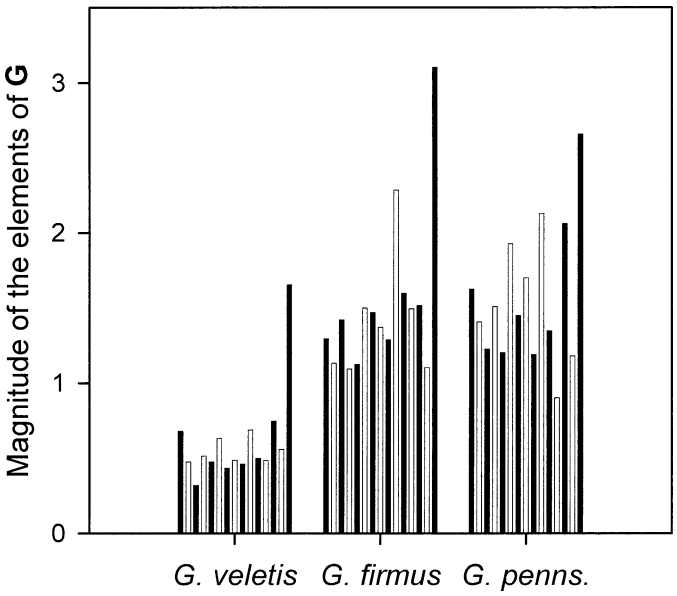

FIG. 1. Comparison of the magnitude of the elements of **G** within and among species *Gryllus veletis*, *G. firmus*, and *G. pennsylvanicus*. Matrix elements were multiplied by 1000 for clarity. Each bar represents an additive genetic variance or an additive genetic covariance. The tallest bar in each species is the variance of ovipositor length (OVIP). Variation in bar color is for visual clarity only. Parameters were estimated by a jackknife procedure on the results of a nested ANOVA or ANCOVA.

individuals (all ANOVAs $P < 0.001$; MANOVA Wilks' $\lambda = 0.60$; approximate $F = 155$; df = 5, 1147; $P < 0.001$).

We tested the significance of the genetic variation within each species. Results from a MANOVA using family and cage nested within family as independent variables revealed that, within a species, the variation among families in all five traits was highly significant (all univariate and multivariate $P < 0.001$). Cage effects were generally significant (10 of the 15 univariate, and two of the three multivariate $P < 0.05$), but generally explained an order of magnitude less variation than the corresponding family effects. The average heritability in each species was in the normal range for morphological traits (Mousseau and Roff 1987), as were the average genetic correlations (Roff 1996; Table 4). There was less genetic and phenotypic variation in *G. veletis* than in the two other species (Tables 5, 6). Figure 1 compares the magnitude of all **G** matrix elements across species and shows that the variance of the trait OVIP (the last bar of each histogram) was the largest element in all three **G** matrices. The laboratory population of *G. veletis* was less phenotypically variable than its wild counterpart (Table 6).

Average genetic correlations were very similar to average phenotypic correlations within each species (Table 4). To further compare these two types of correlation, we plotted the individual genetic correlations against their corresponding phenotypic correlations, using the estimates of all three species in the same plot (Fig. 2A). This graph revealed that these two types of parameters were highly correlated ($r = 0.87$). The same procedure was followed with the elements of the **G** and **P** matrices (Fig. 2B), and a slightly weaker correlation was apparent ($r = 0.74$). Note, however, that the test of statistical significance of such a correlation was not strictly valid because the correlations or (co)variances were not independent of each other; thus, we provided no such probability. All of the above suggested that patterns of genetic and phenotypic variation were similar.

Matrix Comparisons

The T method and the related *T*% statistic indicated that the **G** matrix of *G. veletis* was different from the **G** matrices of the two other species (Table 7). This method is mainly influenced by the magnitude of the **G** matrix elements and the results were thus expected given the smaller variances and covariances found in *G. veletis* (Fig. 1). The Flury hierarchy yielded a pattern of probabilities slightly different from the pattern found with the T method, the main difference

TABLE 4. Heritabilities (h^2, diagonal elements), genetic correlations (r_G, elements above the diagonal), and phenotypic correlations (r_P, elements below the diagonal), followed by their standard error. The average of these parameters is given for each species. Parameters were estimated by a jackknife procedure on the results of a nested ANOVA or ANCOVA, using ln-transformed data.

	FEMUR	HEAD	PTHL	PTHW	OVIP	Average h^2	Average r_G	Average r_P
Gryllus veletis						0.41	0.59	0.59
FEMUR	0.44 (0.08)	0.73 (0.09)	0.48 (0.16)	0.72 (0.07)	0.46 (0.13)			
HEAD	0.76 (0.02)	0.41 (0.07)	0.66 (0.12)	0.72 (0.10)	0.46 (0.14)			
PTHL	0.49 (0.03)	0.54 (0.02)	0.26 (0.06)	0.69 (0.15)	0.46 (0.13)			
PTHW	0.72 (0.02)	0.80 (0.02)	0.51 (0.03)	0.46 (0.10)	0.51 (0.14)			
OVIP	0.56 (0.03)	0.56 (0.03)	0.41 (0.03)	0.57 (0.03)	0.49 (0.09)			
G. firmus						0.52	0.73	0.77
FEMUR	0.48 (0.09)	0.81 (0.05)	0.83 (0.05)	0.78 (0.06)	0.57 (0.15)			
HEAD	0.85 (0.01)	0.48 (0.09)	0.80 (0.07)	0.91 (0.03)	0.61 (0.14)			
PTHL	0.81 (0.02)	0.81 (0.02)	0.59 (0.11)	0.86 (0.05)	0.57 (0.11)			
PTHW	0.86 (0.01)	0.90 (0.01)	0.84 (0.02)	0.46 (0.09)	0.52 (0.17)			
OVIP	0.66 (0.04)	0.71 (0.03)	0.64 (0.03)	0.67 (0.04)	0.56 (0.13)			
G. pennsylvanicus						0.49	0.66	0.61
FEMUR	0.62 (0.16)	0.80 (0.08)	0.67 (0.14)	0.83 (0.08)	0.61 (0.21)			
HEAD	0.77 (0.02)	0.65 (0.12)	0.73 (0.11)	0.86 (0.06)	0.53 (0.19)			
PTHL	0.62 (0.05)	0.62 (0.04)	0.38 (0.10)	0.65 (0.13)	0.37 (0.18)			
PTHW	0.78 (0.03)	0.82 (0.02)	0.63 (0.03)	0.61 (0.14)	0.53 (0.21)			
OVIP	0.52 (0.03)	0.47 (0.04)	0.37 (0.03)	0.46 (0.03)	0.20 (0.09)			

TABLE 5. Additive genetic (co)variance matrix (**G**) for each species, based on ln-transformed data. Each matrix element is followed by its standard error. Matrix elements and standard errors were multiplied by 1000 for clarity. Parameters were estimated by a jackknife procedure on the results of a nested ANOVA or ANCOVA.

