The evolution of the phenotypic covariance matrix: evidence for selection and drift in *Melanoplus*

D. A. ROFF* & T. MOUSSEAU†

*Department of Biology, University of California, Riverside, CA, USA †Department of Biological Sciences, University of South Carolina, Columbia, SC, USA

Keywords:

Flury hierarchy; **G** matrix; Jackknife–MANOVA method; matrix comparisons; **P** matrix; phenotypic evolution; quantitative genetics.

Abstract

Phenotypic variation in trait means is a common observation for geographically separated populations. Such variation is typically retained under common garden conditions, indicating that there has been evolutionary change in the populations, as a result of selection and/or drift. Much less frequently studied is variation in the phenotypic covariance matrix (hereafter, **P** matrix), although this is an important component of evolutionary change. In this paper, we examine variation in the phenotypic means and **P** matrices in two species of grasshopper, *Melanoplus sanguinipes* and *M. devastator*. Using the **P** matrices estimated for 14 populations of *M. sanquinipes* and three populations of *M. devastator* we find that (1) significant differences between the sexes can be attributed to scaling effects; (2) there is no significant difference between the two species; (3) there are highly significant differences among populations that cannot be accounted for by scaling effects; (4) these differences are a consequence of statistically significant patterns of covariation with geographic and environmental factors, phenotypic variances and covariances increasing with increased temperature but decreasing with increased latitude and altitude. This covariation suggests that selection has been important in the evolution of the \mathbf{P} matrix in these populations Finally, we find a significant positive correlation between the average difference between matrices and the genetic distance between the populations, indicating that drift has caused some of the variation in the **P** matrices.

Introduction

There is abundant evidence that populations of the same species differ phenotypically and that such variation is frequently correlated with some geographical variable such as latitude or with an environmental variable such as temperature (e.g. Masaki, 1967; Mousseau & Roff, 1989; Smith *et al.*, 1994; Blanckenhorn & Fairbairn, 1995; Nylin *et al.*, 1996; Conover *et al.*, 1997; Arnett & Gotelli, 1999; Tracy, 1999; Merila *et al.*, 2000; McKay *et al.*, 2001; Thomas *et al.*, 2001). The former type of variable, and in some cases the latter, is not likely to be itself the agent causing variation but rather an indicator variable of some other factor that directly exerts a selective pressure on the

Correspondence: Derek A. Roff, Department of Biology, University of California, Riverside, CA 92521, USA. Tel.: 951 827 2437; fax: 951 827 5903; e-mail: derek.roff@ucr.edu populations. Phenotypic differences among populations are very often maintained when individuals derived from these populations are grown under common garden conditions (see previous citations), indicating that such variation has a genetic basis. Whereas considerable attention has been devoted to measuring and discussing variation in mean trait values, relatively little attention has been given to variation in the phenotypic variances and covariances (Steppan, 1997; Ackerman & Cheverud, 2000). This is unfortunate, because these statistics are an integral component of the description of variation within and among populations and species. If mean trait values are expected to evolve there seems little reason to suppose that the variances and covariances will not also be subject to selection. Selection and drift act upon the phenotype and hence also the phenotypic variances and covariances, and thus any study of evolutionary change must address both the change in phenotypic means and the change in the phenotypic variance-covariance structure.

The evolution of the mean phenotype can be modelled using the multivariate extension of the breeders equation $\Delta \bar{z} = \mathbf{G} \mathbf{P}^{-1} \mathbf{S}$ where $\Delta \bar{z}$ is the vector of mean responses, G is the matrix of additive genetic variances and covariances, P is the matrix of phenotypic variances and covariances and S is the vector of selection differentials (Lande, 1979). It is evident from the forgoing equation that evolutionary change is a function of the phenotypic variance-covariance matrix (hereafter the P matrix) and in general this matrix, as with its counterpart the G matrix, is assumed to be constant. This assumption has been justified by two further assumptions, namely that population size is sufficiently large that genetic drift can be ignored and that selection is sufficiently weak that variation eroded by selection is replaced by mutation (Lande, 1976, 1980, 1984). Both of these assumptions are highly controversial. Effective population sizes are frequently small enough that significant drift can be expected to occur, particularly over long periods of time (Lande, 1976; Lynch, 1990; chapter 8 in Roff, 1997). Estimates of selection coefficients in wild populations show that the strength of selection varies widely from weak to very strong (Endler, 1986; Kingsolver et al., 2001), although the extent to which the estimated values represent a random sample of selection coefficients is uncertain. Nevertheless, these data indicate that the assumption that the G or P matrix will be invariant cannot be justified on theoretical grounds alone but must be verified empirically (Turelli, 1988; Arnold, 1992; Whitlock et al., 2002). The question then is not whether P matrices vary but rather at what level do detectable differences arise and to what effects can these differences be attributed?

In the present paper, we present an analysis of geographic variation in the phenotypic means, variances and covariances of populations of two grasshopper species, *Melanoplus sanguinipes* and *M. devastator*. The purpose of the analysis is to determine: (1) if the phenotypic means and **P** matrix vary among populations and between species; (2) if such variation correlates with geographic or environmental variables, suggesting that selection has shaped the **P** matrix; and (3) if variation in the **P** matrix correlates with variation in neutral genetic markers, an indication that drift has also played a role in shaping the **P** matrix.