	FEMUR	HEAD	PTHL	PTHW	OVIP
Gryllus veletis					
FEMUR	0.68 (0.16)				
HEAD	0.48 (0.13)	0.63 (0.14)			
PTHL	0.32 (0.15)	0.43 (0.13)	0.69 (0.18)		
PTHW	0.52 (0.14)	0.49 (0.15)	0.50 (0.16)	0.75 (0.19)	
OVIP	0.48 (0.20)	0.46 (0.19)	0.49 (0.18)	0.56 (0.23)	1.65 (0.35)
G. firmus					
FEMUR	1.29 (0.27)				
HEAD	1.13 (0.27)	1.50 (0.33)			
PTHL	1.42 (0.35)	1.47 (0.40)	2.28 (0.54)		
PTHW	1.09 (0.28)	1.37 (0.33)	1.60 (0.41)	1.52 (0.35)	
OVIP	1.12 (0.43)	1.29 (0.49)	1.49 (0.52)	1.10 (0.50)	3.10 (0.85)
G. pennsylvanicus					
FEMUR	1.63 (0.53)				
HEAD	1.41 (0.47)	1.93 (0.50)			
PTHL	1.23 (0.52)	1.45 (0.56)	2.13 (0.64)		
PTHW	1.51 (0.54)	1.70 (0.51)	1.35 (0.56)	2.06 (0.60)	
OVIP	1.20 (0.70)	1.19 (0.63)	0.90 (0.50)	1.18 (0.74)	2.66 (1.19)

being that the comparison *G. veletis–G. pennsylvanicus* was not significant with the Flury method. Exploration of the verdicts of this method revealed that *G. veletis* and *G. firmus* shared all of their eigenvectors but not their eigenvalues (common principal components, or CPC model), and the other comparisons were best described by the equality model (Table 7). When all three species were compared simultaneously with the Flury hierarchy, the model equality best explained the similarities between the matrices ($P_{\text{equality}} = 0.06$).

The results of the MANOVA method indicated that none of the comparisons yielded a significant probability (results not shown). However, an exploration of the univariate ANOVA probabilities corresponding to the 15 matrix elements within each pairwise comparison of species (not shown) revealed that, in the case of the *G. veletis–G. firmus* and *G. veletis–G. pennsylvanicus* comparisons, many univariate probabilities were significant (nine and eight of 15, respectively), and most nonsignificant univariate probabilities were between 0.05 and 0.10. This overall significant signal contrasted with the multivariate results and suggested a lack of power at the multivariate level. This type of problem is frequent in MANOVAs when the dependent variables are highly positively correlated (Tabachnick and Fidell 2001, p. 329),

TABLE 6. Phenotypic (co)variance matrix (**P**) for each species, based on ln-transformed data. Each matrix element is followed by its standard error. Matrix elements and standard errors were multiple by 1000 for clarity. Parameters were estimated by a jackknife procedure on the results of a nested ANOVA or ANCOVA.

	FEMUR	HEAD	PTHL	PTHW	OVIP
Gryllus veletis					
FEMUR	1.56 (0.07)				
HEAD	1.19 (0.06)	1.56 (0.07)			
PTHL	0.99 (0.08)	1.10 (0.07)	2.61 (0.11)		
PTHW	1.14 (0.06)	1.27 (0.06)	1.04 (0.07)	1.62 (0.07)	
OVIP	1.28 (0.08)	1.27 (0.08)	1.22 (0.10)	1.34 (0.09)	3.35 (0.16)
G. firmus					
FEMUR	2.69 (0.13)				
HEAD	2.46 (0.12)	3.10 (0.14)			
PTHL	2.59 (0.13)	2.81 (0.14)	3.85 (0.17)		
PTHW	2.54 (0.13)	2.86 (0.13)	2.97 (0.14)	3.26 (0.15)	
OVIP	2.56 (0.15)	2.95 (0.15)	2.93 (0.18)	2.85 (0.16)	5.52 (0.27)
G. pennsylvanicus					
FEMUR	2.64 (0.18)				
HEAD	2.15 (0.17)	2.97 (0.20)			
PTHL	2.35 (0.23)	2.52 (0.25)	5.50 (0.41)		
PTHW	2.33 (0.19)	2.58 (0.21)	2.70 (0.28)	3.35 (0.29)	
OVIP	3.11 (0.36)	2.97 (0.35)	3.19 (0.42)	3.06 (0.38)	13.23 (1.86)
G. veletis field					
FEMUR	2.27 (0.39)				
HEAD	2.05 (0.39)	2.44 (0.42)			
PTHL	1.94 (0.50)	2.01 (0.47)	3.47 (0.76)		
PTHW	1.90 (0.40)	2.11 (0.38)	2.01 (0.51)	2.33 (0.40)	
OVIP	1.56 (0.41)	1.74 (0.42)	0.98 (0.57)	1.60 (0.41)	4.10 (0.76)

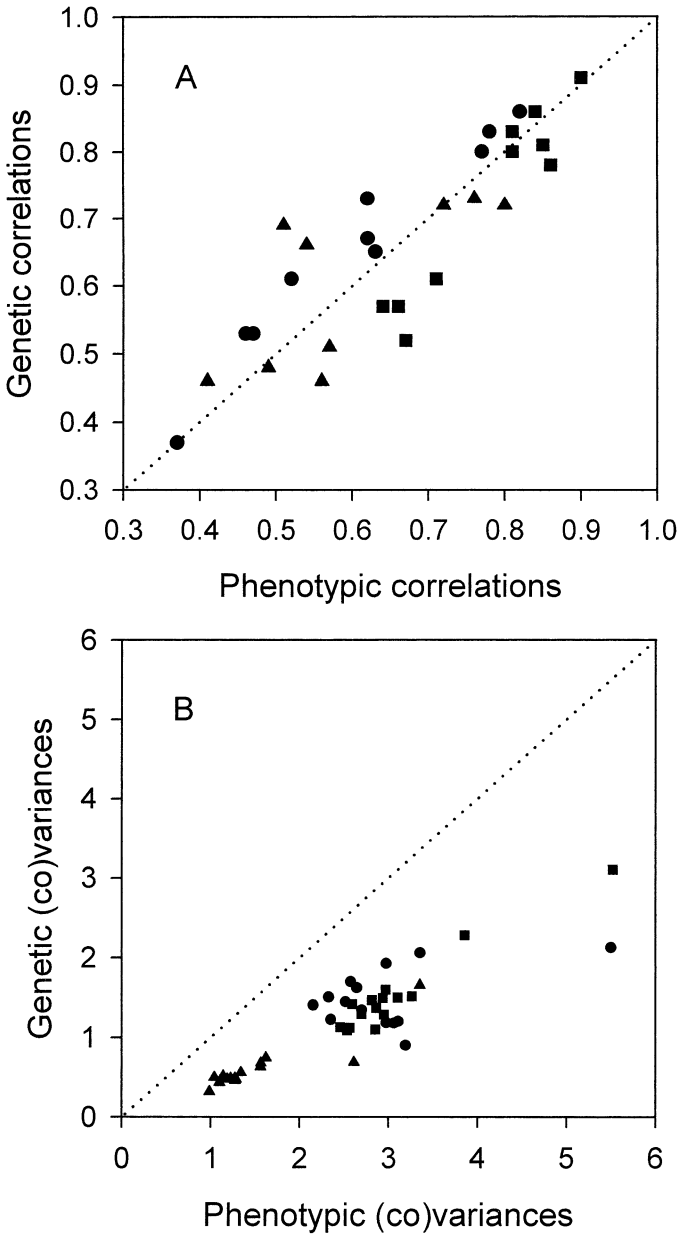

FIG. 2. Comparison of (A) genetic correlations with their corresponding phenotypic correlations and (B) genetic (co)variances with their corresponding phenotypic (co)variances. One datapoint was omitted from panel (B) for clarity (the variance of the ovipositor length, coordinates **P**: 13.2, **G**: 2.7). All (co)variances were multiplied by 1000. Triangles represent *Gryllus veletis*, squares represent *G. firmus*, and circles represent *G. pennsylvanicus*. The dotted line is the 1:1 line. Parameters were estimated by a jackknife procedure on the results of a nested ANOVA or ANCOVA.

as is the case with the **G** matrix pseudovalues. Additionally, the distribution of the pseudovalues often was not normal and several multivariate outliers were found with the Mahalanobis distance statistic (no transformation seemed to make the distribution more adequate for statistical testing). To solve this problem of low multivariate power, we reduced the dimensionality of the dataset by using a principal components analysis on the pseudovalues of the **G** matrix (Tabachnick and Fidell 2001, p. 357) and by only retaining the scores of the first principal component as data. This first principal component explained approximately 70% of the total variation in all three cases, which suggests that only a minor portion of the differences between species would be lost to the data reduction. This procedure was applied to each of the pairwise comparisons (i.e., the principal component analysis was performed on the pooled pseudovalue data from the two species being compared). Results (referred to as MANOVA first PC) were extremely similar to the ones of the T method (Table 7), with *G. veletis* differing from the two other species. The three species were compared simultaneously using this procedure and the results were marginally significant ($P = 0.04$), which is consistent with the result obtained using the Flury hierarchy.

Overall, the two main results that emerged from the comparison of **G** matrices are that: (1) the **G** matrix of *G. veletis* differed from the **G** matrix of the two other species, which is qualitatively consistent with the hypothesis of a phylogenetic structure of **G** matrix variation (Table 1); and (2) the observed differences between species mainly reflected differences in the magnitude of genetic variation, not in the principal components structure. These results were unchanged (not shown) when the trait OVIP, which is the most variable trait (Table 5, Fig. 1), was removed from the analysis.

Comparisons of **P** matrices across laboratory-reared species indicated that matrices generally differed highly significantly from each other (Table 7). The **P** matrices of the field-caught and laboratory-reared samples of *G. veletis* were also found to differ significantly (Table 7). The power available for **P** matrix comparisons appeared to be extremely high, which implied that very small differences could be found to be statistically significant. The two methods based on randomization (the T method and the Flury hierarchy) generally yielded probabilities of zero because no single randomized dataset had a statistic larger than the one of the original dataset. This has the disadvantage of preventing the comparison of probabilities across species comparisons, and therefore prevents the assessment of which matrices are the most different. To investigate relative differences between matrices, we therefore relied only on the *T*% statistic and the *F* statistic of the MANOVA approach, because these statistics have no base limit. These statistics suggested that the **P** matrix of *G. veletis* differed from the **P** matrices of the other two species, whereas the two latter were relatively similar to each other (Table 7), which is in accordance with the results of **G** matrix comparisons. By contrast, the Flury hierarchy yielded very different verdicts for the **P** and **G** analyses. This result was probably not informative, however, because the determination of the verdict by the Flury hierarchy is heavily determined by degrees of freedom, a problem acknowledged by Phillips and Arnold (1999, p. 1513) and reported in several other studies (Steppan 1997a; Ackermann and Cheverud 2000; Marroig and Cheverud 2001). Table 7 also revealed that the **P** matrix of laboratory-reared *G. veletis* differed much less from the **P** matrix of field-caught *G. veletis* than from the **P** matrices of the two other species.

TABLE 7. Comparisons of the **G** and **P** matrices of pairs of species using the T method, the Flury hierarchy, and the MANOVA method. All probabilities correspond to the null hypothesis of no difference between matrices. The *T*% statistic is the absolute average percent difference between two matrices; the higher the value of *T*%, the larger the difference. The verdict of the Flury hierarchy is the model that best explains the difference between two matrices. The probabilities given for the Flury hierarchy correspond to the test of equality of two matrices, not necessarily to the verdict. In the case of the comparison of laboratory and field individuals of *Gryllus veletis*, the parametric version of the Flury hierarchy was used. The method MANOVA (first PC) refers to the results obtained by using the scores of the first principal component of the pseudovalues of the **G** matrices. The numbers in parentheses for the MANOVA method correspond to the *F*-statistic and are given to provide an idea of the magnitude of the differences between pairs of matrices when probabilities are below 0.001. The results of the T and MANOVA methods were based on estimations of the **G** matrix that are not inflated by common environmental effects, as opposed to the results of the Flury hierarchy.