Materials and methods

Species descriptions

Melanoplus sangiunipes is a very wide ranging grasshopper species throughout North America. In California, it is found at middle to high elevations in the Sierra Nevada, along the northern coast to Monterey Bay and from Los Angeles south (Gurney & Brooks, 1959; Dingle & Mousseau, 1994). Enclosed within the range of *M. sanguinipes*, the devastating grasshopper, *M. devastator*, inhabits the Central Valley northwards along a narrow band to Oregon, at lower altitudes in the Coast Ranges and the foothills of the Sierra Nevada, and along the coast from Monterey Bay to just north of Los Angeles (Dingle & Mousseau, 1994). The differences in distribution are reflected in life history differences that are maintained under common garden conditions (Dingle & Mousseau, 1994; Orr, 1996; Tatar et al., 1997). Morphological, life historical and physiological differences are also to be found under common garden conditions among populations of M. sanguinipes (Putnam, 1968; Chapco, 1987; Dingle et al., 1990; Gibbs et al., 1991; Dingle & Mousseau, 1994; Gibbs & Mousseau, 1994). Finally, there is also significant molecular genetic variation among populations of M. sanguinipes (Chapco & Bidochka, 1986; Orr et al., 1994).

Field collections and husbandry

Grasshoppers for this study were collected as adults from 17 sites throughout California in August and September (Fig. 1). Latitude, altitude and mean annual temperature for each site were taken from Table 1 of Dingle & Mousseau (1994). Upon return to the laboratory adult pairs were set up in individual breeding cages, consisting of a styrofoam cup filled with damp sand and a ventilated clear plastic cup placed over the sand. Fresh dandelion leaves were provided for food. All cages were maintained at 32 °C and a photoperiod of 13L : 11D. Eggs laid in the sand were collected by sieving, placed in capped vials containing moistened vermiculite and incubated at 27 °C. Nymphs that had not hatched after 32 days were deemed to be in diapause. After 45 days the diapausing eggs were transferred to 4 °C for approximately 100 days after which they were returned to 27 °C. Newly hatched nymphs were transferred to plastic boxes $(10 \text{ cm} \times 10 \text{ cm} \times 10 \text{ cm})$, which were screened on top and sides to provide ventilation. Each cage contained the offspring from a single cross. Vials containing eggs and nymphal development cages were overdispersed on trays (i.e. quasi-regular distribution to avoid clumping of populations) which were rotated and shuffled (with respect to location in incubator) on a regular schedule (every day for eggs and 2-3 days for developing nymphs).

We reared a maximum of 20 nymphs per cage. Some females produced multiple egg masses thus requiring several cages per family. Nymphs were provided with sticks for climbing and molting and were fed with a mixture of bran, rolled oats and cracked corn, and fresh dandelion leaves or wheat seedlings. Upon final eclosion adults were preserved for later measurement.

Traits measured and statistical methods

All measurements were made using a dissecting scope with ocular micrometer. Ocular micrometer units were



Fig. 1 Relief map of California showing collecting sites. For consistency, the sites are labelled according to Dingle & Mousseau (1994). Latitudes, altitudes, and mean annual temperatures are given in Table 1 of Dingle & Mousseau (1994). Collection sites for *M. sanguinipes* shown in boxes, sites for *M. devastator* shown in circles. Two populations were from site Q (Mount Palomar).

converted to mm using a stage micrometer. Three morphological measurements were taken: femur, pronotum and wing length. Femur length was measured as the maximum length of the hind femur from the trochanter to the tibia. Pronotum length was measured as the maximum length of the midline along the dorsal surface. Wing length was measured as the maximum length of the forewing, from the abdominal attachment point to the tip of the wing. In total 3441 adult offspring were measured from 511 families, with a mean of 6.7 individuals per family (range 1–59, SD = 8.3) and 202.4 individuals per population (range 61–439, SD = 112.5).

Although the data are structured by family, the majority of siblings were reared in a common cage (at least 89% were so reared but, unfortunately, cages within families were not recorded) and hence it is not strictly possible to separate cage effects from family effects. For this reason, we did not attempt an analysis based on a presumed genetic structure. To take into account any family/cage effects we calculated the phenotypic covariances (as the variance is itself the covariance between a trait and itself we use the term 'covariance' to

apply to both variances and covariances) using a multivariate analysis of variance with family/cage as an independent variable (Bulmer, 1985). The phenotypic covariances were then estimated as the sum of the covariance components. The individuals used in the present analysis were the offspring of adults taken from the field and hence we cannot preclude maternal effects as a potential source of differences among populations. Maternal effects on adults size appear to be very uncommon (Mousseau & Dingle, 1991a,b; Shaw & Byers, 1998). Further, Putnam (1968) reared three populations of *M. sanguinipes* for four generations in the laboratory and found no difference between the F1 generations and the other generations. Therefore, while we cannot definitively rule out such effects in the present analysis, the analysis by Putnam (1968) make such effects highly unlikely.

Analysis of phenotypic means

A preliminary analysis showed significant effects due to family/cage: therefore, we used cage/family means in the present analysis. We analysed the data using MANOVA,

with the variable 'population' nested within 'species'. In a nested analysis of variance the effect of X when Y is nested within X and is significant is determined by the mean square of X over the mean square of the nested term and not the error term (Zar, 1999). The F-ratio in MANOVA is constructed somewhat differently from that in ANOVA and precludes readily making the appropriate test when the nested term is significant. Therefore, to test the effect of species and its interaction with sex we used the residuals from the MANOVA of sex, population (species) and their interaction, which we refer to as a 'step-wise' MANOVA. To ask how well sex and the three morphological characters discriminate between the two species we ran a discriminant function analysis.

We next analysed the data with respect to species, sex and the three geographic variables latitude, altitude and temperature. To prevent pseudo-replication, we used the mean of the cage means for each population. In a saturated model there are a total of 21 terms (five additive, nine two-way interactions, six three-way interactions, one four-way interaction). Given that there are only 34 data points (17 populations and two sexes) such a model is inappropriate. The question, we are interested in asking is 'Are there main effects of variables and any indication of interactions?' We, therefore, used a model that included all additive terms and all two-way interactions involving species or sex (six terms).

An alternate approach to MANOVA is to reduce the number of independent variables by use of principal components analysis. We subjected the factor scores using the first principal component to an analysis of variance with the independent variables sex, species and population nested in species. Next we used the grand cage means to test for an association between the three geographic variables, keeping both sex and species in the model. As before, we used only the two-way interaction terms that included sex or species.

Matrix comparison: Flury method

There is no single satisfactory method of analysing matrix variation. Here we use two approaches that address different aspects of such variation: the Flury hierarchical method and the Jackknife followed by MANOVA method. The Flury approach considers a hierarchy of possible differences among a set of matrices. The method looks at the matrices from the perspective of their common principal components: starting from the 'top' of the hierarchy the matrices can be (a) identical; (b) proportional, in which case the matrices share identical principal components (identical eigenvectors) but their eigenvalues differ by a proportional constant; (c) share principal components but differ without pattern in their eigenvalues (CPC model); (d) they can have one or more, but not all, principal components in common [CPC(I) model, where I is the number of principal components in common]; or (e) they can have completely unrelated structures (Phillips & Arnold, 1999). Two approaches to the testing of the hierarchical structure are the step-up approach and the jump-up approach (Phillips & Arnold, 1999): we employed both methods in the present analysis. The argument for analysing the principle components is that evolutionary change will be influenced by the size of the eigenvalues corresponding to the principal components (Bjorklund, 1996; Schluter, 1996), that is, evolution will tend to proceed in the direction of greatest genetic variance. On the other hand if the principal components underlying differences between matrices cannot biologically be decomposed into causal factors the interpretation of the Flury analysis can be misleading (Houle *et al.*, 2002).

Matrix comparison: Jackknife-MANOVA method

Another limitation of the Flury method is that it does not lend itself readily to a consideration of how variation among matrices might be structured with respect to such variables as sex and population. To do this analysis, we used the Jackknife-MANOVA method (Roff, 2002a), the statistical validity of which has been verified by simulation (Bégin et al., 2004). The procedure is as follows: for each group (population, sex within population, etc.) calculate the **P** matrix. Next, delete in turn one sampling unit, which in this case is a 'family' unit, and calculate the pseudovalues according to the usual Jackknife procedure (Potvin & Roff, 1993; Manly, 1997). The final data matrix was arranged such that the columns comprised the pseudovalues of each covariance and the rows the results for the deletion of a given 'family' (so the *i*th row *j*th column is the pseudovalue for the *j*th covariance for the sample with the *i*th 'family' deleted). These data were then used in a multivariate analysis of variance with independent variables species, sex, and population. As with the phenotypic means, we adopted a stepwise approach to account for the nesting of population within species.

Correlation of covariances with environmental variables As with the phenotypic means, to analyse the relationship between the **P** matrices and the geographic variables we used a single estimate per population. As there were no significant effects due to sex we averaged the two matrices. Because the phenotypic means differed we could not simply lump males and females to calculate the covariances. Averaging the two matrices is equivalent to using the sex-specific pseudovalues as independent observations.

Testing for drift

Divergence in the **P** matrix could come about because of selection or drift. In principle, drift will change the matrices in such a manner that they will remain proportional, but experimental evidence indicates that can change other components of matrix structure (Whitlock *et al.*, 2002). Further, proportionality can itself result from selection if the traits are highly correlated

(Roff, 2004). Hence, the type of variation among matrices cannot be used to determine if such changes were the result of selection or drift. However, in the case of drift there should be a positive correlation between pair-wise differences in the P matrices and the pair-wise differences in neutral genetic markers. Similarly, if drift is responsible for differences between trait means then these also should be positively correlated with differences in neutral genetic markers. A phylogenetic analysis of the Californian populations of M. sanguinipes and M. devastator, was produced by Orr et al. (1994). This study and the present study have 11 populations in common: we, therefore, used Nei's genetic distances from the Orr study to address the question of whether genetic divergence in these 11 populations is mirrored in the divergence of the trait means and P matrices. For all possible pair-wise contrasts we calculated the mean percentage difference between elements of the two **P** matrices:

$$T\% = 100 \frac{\sum_{i=1}^{n} |\rho_{i1} - \rho_{i2}|/c}{(\bar{\rho}_1 + \bar{\rho}_2)/2},$$

where ρ_{ij} is the *i*th element of the *j*th matrix, *c* is the number of elements in each matrix and $\bar{\rho}_j$ is the overall average of the elements of the *j*th matrix. We then correlated this set of differences with the set of genetic distances between the respective two matrices. Because of nonindependence, we tested for significance using the Mantel test (Manly, 1997).

Results

Variation in phenotypic means

Melanoplus devastator is larger than *M. sanguinipes* and females of *M. sanguinipes* are larger than the males (Table 1). The initial MANOVA analysis shows that population, is highly significant (Table 2). Because this nested variable is significant, we use the 'stepwise' MANOVA to test for other effects. This analysis shows that all effects are significant except the interaction between sex and species (Table 2). Thus, to account for variation in body components, we must take into account sex, species and the population from which the individuals came. The discriminant function, using sex and morphology, correctly identified 80% of *M. devastator* and 83% of *M. sanguinipes*. We next analysed the data with respect to species, sex and the three geographic variables latitude, altitude and temperature. Altitude, temperature and latitude are themselves correlated (Alt. vs. Temp., r = -0.774; Alt. vs. Lat., r = 0.008; Alt. vs. Temp., r = -0.470) but not sufficiently so as to cause problems with colinearity (Tabachnick & Fidell, 2001, p. 83). All main effects are significant (P < 0.01 in all cases) but none of the interactions are significant (P > 0.35 in all cases). Overall, the morphological traits decrease as altitude increases, increase with temperature and decrease with latitude (Fig. 2).