Comparison	*T* *P*	*T*%	Flury hierarchy *P*	Flury hierarchy verdict	MANOVA (first PC) *P*
G matrices					
G. veletis–G. firmus	0.007	85.7	0.01	CPC	0.01
G. veletis–G. pennsylvanicus	0.007	88.2	0.13	equality	0.02
G. firmus–G. pennsylvanicus	0.89	18.3	0.39	equality	0.86
P matrices					
G. veletis–G. firmus	0	68.3	0	PCPC 1	<0.001 (139)
G. veletis–G. pennsylvanicus	0	83.2	0	unrelated	<0.001 (84)
G. firmus–G. pennsylvanicus	0.02	24.4	0	unrelated	0.94 (0.005)
G. veletis lab–*G. veletis* field	0.02	38.0	0.04	CPC	0.01 (7)

Comparison of $\mathbf{g}_{max}$ and $\Delta\bar{\mathbf{z}}$

The vector $\mathbf{g}_{max}$ for each species and the vector $\Delta\bar{\mathbf{z}}$ for each pair of species are shown in Table 8. The values of θ and of r_V corresponding to the comparison of the direction of the vectors $\mathbf{g}_{max}$ and $\Delta\bar{\mathbf{z}}$ in each pairwise comparison of species are given in Table 9. Note that two values each of θ and r_V are given for each comparison of species, the reason being that the calculation of θ and r_V are based on a single $\mathbf{g}_{max}$. We therefore estimated the angle between $\Delta\bar{\mathbf{z}}$ and each of the two $\mathbf{g}_{max}$ separately (one for each species). In Table 9, each estimation of θ and r_V were based on the $\mathbf{g}_{max}$ of the first species listed. Table 9 indicated that, for a given pairwise comparison of species, θ and r_V were not greatly influenced by the identity of the species from which $\mathbf{g}_{max}$ was taken, which was to be expected given that the **G** matrices of all species were found to share a common principal component structure (see above). In the two comparisons that included *G. firmus*, the direction of $\Delta\bar{\mathbf{z}}$ was similar to the direction of $\mathbf{g}_{max}$, although sometimes statistically distinguishable from it (Table 9). By contrast, the direction differed much more in the case of the *G. veletis–G. pennsylvanicus* comparison. Notice that the probability of rejecting the null hypothesis differed depending on which $\mathbf{g}_{max}$ is bootstrapped, with *G. firmus* being the least variable in its $\mathbf{g}_{max}$.

To illustrate the comparison of the direction of $\mathbf{g}_{max}$ and $\Delta\bar{\mathbf{z}}$, the dataset had to be reduced in dimensionality from five to two dimensions. All bivariate combinations of traits are plotted in Figure 3, on which the mean phenotypic value and the $\mathbf{g}_{max}$ of each species are given for each combination of traits. For Figure 3 only, the direction of $\mathbf{g}_{max}$ for a pair of traits was estimated with a principal component analysis on the two genetic variances and the genetic covariance corresponding to the two traits, not the whole **G** matrix. The slope of $\mathbf{g}_{max}$ is equal to the quotient of the eigenvector loading of the two traits. Because this procedure uses only two traits at a time, the direction of $\mathbf{g}_{max}$ thus obtained does not correspond exactly to the direction of the five-dimensional vector, but it is a relatively close approximation. It can be seen from the approximate parallelism of the direction of $\mathbf{g}_{max}$ that the orientation of this vector, and thus of **G,** was relatively similar across species (Fig. 3), confirming the results of the comparison of **G** matrices (see above). Second, Figure 3 reveals that, in most cases, there was some discrepancy between the different bivariate $\mathbf{g}_{max}$ and the bivariate direction of phe-

TABLE 8. Estimation of the vectors $\Delta\bar{\mathbf{z}}$, the multivariate difference between the means of two species, and $\mathbf{g}_{max}$, the multivariate direction of greatest quantitative genetic variation within a species. Each of the five numbers within the vector $\Delta\bar{\mathbf{z}}$ corresponds to a standardized difference in the phenotypic mean of a trait between two species. Each of the five numbers within the vector $\mathbf{g}_{max}$ corresponds to the loading of a trait on the first eigenvector of the **G** matrix of a species. The vector $\mathbf{g}_{max}$ was estimated from a principal component analysis on **G,** which was estimated by a jackknife procedure on the results of a nested ANOVA or ANCOVA.

	FEMUR	HEAD	PTHL	PTHW	OVIP
$\Delta\bar{\mathbf{z}}$					
G. veletis–G. firmus	0.549	0.361	0.386	0.344	0.549
G. veletis–G. pennsylvanicus	0.599	0.156	−0.096	0.344	0.699
G. firmus–G. pennsylvanicus	0.465	0.423	0.581	0.307	0.416
$\mathbf{g}_{max}$					
G. veletis	0.374	0.375	0.368	0.430	0.632
G. firmus	0.370	0.413	0.510	0.408	0.516
G. pennsylvanicus	0.425	0.470	0.429	0.478	0.431

TABLE 9. Estimation of the angle θ in degrees between the vectors $\Delta\bar{\mathbf{z}}$ and $\mathbf{g}_{max}$ and of the correlation between the two same vectors for all pairwise comparisons of species. Two θ and r_V are given for each species comparison because calculations are based on the $\mathbf{g}_{max}$ of the species that is listed first.

Comparison	θ	r_V
G. veletis–G. firmus	12.2	0.98
G. firmus–G. veletis	13.4*	0.97
G. veletis–G. pennsylvanicus	33.2**	0.84
G. pennsylvanicus–G. veletis	41.2***	0.75
G. firmus–G. pennsylvanicus	10.6*	0.98
G. pennsylvanicus–G. firmus	13.7	0.97

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

notypic divergence across species. It can also be observed that the cases in which the direction of $\mathbf{g}_{max}$ and $\Delta\bar{\mathbf{z}}$ are the less similar are the ones that included PTHL for *G. veletis* and *G. pennsylvanicus*. This figure thus helped identify the cause of the large θ value obtained for the *G. veletis–G. pennsylvanicus* comparison (Table 9).

Overall, the results of the comparison of the direction of $\mathbf{g}_{max}$ and $\Delta\bar{\mathbf{z}}$ suggested that these two vectors are similarly oriented, except for one of the five dimensions of the *G. veletis–G. pennsylvanicus* comparison.

DISCUSSION

The conclusions of this study are dependent on the assumption that the parameters estimated in the laboratory are representative of natural population parameters. We have several lines of evidence to support such an assumption. First, the relative size of the three species of crickets reared in the laboratory for the current study reflected their relative size in nature (Alexander 1957): *Gryllus firmus* is the largest of these species, and *G. veletis* and *G. pennsylvanicus* are of similar size. Concerning the estimates of **G,** and therefore of $\mathbf{g}_{max}$, Bégin and Roff (2001) have shown for *G. pennsylvanicus* that the **G** matrix of individuals reared in the laboratory was not different from the **G** matrix of the same families reared in plastic buckets placed in the field. In the current study, we showed that the **P** matrix of *G. veletis* reared in the laboratory was more similar to the **P** matrix of wild-caught *G. veletis* than to the **P** matrix of the other two species (Table 7). Additionally, Weigensberg and Roff (1996) showed that heritabilities estimated in the laboratory are often similar to their natural counterpart. We therefore believe that our laboratory estimates provide a good representation of natural population parameters.

Matrix Comparisons

This study is only the second (after Lofsvold 1986) to use more than two well-differentiated taxonomic groups (species or subspecies) in a study of the variation of **G.** The first important result obtained in the current study is that the **G** matrix of *G. veletis*, which is the most distantly related of the three species (Table 1), is different from the matrices of *G. pennsylvanicus* and of *G. firmus*, whereas the two latter are very similar. This result is qualitatively consistent with the hypothesis that differences in **G** matrices increase with genetic distance, that is, that the pattern of **G** matrix variation is phylogenetically structured. However, the material available in the present study does not allow great confidence in the inference of a phylogenetic pattern because only three species were sampled and because only one population was used as a representative of its species. The observed structure of the **G** matrix variation could have several possible causes. It could indeed reflect a clocklike accumulation of changes with time as expected by the phylogenetic hypothesis. Alternatively, it could be the result of some event (selection or drift) that caused an evolutionary episode in *G. veletis* and made it diverge from the other two species. Finally, it could result from a population effect such as inbreeding that may have affected the genetic variances of the particular population of *G. veletis* that was sampled for this study. Replication and denser phylogenetic investigation of **G** matrix variation are needed to test these alternatives.

Using **P** as a surrogate for **G** is a very attractive alternative (Cheverud 1988; Roff 1997) because it greatly reduces the amount of work required and therefore allows the estimation of the covariance matrix of more groups. Several studies have tested the hypothesis of a phylogenetic structure of **P** matrices, but no clear pattern yet emerges. A few studies have shown evidence for some phylogenetic pattern (Goodin and Johnson 1992; Steppan 1997b; Ackermann and Cheverud 2000), whereas others have found no pattern (Cheverud 1989; Badyaev and Hill 2000; Marroig and Cheverud 2001). However, little is known about the general validity of replacing **G** by **P** (Willis et al. 1991). In the present study, we compared the pattern of **G** matrix variation to the pattern of **P** matrix variation and found similarities. This provides support for the use of **P** instead of **G** in the case of morphological traits in crickets.

An alternative framework for the study of **G** matrix variation is to assume that changes in **G** are associated with adaptation to some ecological variables (which may or may not be compatible with a phylogenetic hypothesis). Only one study has made such an investigation and showed, for two populations of amphipods, that the effect of the adaptation to habitat type on **G** matrices was larger than the effect of phylogeny (Fong 1989; re-analyzed in Roff 2002). Additionally, a few studies have looked at the effect of adaptation to environmental variables on **P** matrix variation. These studies found a correlation between **P** matrix variation and feeding habits in new world monkeys (Marroig and Cheverud 2001), no geographical patterns of **P** matrix variation in aphids (Riska 1985), and a geographical pattern in mangroves (Dodd et al. 2000). In the present study, if adaptation to ecological variables had been important in shaping **G** matrix differences in crickets, we would have expected *G. firmus* to have a **G** matrix different from both *G. veletis* and *G. pennsylvanicus* because the latter two species are very similar morphologically, are found in the same microhabitats (Alexander and Bigelow 1960; Alexander 1968), and are under similar selective pressures for a phosphoglucose isomerase electromorph (Harrison 1977; Katz and Harrison 1997). Because we did not observe a pattern consistent with ecological information, we rejected the hypothesis that **G** matrix variation is determined by an adaptation to the current ecology of these species. The morphological similarity between *G. veletis* and *G. pennsylvanicus* may result from character con-