The first principal component accounts for 84.7% of the total variation and hence captures a very large fraction of the variability in morphology. We tested for variation in the first principal component using ANOVA. As with the previous stepwise MANOVA, sex, population, species and species × sex are all significant. Based on the grand cage means, there is a marginally nonsignificant effect of sex ($F_{1,23} = 3.72$, P = 0.0662) but both species and temperature are highly significant ($F_{1,23} = 15.47$, P < 0.001, $F_{1,23} = 19.45$, P < 0.001, respectively). None of the other terms approach statistical significance (P > 0.2).

In summary, the MANOVA shows that body components differ significantly between the two species, between the sexes, among the populations and as a function of altitude, latitude and temperature. Reducing the number of dependent traits by use of PCA produces more or less the same results except that the significant association with latitude and altitude is lost.

Variation among the P matrices

Flury method

Because of the possibility of differences due to species or sex we used only data from females and did separate analyses for the two species using the Flury method. Both step-up and jump-up analyses gave the same results. For *M. sanguinipes* the Flury analysis indicates that the **P** matrices differ with respect to equality, proportionality, and share no common principal components (P < 0.01 in all cases). For *M. devastator* the matrices are not proportional (P < 0.05 using either approach) but do share common principal components (P = 0.0545 using the step-up approach, P = 0.0826 using the jump-up approach). Thus in both species there

Table 1 Mean* trait values (SE) and abioticdata (mean, SE) for Melanoplus sanguinipesand M. devastator

	Femur	Wing	Prothorax	Latitude (m)	Altitude (m)	Temperature (°C)†
M. sanguinipes, 3	71.3 (1.4)	95.3 (3.7)	73.3 (1.2)	37.0 (0.9)	1496 (251.5)	10.5 (1.2)
<i>M.</i> sanguinipes, ♀	75.8 (1.1)	97.1 (3.4)	76.4 (0.9)			
M. devastator, 3	79.5 (1.6)	108.8 (5.0)	83.4 (1.8)	36.4 (1.2)	15.0 (2.9)	15.3 (1.0)
<i>M. devastator</i> , Q	81.3 (2.1)	105.9 (5.3)	82.0 (2.02)			

*Means (mm × 10) based on population means calculated using cage means. †Mean annual temperature.

	Wilks λ	Approx. F	Numerator d.f.	Denominator d.f.	P-value
Species*	0.799	72.0	3	858	< 0.001
Sex	0.770	82.6	3	829	< 0.001
Population (species)	0.053	92.7	45	2463	< 0.001
Species \times sex*	0.995	1.57	3	858	0.196
Population	0.872	2.59	45	2463	< 0.001
$(species) \times sex$					

 Table 2 Results of MANOVA on cage means of the thee morphological traits (femur, pronotum and wing lengths)

*Tested using residuals from MANOVA with sex, population (species) and their interaction.

is statistically significant variation in the structure of the **P** matrices.

Jackknife–manova method

To determine if this variation in structure of the **P** matrices is associated with sex, species or population of origin we used the Jackknife–MANOVA method. For both the untransformed and log-transformed data there

is a highly significant effect of population (Table 3). There is a marginally nonsignificant effect due to sex in the untransformed data that is not even close to significance in the log-transformed data (Table 3). These results indicate that differences due to sex are a result of allometric scale effects but the variation among populations is not. We confirmed that a logarithmic transformation removed any allometric association between mean and variance by calculating the correlation between population means and variances for the raw and transformed data: with the raw data there is one significant correlation and two that approach significance, but none with the log-transformed data (Table 4).

Correlation between covariances and environmental variables

As the effect of species was not significant we excluded this category from further analysis. For the purposes of comparison with the phenotypic means analysis we first ran a MANOVA using only the additive model: all three variables are significant (P < 0.01 in all cases). The general pattern of association is the same as with the



Fig. 2 Population mean morphological traits as a function of altitude, latitude and temperature. Circles represent *M. sanguinipe*, triangles are *M. devastator*. Females shown in filled symbols, males in open symbols.

Table 3	Summary	of	MANOVA	on	Р	matrices
---------	---------	----	--------	----	---	----------

	Approx. F		Numerator	Denominator	P-value		
	Raw	Log	d.f.	d.f.	Raw	Log	
Species	1.65	1.01	6	855	0.131	0.418	
Sex	2.08	1.25	6	826	0.053	0.279	
Population (species)	2.53	1.91	90	4651	<0.001	<0.001	
Species \times sex	0.36	0.26	6	855	0.906	0.956	
Sex × population (species)	0.61	0.62	90	4651	0.999	0.998	

Table 4 Correlations between means and variances of raw and logtransformed data in *Melanoplus*

	Untransfor	med	Log-transformed		
Trait	r	P-value	r	P-value	
Femur	0.563	0.001	0.206	0.241	
Wing	0.297	0.088	0.196	0.266	
Pronotum	0.323	0.063	0.008	0.966	

phenotypic means (cf. Figs 2 and 3). The saturated model indicates that there are significant interactions (Table 5).

Testing for drift

There is no significant correlation between the difference in trait means and the difference in Nei's genetic distance (P > 0.3 in all cases, including using the first pc score), suggesting that variation in trait means cannot be accounted for by drift. On the other hand, there is a highly significant correlation between Nei's genetic distance and, T%, the average difference between the corresponding **P** matrices (P < 0.01, Mantel's test using 5000 randomizations: Fig. 4. We also note that the Pearson correlation is 0.36 with P < 0.01). There is no consistent difference between combinations of different species (Fig. 4), consistent with the apparent lack of genetic differentiation between the two species (Orr et al., 1994). We also ran correlations between the genetic distance and geographic distance and between T% and geographic distance. In neither case is the correlation significant (P > 0.3 in both cases).