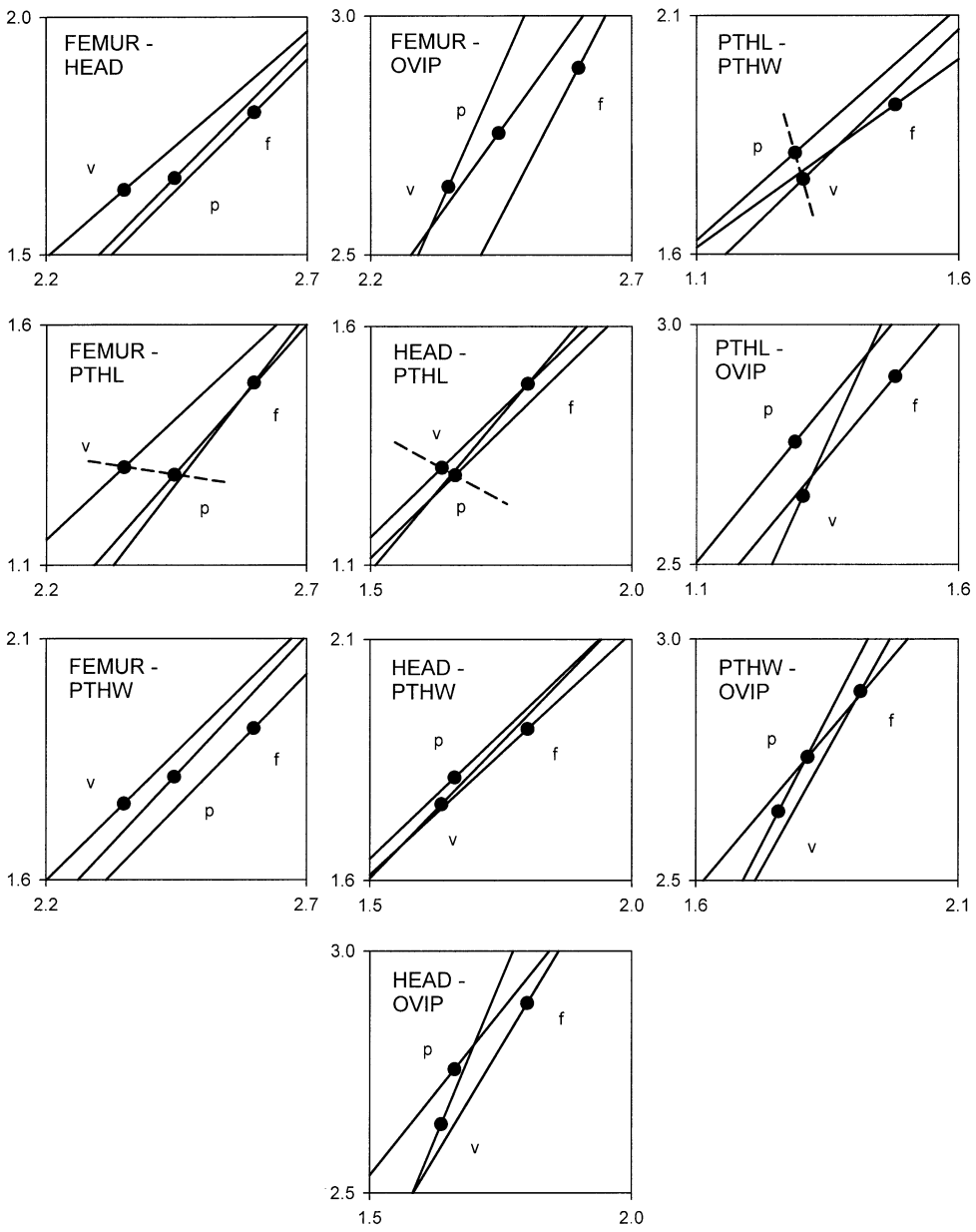

FIG. 3. Two-dimensional comparisons of the direction of greatest genetic variation within each species ($\mathbf{g}_{max}$) to the direction of phenotypic divergence between species ($\mathbf{\Delta}\bar{\mathbf{z}}$). Each graph plots the mean phenotypic value in millimeters (black dots) of a pair of traits for each of the three species. In each plot, the first trait named is on the *x*-axis and the second trait is on the *y*-axis. The solid line extending on each side of the species means corresponds to $\mathbf{g}_{max}$, which was calculated from a principal component analysis on the two corresponding traits only. The direction of $\mathbf{\Delta}\bar{\mathbf{z}}$ (dotted line) is illustrated only in the case where its slope is negative, and therefore clearly different from the direction of $\mathbf{g}_{max}$. The direction of $\mathbf{\Delta}\bar{\mathbf{z}}$ in all other cases can be visualized by linking pairs of mean trait values. v, *Gryllus veletis*; p, *G. pennsylvanicus*; and f, *G. firmus*.

vergence controlled by different genetic mechanisms. It therefore appears that, for this study, knowledge of the phylogeny allows a better prediction of **G** matrix variation than does knowledge of the current habitat and selection pressures. More work is needed to see if this conclusion reflects a general pattern in field crickets.

The second important result found in the current study is that the observed differences in **G** across species were not large and seemed to be mainly caused by a lower genetic variation in *G. veletis*, not by important structural differences. In particular the **G** matrices of the two most closely related species, *G. pennsylvanicus* and *G. firmus*, were statistically indistinguishable, which confirms a previous result (Bégin and Roff 2001). The above evidence therefore suggests that little differentiation has occurred in the quantitative genetic architecture of size traits in these three congeneric species, despite significant changes in mean trait values. Reviews of the literature (Arnold and Phillips 1999; Roff 2000; Steppan et al. 2002) indicated that **G** matrices are often conserved across populations or closely related species. If this trend is

real, the assumption of a constant **G** matrix (Lande 1979) could be valid at low taxonomic levels, and the predictive power of quantitative genetics could potentially be realized for these intermediate time scales. However, the reliability of such predictions still needs to be verified because numerous theoretical arguments have been advanced to suggest that they may be misleading (e.g., Riska 1989; Shaw et al. 1995; Agrawal et al. 2001).

Comparing **G** matrices is a central focus of quantitative genetic research. However, the comparison of matrices is a complex problem and the statistical tools currently available are not ideal (Steppan et al. 2002). Using different methods within a study was therefore important because no single method is likely to provide all the relevant information. In the present study, the results from the T and MANOVA methods were extremely similar and differed slightly from the results of the Flury hierarchy. The only previous comparison of the T method and Flury hierarchy (Bégin and Roff 2001) empirically showed that the two methods provide sometimes similar, sometimes different results for a given **G** matrix comparison. The T method is mainly influenced by the magnitude of the elements of the matrices, whereas the Flury hierarchy uses a more complete approach that incorporates both the magnitude and the structure of the variation (Bégin and Roff 2001). Roff (2002) empirically compared the Flury hierarchy with the results of the MANOVA method and concluded that these two approaches provide qualitatively similar probabilities for the test of matrix equality. The greatest strength of the MANOVA method is that any type of variable (e.g., latitude, time since divergence, population size, habitat type nested within species) can be incorporated into an analysis of matrix variation, making this method ideal to answer various ecological and evolutionary questions (Roff 2002). In the present study, low power issues made the results of the initial MANOVA uninteresting (low power may also be a problem with the two other approaches, but is more difficult to investigate than with the MANOVA method). To get around this power problem, we introduced an apparently successful modification to the protocol of the MANOVA method. The modification consists of using the scores of the first principal component of the pseudovalues. Very little is known about the properties of the various matrix comparison methods (Steppan et al. 2002), and simulation studies are needed (e.g., Houle et al. 2002).

Comparison of $\mathbf{g}_{max}$ and $\Delta\bar{\mathbf{z}}$

Because the **G** matrices of these three cricket species all share a common principal component structure, and thus can be assumed to be similar to the ancestral **G** matrix, the comparison of $\mathbf{g}_{\max}$ and $\Delta\bar{\mathbf{z}}$ may provide insight into the importance of quantitative genetic constraints during phenotypic evolution (but see next paragraph). The present study is the first one in which the direction of $\mathbf{g}_{\max}$ and $\Delta\bar{\mathbf{z}}$ is compared using the **G** matrix of each of the studied species. We found that the angle between the two vectors was small in the cases of the *G. firmus–G. veletis* and *G. firmus–G. pennsylvanicus* comparisons, which means that most of the trait evolution between these species has occurred in a direction approximately consistent with $\mathbf{g}_{\max}$. The larger angle corresponding to the *G. veletis–G. pennsylvanicus* was mostly caused by only one of the five traits (prothorax length or PTHL) and may therefore also provide support for the importance of $\mathbf{g}_{\max}$ in shaping phenotypic evolution. Moreover, it is predicted that evolution in a direction not compatible with $\mathbf{g}_{\max}$ will proceed more slowly than evolution along the genetic line of least resistance (Lande 1979; Schluter 1996). In the current study, the difference in the mean value of the trait prothorax length between *G. veletis* and *G. pennsylvanicus* is small compared to the difference corresponding to all other pairs of traits (Table 8). This result suggests that the divergence of the trait prothorax length between *G. veletis* and *G. pennsylvanicus* has occurred at a lower rate because this trait has evolved in a direction that does not correspond to $\mathbf{g}_{\max}$. Overall, $\mathbf{g}_{\max}$ seems to provide a relatively good prediction of phenotypic evolution in crickets. However, the characterization of species differentiation and of the ancestral **G** matrix is based on only three datapoints, which may or may not provide an adequate picture of the evolution of this group. A similar result of concordance of phenotypic evolution to the covariance matrix has been found in some other studies (Mitchell-Olds 1996; Badyaev and Foresman 2000; see review in Schluter 2000), but the opposite pattern of no concordance has also been found (Venable and Burquez 1990; Mitchell-Olds 1996; Merilä and Bjorklund 1999; Badyaev and Hill 2000).