Discussion

There are statistically significant quantitative differences in morphology between the sexes and the two species (Table 1), as has been found previously for these two species (Orr *et al.*, 1994) and between other pairs of *Melanoplus* species (Guenther & Chapco, 1990). Neverthe-less, there is considerable overlap in morphology between *M. sanguinipes* and *M. devastator* (Table 1). Morphology covaries with latitude, altitude and temperature but species still remains a significant predictor,



Fig. 3 Plots of the **P** matrix elements for each population against the three environmental variables. For simplicity the axes have been left numerically unlabeled. Symbols indicate *M. sanguinipes* (S) and *M. devastator* (D). The ellipses are drawn to show the general orientation of the covariation. Each ellipse is centred on the sample means of the *x* and *y* variables. The unbiased sample standard deviations of *x* and *y* determine its major axes and the sample covariance between *x* and *y*, its orientation. The size of the ellipse was set to a probability value of 0.6827. The apparent outlier does not represent a single population.

indicating that the differences between the species cannot be accounted for solely by differences in the variables measured in this study. It is possible that the effect of species might disappear with the addition of other environmental variables. The two species hybridize but hatching rates between populations widely separated are very low (Orr *et al.*, 1994; Orr, 1996). Molecular

	Wilks' λ	F	P-value
Temperature	0.111	5.316	0.064
Altitude	0.052	12.254	0.015
Latitude	0.139	4.119	0.096
Latitude × altitude	0.048	13.251	0.013
Latitude × temperature	0.091	6.675	0.044
Temperature \times altitude	0.053	11.971	0.016
Latitude \times temperature \times altitude	0.053	11.956	0.016

 Table 5 MANOVA results for association between the grand P

 matrices and three geographic variables



Fig. 4 *T*% vs. Nei's genetic distance for 11 populations of *Melanoplus*. Circles indicate combinations in which one is *M. devastator* [populations Davis (T) and Santa Barbara (U)]. Triangles show pairs in which both species were *M. sanguinipes*.

analysis cannot clearly distinguish between the two species and gene flow analysis indicates that the two species are not reproductively isolated (Orr et al., 1994). Thus we have a situation in which there is considerable morphological variation among populations and species, but at the same time there is gene flow both among populations and species. This flow is not sufficient to eliminate both morphological and life history differences among populations (Dingle et al., 1990; Gibbs et al., 1991; Dingle & Mousseau, 1994; Orr, 1996; Tatar et al., 1997). The statistical association between trait variation in these studies and the present analysis and variation in such factors as latitude, altitude and temperature argues for the direct effects of selection in the production of these population differences. At the same time there is considerable geographic distance among some of the populations (Fig. 1) allowing for the possibility of differences arising via genetic drift. We found no correlation between geographic distance and genetic divergence but this most likely reflects the fact that gene flow is generally around the central valley of California and thus the shortest distance between locations is a poor indicator of the true distance between populations (Fig. 1).

Of the three environmental variables used in the present study, temperature is probably the most closely connected to the actual agent of selection. Temperature can act directly on trait values and indirectly by its involvement in the determination of season length. With regard to the former, water loss has been shown to vary among populations with grasshoppers from high-elevation populations exhibiting the greatest water loss and metabolic rates (Rourke, 2000). This water loss is a function of the biophysical properties of the cuticle, properties that vary among populations raised under common garden conditions (Gibbs et al., 1991) and among families (Gibbs & Mousseau, 1994). Gibbs et al. (1991) suggested that this variation is a direct consequence temperature selection, and Rourke (2000) showed that grasshopper body temperatures can exceed 40 °C, putting them in a thermal regime where cuticular construction could be very important.

Mean annual temperature is an index of the duration of time over which conditions are suitable for growth and reproduction. Thus low elevation populations face a very hot, dry summer and M. devastator females pass this period in reproductive diapause (Middelkauf, 1964). In contrast, at high elevations the summers are much cooler and the *M. sanguinipes* populations display rapid hatching and growth and no reproductive diapause (Orr et al., 1994). A shortened period for growth means less time to increase in body size and hence, all other things being equal, body size should decrease with increases in altitude and latitude and increase with increases in temperature (Masaki, 1967; Roff, 1980, 1983). This pattern has been observed in a wide range of insect species (reviewed in Roff, 2002b) and is found in the present species (Orr, 1996; Tatar et al., 1997; this study; Fig. 2).

The present analysis indicates that there are statistically highly significant differences among the **P** matrices. This strongly suggests that there will also be variation in the **G** matrix but the problem of disentangling cage effects from family effects precludes a direct test of this. However, there is a high correlation (r = 0.79, n = 102, P < 0.001) between the 'among-family' covariances (which in the absence of common environmental and nonadditive effects would be equal to the additive genetic covariances) and the phenotypic covariances, which is consistent with other studies showing a high correlation between the **P** and **G** matrices for morphological traits (Roff, 1996). It seems highly unlikely that the variation among the populations grown in a common garden setting will be a consequence of variation in the environmental covariances alone. There is now abundant evidence that the covariances, both genetic and phenotypic evolve (Roff, 2000; Steppan et al., 2002). What we seriously lack is empirical data on the nature of the variation in the covariances: it is this question that is addressed in the present paper. If the **P** and **G** matrices are proportional then the former will provide a more precise estimate of the form of **G** (Steppan *et al.*, 2002). In this respect, it is pertinent to note that the regression of the 'among-family' covariances (*Y*) on the phenotypic covariances (*X*) is Y = -0.0005 + 1.29X, which supports the hypothesis of proportionality.