However, this type of analysis is only correlative and does not differentiate between several hypotheses that explain the concordance of $\mathbf{g}_{\max}$ and $\Delta\bar{\mathbf{z}}$. The first possible explanation is the classical quantitative genetic view of the **G** matrix as a constraint that redirects and slows down the effect of selection (Arnold 1992), a hypothesis that is sensible mostly in the case of complex adaptive landscapes where small deviations in the evolutionary trajectory can substantially affect the outcome (Lande 1979; Bürger 1986; Price et al. 1993). The second possible hypothesis is that population differentiation is caused by random genetic drift. The model corresponding to this situation predicts that phenotypic evolution will occur more frequently in the direction that is most genetically variable (Lande 1979; Arnold et al. 2001). A third possibility is that the concordance between **G** and $\Delta\bar{\mathbf{z}}$ results from the evolutionary response of quasineutral traits that are correlated to a trait under selection (Lande 1979). These three general models posit that **G** acts as a genetic constraint. Alternatively, if the **G** matrix is directly shaped by the long-term pattern of correlational selection (Cheverud 1982, 1984; Arnold 1992; Deng et al. 1999) and if phenotypic evolution occurs mostly along the selective line of least resistance, that is, peak shifts occur along the principal axis of the adaptive landscape (Arnold et al. 2001), the similarity in the direction of $\mathbf{g}_{\max}$ and $\Delta\bar{\mathbf{z}}$ would not represent a causal relation between the two, but rather two separate results of the same causal factor: selection. Distinguishing between these explanations is difficult because detailed information on the **G** matrix and adaptive landscape are required. No investigation of this type has yet been published, but such studies could be instrumental in understanding the relative importance of selection and genetic constraints in shaping phenotypic evolution.

Conclusion

This study provided some support for the validity of the predictive power of quantitative genetics over evolutionary time scales. We have shown that the **G** matrix of three cricket species is relatively conserved despite significant phenotypic trait divergence. We have also shown that the small observed **G** matrix differentiation between species appeared to be qualitatively consistent with a phylogenetic pattern, but is not predictable using ecological information. Finally, we have shown that the direction of phenotypic evolution can be partly predicted from the **G** matrix, whether or not a causal relation exists.

Acknowledgments

We thank A. Simons for providing the *Gryllus pennsylvanicus* data, G. Orti for providing the genetic distance data, and McGill University for allowing access to the Gault estate for sampling. Invaluable laboratory assistance was provided by E. Geoffroy, R. Roff, and A. Mejia. Assistance with statistical programming was provided by P. Phillips and K. Emerson. This work benefited from discussions with D. Reznick, D. Houle, P. Phillips, S. Arnold, and an anonymous reviewer. M. Bégin was funded by postgraduate scholarships from the National Sciences and Engineering Council of Canada and from the Fonds Québecois de la Recherche sur la Nature et les Technologies. D. A. Roff was funded by an operating grant from the National Sciences and Engineering Council of Canada.

Literature Cited

Ackermann, R. R., and J. M. Cheverud. 2000. Phenotypic covariance structure in tamarins (genus *Saguinus*): a comparison of variation patterns using matrix correlation and common principal component analysis. Am. J. Phys. Anthropol. 111:489–501.

Agrawal, A. F., E. D. I. Brodie, and L. H. Rieseberg. 2001. Possible consequences of genes of major effect: transient changes in the **G**-matrix. Genetica 112/ 113:33–43.

Alexander, R. D. 1957. The taxonomy of the field crickets of the eastern United States (Orthoptera: Gryllidae: *Acheta*). Ann. Entomol. Soc. Am. 50:584–602.

———. 1968. Life cycle origins, speciation, and related phenomena in crickets. Q. Rev. Biol. 43:1–41.

Alexander, R. D., and R. S. Bigelow. 1960. Allochronic speciation in field crickets, and a new species, *Acheta veletis*. Evolution 14:334–346.

Arnold, S. J. 1992. Constraints on phenotypic evolution. Am. Nat. 140:S85–S107.

Arnold, S. J., and P. C. Phillips. 1999. Hierarchical comparison of genetic variance-covariance matrices. II. Coastal-inland divergence in the garter snake, *Thamnophis elegans*. Evolution 53: 1516–1527.

Arnold, S. J., M. E. Pfrender, and A. G. Jones. 2001. The adaptive landscape as a conceptual bridge between micro- and macroevolution. Genetica 112/113:9–32.

Badyaev, A. V., and K. R. Foresman. 2000. Extreme environmental change and evolution: stress-induced morphological variation is strongly concordant with patterns of evolutionary divergence in shrew mandibles. Proc. R. Soc. Lond. B. Biol. Sci. 267:371–377.

Badyaev, A. V., and G. E. Hill. 2000. The evolution of sexual dimorphism in the house finch. I. Population divergence in morphological covariance structure. Evolution 54:1784–1794.

Bégin, M., and D. A. Roff. 2001. An analysis of **G** matrix variation in two closely related cricket species, *Gryllus firmus* and *G. pennsylvanicus*. J. Evol. Biol. 14:1–13.

Bigelow, R. S. 1960. Developmental rates and diapause in *Acheta pennsylvanicus* (Burmeister) and *Acheta veletis* (Alexander and Bigelow) (Orthoptera: Gryllidae). Can. J. Zool. 38:973–988.

Bürger, R. 1986. Constraints for the evolution of functionally coupled characters: a nonlinear analysis of a phenotypic model. Evolution 40:182–193.

Cheverud, J. M. 1982. Phenotypic, genetic and environmental morphological integration in the cranium. Evolution 36:499–516.

———. 1984. Quantitative genetics and developmental constraints on evolution by selection. J. Theor. Biol. 110:155–171.

———. 1988. A comparison of genetic and phenotypic correlations. Evolution 42:958–968.

———. 1989. A comparative analysis of morphological variation patterns in the papionins. Evolution 43:1737–1747.

Cheverud, J. M., and L. J. Leamy. 1985. Quantitative genetics and the evolution of ontogeny. III. Ontogenetic changes in correlation structure among live-body traits in random-bred mice. Genet. Res 46:325–335.

Deng, H.-W., V. Haynatzka, K. Spitze, and G. Haynatzki. 1999. The determination of genetic covariances and prediction of evolutionary trajectories based on a genetic correlation matrix. Evolution 53:1592–1599.

Dodd, R. S., Z. Afzal Rafii, and A. Bousquet-Mélou. 2000. Evolutionary divergence in the pan-Atlantic magrove *Avicennia germinans*. New Phytol. 145:115–125.

Falconer, D. S., and T. F. C. Mackay. 1996. Introduction to quantitative genetics. Longman, Harlow, U.K.

Fong, D. W. 1989. Morphological evolution of the amphipod *Gammarus minus* in caves: quantitative genetic analysis. Am. Midl. Nat. 121:361–378.