The Flury method indicates that the **P** matrices for the 14 populations of M. sanguinipes share no principal components in common, whereas the three populations of M. devastator do share principal components. The environmental range of the M. devastator populations is much less than M. sanguinipes, the mean annual temperatures for the three M. devastator populations ranging only from 13.3 °C (Big Sur, site N) to 16.4 °C (Santa Barbara, site U), whereas the range for M. sanguinipes is from 1 °C (Tioga Pass, site A) to 18 °C (Whispering Palms, site W). Similarly, the coefficient of variation for the P matrix elements in M. devastator is 10.6% whereas for M. sanguinipes it is 26.6%. Thus, the relative lack of differences among the **P** matrices of *M. devastator* could be related to the lack of variation in environmental conditions. This hypothesis was explored using the MANOVA method of matrix analysis.

In the first part of the analysis, we examined variation due to sex, species and population of origin. We found no significant effect due to species, differences due to sex that could be attributed to scaling effects, and significant variation among populations. This results indicates that variation in the **P** matrix within a species can be greater than that among species and suggests that the two species are very closely related, a result that is consistent with the genetic evidence (Orr et al., 1994). The finding of significant variation among populations is in accord with the results from the Flury analysis. In the second part of the analysis, we examined the relationship between the P matrix and the three environmental variables, and found the same general pattern as with trait means, namely covariances increase with increased temperature but decrease with increased latitude or altitude (Fig. 3). This covariation is not a result of scaling (Table 4) and suggests that selection has played a role in shaping the phenotypic variances and covariances in much the same manner as it has shaped the phenotypic means.

The clinal pattern of variation in the **P** matrix of the two *Melanoplus* species suggests that equivalent patterns are likely to exist in the genetic components. This prediction is supported by data from three different insect species. Latitudinal variation in the heritability of critical photoperiod has been observed in the pitcherplant mosquito (Bradshaw & Holzapfel, 2001) and in morphological traits of the *Allonemobius fasciatus/socius* complex (Roff & Mousseau, 1999). Latitudinal variation in wing length and wing area was found in South American populations of *Drosophila melanogaster* (van 'T Land *et al.*, 1999): heritabilities of both traits show positive clinal variation, with that of wing area being

highly significant (t = 3.99, n = 10, P = 0.004, analysis of data presented in paper). Habitat components have also been correlated with variation in genetic/phenotypic variances and covariances in the amphipod, *Gammarus minus* (Fong, 1989; Jernigan *et al.*, 1994; Roff, 2002a,b), New World monkeys (Marroig & Cheverud, 2001) and the mangrove *Avicennia germinans* (Dodd *et al.*, 2000).

If drift has played a role in the current phenotypic means, variances and covariances we should expect to find differences between populations being positively correlated with differences in neutral genetic markers. We found no differences with respect to the phenotypic means but did find highly significant correlation between Nei's genetic distance and T% (measure of difference between two **P** matrices: Fig. 4). Further, as with the molecular genetic analysis, there is no obvious separation in the two species. The covariation of the phenotypic covariances with temperature and with neutral genetic markers argues for the combined effects of both drift and selection in moulding the **P** matrix, whereas selection appears to the primary determinant of variation in the trait means.

Bégin & Roff (2004) examined matrix variation in morphology in seven species of crickets, belonging to two different genera (six *Gryllus* spp. and one *Teleogryllus* spp.). The average percentage pair-wise difference between the elements of the **P** matrices (*T*%) was 38% (SE = 5.3, range 9–92%, median = 31%) compared with 33% (SE = 1.65, range 5–92%, median = 30%) for the *Melanoplus* populations and species. The variation among the cricket species did not reflect phylogenetic divergence, which agrees with the present analysis in that there was no significant difference overall between *M. devastator* and *M. sanguinipes* in their **P** matrices (Table 3), although there was highly significant differences in mean morphology (Table 2), as was also found in the cricket study (Bégin & Roff, 2004).

The present analysis indicates, as do numerous other analyses of both P and G matrix variation (reviewed in Roff, 2000; Steppan et al., 2002; Bégin & Roff, 2004) that, contrary to the 'classical' evolutionary quantitative genetic models, matrix evolution is the norm not the exception and that significant variation is likely to be general even at the level of intra-specific population comparisons. The present analysis indicates that this variation is driven both by drift and selection. The challenge is to develop a theory of how selection concurrently changes both trait means and covariances between traits. At the same time, we need to collect more empirical data on matrix variation in relation to ecologically relevant variables. Because selection acts on the phenotype, while response to selection depends on the genotype, we require information on variation in both the **P** and **G** matrices. Further, while most studies to date have focused on morphological traits we very clearly need more studies directed to life history variation, where selection can be expected to be particularly strong.

Acknowledgments

We are very grateful for the constructive comments of Mattieu Bégin. This work was supported by a grant from the University of California.

References

- Ackerman, R.R. & Cheverud, J.M. 2000. Phenotypic covariance structure in tamarins (Genus *Saguinus*): a comparison of variation patterns using matrix correlation and common principal components. *Am. J. Phys. Anthrop.* 111: 489–501.
- Arnett, A.E. & Gotelli, N.J. 1999. Geographic variation in lifehistory traits of the ant lion, *Myrmeleon immaculatus*: evolutionary implications of Bergmann's rule. *Evolution* **53**: 1180– 1188.
- Arnold, S.J. 1992. Constraints on phenotypic evolution. *Am. Nat.* **140**(Suppl.):s85–s107.
- Bégin, M., Roff, D.A. & Debat, V. 2004. The effect of temperature and wing morphology on quantitative genetic variation in the cricket *Gryllus firmus*, with an appendix examining the statisticl properties of the Jackknife-MANOVA method of matrix comparison. J. Evol. Biol. 17: 1255–1267.
- Bégin, M. & Roff, D.A. 2004. From micro- to macroevolution through quantitative genetic variation: positive evidence from field crickets. *Evolution*. **58**: 2287–2304.
- Bjorklund, M. 1996. The importance of evolutionary constraints in ecological time scales. *Evolution. Ecol.* **10**: 423–431.
- Blanckenhorn, W.U. & Fairbairn, D.J. 1995. Life history adaptation along a latitudinal cline in the water strider Aquarius remigis (Heteroptera: Gerridae). J. Evol. Biol. 8: 21–41.
- Bradshaw, W.E. & Holzapfel, C.M. 2001. Phenotypic evolution and the genetic architecture underlying photoperiodic time measurement. J. Insect Physiol. 47: 809–820.
- Bulmer, M.G. 1985. *The Mathematical Theory of Quantitative Genetics*. Claredon Press, Oxford.
- Chapco, W. 1987. Genetic variation in three subspecies of *Melanoplus sanguinipes* F. I. Morphometric traits. *Genome* 29: 793–801.
- Chapco, W. & Bidochka, M.J. 1986. Genetic variation in prairie populations of *Melanoplus sanguinipes* the migratory grasshopper. *Heredity* 56: 397–408.
- Conover, D.O., Brown, J.J. & Ehtisham, A. 1997. Countergradient variation in growth of young striped bass (*Morone saxatilis*) from different latitudes. *Can. J. Fish. Aquatic Sci.* **54**: 2401–2409.
- Dingle, H. & Mousseau, T.A. 1994. Geographic variation in embryonic development time and stage of diapause in a grasshopper. *Oecologia* 97: 179–185.
- Dingle, H., Mousseau, T.A. & Scott, S.M. 1990. Altitudinal variation in life cycle syndromes of California USA populations of the grasshopper *Melanoplus sanguinipes* F. *Oecologia* 84: 199–206.
- Dodd, R.S., Afzal Rafii, Z. & Bousquet-Melou, A. 2000. Evolutionary divergence in the pan-Atlantic mangrove Avicennia germinans. New Phytol. 145: 115–125.
- Endler, J.A. 1986. *Natural Selection in the Wild*. Princeton University Press, Princeton, NJ.
- Fong, D.W. 1989. Morphological evolution of the amphipod *Gammarus minus* in caves: quantitative genetic analysis. *Am. Midl. Nat.* **121**: 361–378.

- Gibbs, A. & Mousseau, T.A. 1994. Thermal acclimation and genetic variation in cuticular lipids of the lesser migratory grasshopper (*Melanoplus sanguinipes*): Effects of lipid composition on biophysical properties. *Physiol. Zool.* **67**: 1523–1543.
- Gibbs, A., Mousseau, T.A. & Crowe, J.H. 1991. Genetic and acclimatory variation in biophysical properties of insect cuticle lipids. *Proc. Natl Acad. Sci. U S A* **88**: 7257–7260.
- Guenther, S.J. & Chapco, W. 1990. A morphometric analysis of *Melanoplus* females (Orthoptera: Acrididae). *Pan-Pacific Ent.* 66: 39–42.
- Gurney, A.B. & Brooks, A.R. 1959. Grasshoppers of the Mexicanus group, genus *Melanoplus* (Orthoptera: Acrididae). *Proc. U. S. Nat. Museum* 110: 1–93.
- Houle, D., Mezey, J. & Galperna, P. 2002. Interpretation of the results of common principal components. *Evolution* **56**: 433–440.
- Jernigan, R.W., Culver, D.C. & Fong, D.W. 1994. The dual role of selection and evolutionary history as reflected in genetic correlations. *Evolution* **48**: 587–596.
- Kingsolver, J.G., Hoekstra, H.E., Hoekstra, J.M., Berrigan, D., Vignieri, S.N., Hill, C.E., Hoang, A., Gilbert, P. & Beerli, P. 2001. The strength of phenotypic selection in natural populations. *Am. Nat.* **157**: 245–261.
- van 't Land, J., van Putten, P., Zwaan, B.J., Kamping, A. & van Delden, W. 1999. Latitudinal variation in wild populations of *Drosophila melanogaster*: heritabilities and reaction norms. *J. Evol. Biol.* **12**: 222–232.
- Lande, R. 1976. Natural selection and random genetic drift in phenotypic evolution. *Evolution* **30**: 314–334.
- Lande, R. 1979. Quantitative genetic analysis of multivariate evolution applied to brain:body size allometry. *Evolution* **33**: 402–416.
- Lande, R. 1980. The genetic covariance between characters maintained by pleiotropic mutations. *Genetics* **94**: 203–215.
- Lande, R. 1984. The genetic correlation between characters maintained by selection, linkage and inbreeding. *Genet. Res.* **44**: 309–320.
- Lynch, M. 1990. The rate of morphological evolution in mammals from the standpoint of the neutral expectation. *Am. Nat.* **136**: 727–741.
- Manly, B.F.J. 1997. Randomization, Bootstrap and Monte Carlo Methods in Biology. Chapman and Hall, New York.
- Marroig, G. & Cheverud, J.M. 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.
- Masaki, S. 1967. Geographic variation and climatic adaptation in a field cricket (Orthoptera: Gryllidae). *Evolution* **21**: 725–741.
- McKay, J.K., Bishop, J.G., Lin, J.-Z., Richards, J.H., Sala, A. & Mitchell-Olds, T. 2001. Local adaptation across a climatic gradient despite small effective population size in the rare sapphire rockcress. *Proc. R. Soc. Biol. Sci. Ser. B* 268: 1715–1721.
- Merila, J., Laurila, A., Laugen, A.T., Rasanen, K. & Pahkala, M. 2000. Plasticity in age and size at metamorphosis in *Rana temporaria*: Comparison of high and low latitude populations. *Ecography* 23: 457–465.
- Middelkauf, W.W. 1964. Effects of photoperiod upon oogenesis in *Melanoplus devastator*. J. Kansas Ent. Soc. **37**: 163–168.
- Mousseau, T.A. & Dingle, H. 1991a. Maternal effects in insects: examples, constraints, and geographic variation. In: *The Unity of Evolutionary Biology* (E. C. Dudley, ed.), pp. 745–761. Discoides Press, Portland, OR.