Goodin, J. T., and M. S. Johnson. 1992. Patterns of morphological covariation in *Partula*. Syst. Biol. 41:292–304.

Grant, P. R., and B. R. Grant. 1995. Predicting microevolutionary responses to directional selection on heritable variation. Evolution 49:241–251.

Harrison, R. G. 1977. Parallel variation at an enzyme locus in sibling species of field crickets. Nature 266:168–170.

Harrison, R. G., and J. Arnold. 1982. A narrow hybrid zone between closely related cricket species. Evolution 36:535–552.

Houle, D., J. Mezey, and P. Galpern. 2002. Interpretation of the results of common principal components analyses. Evolution 56: 433–440.

Huang, Y., G. Orti, M. Sutherlin, A. Duhachek, and A. Zera. 2000. Phylogenetic relationships of North American field crickets inferred from mitochondrial DNA data. Mol. Phylogenet. Evol. 17:48–57.

Katz, L. A., and R. G. Harrison. 1997. Balancing selection on electrophoretic variation of phosphoglucose isomerase in two species of field cricket: *Gryllus veletis* and *G. pennsylvanicus*. Genetics 147:609–621.

Kohn, L. A. P., and W. R. Atchley. 1988. How similar are genetic correlation structures? Data from mice and rats. Evolution 42: 467–481.

Lande, R. 1976. Natural selection and random genetic drift in phenotypic evolution. Evolution 30:314–343.

———. 1979. Quantitative genetic analysis of multivariate evolution applied to brain:body size allometry. Evolution 33: 402–416.

———. 1980. The genetic covariance between characters maintained by pleiotropic mutations. Genetics 94:203–215.

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

Lofsvold, D. 1986. Quantitative genetics of morphological differentiation in *Peromyscus*. I. Tests of the homogeneity of genetic covariance structure among species and subspecies. Evolution 40:559–573.

Manly, B. F. J. 1997. Randomization, bootstrap and Monte Carlo methods in biology. 2d ed. Chapman and Hall, New York.

Marroig, G., and J. M. Cheverud. 2001. A comparison of phenotypic variation and covariation patterns and the role of phylogeny, ecology, and ontogeny during cranial evolution of New World monkeys. Evolution 55:2576–2600.

Merilä, J., and M. Bjorklund. 1999. Population divergence and morphometric integration in the greenfinch (*Carduelis chloris*); Evo-

lution against the trajectory of least resistance? J. Evol. Biol. 12:103–112.
Mitchell-Olds, T. 1996. Pleiotropy causes long-term genetic constraints on life-history evolution in *Brassica rapa*. Evolution 50: 1849–1858.
Mousseau, T. A., and D. A. Roff. 1987. Natural selection and the heritability of fitness components. Heredity 59:181–197.
Phillips, P. C. 1998a. CPCrand: randomization test of the CPC hierarchy. Univ. of Oregon, Eugene. Software available at darkwing.uoregon.edu/~pphil/software.html.
———. 1998b. CPC: common principal components analysis. Univ. of Oregon, Eugene. Software available at darkwing.uoregon.edu/~pphil/software.html.
Phillips, P. C., and S. J. Arnold. 1999. Hierarchical comparison of genetic variance-covariance matrices. I. Using the Flury hierarchy. Evolution 53:1506–1515.
Price, T., M. Turelli, and M. Slatkin. 1993. Peak shifts produced by correlated response to selection. Evolution 47:280–290.
Reeve, J. P. 2000. Predicting long-term response to selection. Genet. Res. 75:83–94.
Riska, B. 1985. Group size factors and geographic variation of morphometric correlation. Evolution 39:792–803.
———. 1989. Composite traits, selection response, and evolution. Evolution 43:1172–1191.
Roff, D. A. 1996. The evolution of genetic correlations: an analysis of patterns. Evolution 50:1392–1403.
———. 1997. Evolutionary quantitative genetics. Chapman and Hall, New York.
———. 1998. Effects of inbreeding on morphological and life history traits of the sand cricket, *Gryllus firmus*. Heredity 81:28–37.
———. 2000. The evolution of the **G** matrix: selection or drift? Heredity 84:135–142.
———. 2002. Comparing **G** matrices: a MANOVA approach. Evolution 56:1286–1291.
Roff, D. A., and R. Preziosi. 1994. The estimation of the genetic correlation: the use of the jackknife. Heredity 73:544–548.
Roff, D. A., T. A. Mousseau, and D. J. Howard. 1999. Variation in genetic architecture of calling song among populations of *Allonemobius socius*, *A. fasciatus*, and a hybrid population: drift or selection? Evolution 53:216–224.
Schluter, D. 1996. Adaptive radiation along genetic lines of least resistance. Evolution 50:1766–1774.
———. 2000. The ecology of adaptive radiation. Oxford Univ. Press, Oxford, U.K.
Shaw, F. H., R. G. Shaw, G. S. Wilkinson, and M. Turelli. 1995. Changes in genetic variances and covariances: **G** whiz? Evolution 49:1260–1267.
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.
Steppan, S. J. 1997a. Phylogenetic analysis of phenotypic covariance structure. I. Contrasting results from matrix correlation and common principal component analyses. Evolution 51:571–586.
———. 1997b. Phylogenetic analysis of phenotypic covariance structure. II. Reconstructing matrix evolution. Evolution 51: 587–594.
Steppan, S. J., P. C. Phillips, and D. Houle. 2002. Comparative quantitative genetics: evolution of the **G** matrix. Trends Ecol. Evol. 17:320–327.
Tabachnick, B. G., and L. S. Fidell. 2001. Using multivariate statistics. 4th ed. Allyn and Bacon, Boston.
Turelli, M. 1988. Phenotypic evolution, constant covariances and the maintenance of additive variance. Evolution 42:1342–1347.
Venable, D. L., and M. A. Burquez. 1990. Quantitative genetics of size, shape, life-history, and fruit characteristics of the seed heteromorphic composite *Heterosperma pinnatum*. II. Correlation structure. Evolution 44:1748–1763.
Via, S., and R. Lande. 1985. Genotype-environment interaction and the evolution of phenotypic plasticity. Evolution 39:505–522.
Weigensberg, I., and D. A. Roff. 1996. Natural heritabilities: Can they be reliably estimated in the laboratory? Evolution 50: 2149–2157.
Willis, J. H., J. A. Coyne, and M. Kirkpatrick. 1991. Can one predict the evolution of quantitative characters without genetics? Evolution 45:441–444.
Zeng, Z.-B. 1988. Long-term correlated response, interpopulation covariation, and interspecific allometry. Evolution 42:363–374.

Corresponding Editor: P. Phillips