- Mousseau, T.A. & Dingle, H. 1991b. Maternal effects in insect life histories. Ann. Rev. Ent. 36: 511–534.
- Mousseau, T.A. & Roff, D.A. 1989. Adaptation to seasonality in a cricket: patterns of phenotypic and genotypic variation in body size and diapause expression along a cline in season length. *Evolution* **43**: 1483–1496.
- Nylin, S., Gotthard, K. & Wiklund, C. 1996. Reaction norms for age and size at maturity in *Lasiommata* butterflies: Predictions and tests. *Evolution* **50**: 1351–1358.
- Orr, M.R. 1996. Life-history adaptation and reproductive isolation in a grasshopper hybrid zone. *Evolution* **50**: 704–716.
- Orr, M.R., Porter, A.H., Mousseau, T.A. & Dingle, H. 1994. Molecular and morphological evidence for hybridization between two ecologically distinct grasshoppers (*Melanoplus sanguinipes* and *M. devastator*) in California. *Heredity* **72**: 42–54.
- Phillips, P.C. & Arnold, S.J. 1999. Hierarchical comparison of genetic variance-covariance matrices. I. Using the Flury hierachy. *Evolution* **53**: 1506–1515.
- Potvin, C. & Roff, D.A. 1993. Distribution-free and robust statistical methods: viable alternatives to parametric statistics? *Ecology* **74**: 1617–1628.
- Putnam, L.G. 1968. Morphometric polymorphism in *Melanoplus* sanguinipes in Western Canada. Ann. Ent. Soc. Am. 61: 540–541.
- Roff, D.A. 1980. Optimizing development time in a seasonal environment: the "ups and downs" of clinal variation. *Oecologia* **45**: 202–208.
- Roff, D.A. 1983. Phenological adaptation in a seasonal environment: a theoretical perspective. In: *Diapause and Life Cycle Strategies in Insects* (V. K. Brown & I. Hodek, eds), pp. 253–270. Dr W. Junk, The Hague.
- Roff, D.A. 1996. The evolution of genetic correlations: an analysis of patterns. *Evolution* **50**: 1392–1403.
- Roff, D.A. 1997. *Evolutionary Quantitative Genetics*. Chapman and Hall, New York.
- Roff, D.A. 2000. The evolution of the G matrix: selection or drift? *Heredity* **84**: 135–142.
- Roff, D.A. 2002a. Comparing G matrices: a MANOVA method. *Evolution* **56**: 1286–1291.
- Roff, D.A. 2002b. *Life History Evolution*. Sinauer Associates, Sunderland, MA.
- Roff, D.A. 2004. The evolution of genetic architecture. In: *Phenotypic Integration* (M. Pigliucci & K. Preston, eds), pp. 345–365. Oxford University Press, Oxford.

- Roff, D.A. & Mousseau, T.A. 1999. Does natural selection alter genetic architecture? An evaluation of quantitative genetic variation among populations of *Allonemobius socius* and *A. fasciatus. J. Evol. Biol.* **12**: 361–369.
- Rourke, B.C. 2000. Geographic and altitudinal variation in water balance and metabolic rate in a California grasshopper, *Melanoplus sanguinipes. J. Exp. Biol.* **203**: 2699–2712.
- Schluter, D. 1996. Adaptive radiation along genetic lines of least resistance. *Evolution*. **50**: 1766–1774.
- Shaw, R.G. & Byers, D.L. 1998. Genetics of maternal and paternal effects. In: *Maternal Effects as Adaptations* (T. A. Mousseau & C. W. Fox, eds), pp. 97–111. Oxford University Press, New York.
- Smith, G.R., Ballinger, R.E. & Nietfeldt, J.W. 1994. Elevational variation of growth rates in neonate *Sceloporus jarrovi*: an experimental evaluation. *Funct. Ecol.* **8**: 215–218.
- Steppan, S.J. 1997. Phylogenetic analysis of phenotypic covariance structure. 2. Reconstructing matrix evolution. *Evolution* 51: 587–594.
- Steppan, S.J., Phillips, P.C. & Houle, D. 2002. Comparative quantitative genetics: evolution of the G matrix. *TREE* 17: 320–327.
- Tabachnick, B.G. & Fidell, L.S. 2001. *Using Multivariate Statistics*. Allyn and Bacon, Boston.
- Tatar, M., Gray, D.W. & Carey, J.R. 1997. Altitudinal variation for senescence in *Melanoplus* grasshoppers. *Oecologia* 111: 357–364.
- Thomas, G., Joseph, L., Varghese, G., Sreejayan, Kalyanaraman, K. & Kuriachan, P. 2001. Analysis of phenotypic and genetic variations among populations of *Oryza malampuzhaensis* show evidence of altitude-dependent genetic changes. *Can. J. Bot.* **79**: 1090–1098.
- Tracy, C.R. 1999. Differences in body size among chuckwalla (*Sauromalus obesus*) populations. *Ecology* **80**: 259–271.
- Turelli, M. 1988. Phenotypic evolution, constant covariance, and the maintenance of additive variance. *Evolution* **42**: 1342–1347.
- Whitlock, M.C., Phillips, P.C. & Fowler, K. 2002. Persistance of changes in the genetic covariance matrix after a bottleneck. *Evolution* 56: 1968–1975.
- Zar, J.H. 1999. *Biostatistical Analysis*. Prentice-Hall, Englewood Cliffs, NJ.

Received 23 July 2004; revised 11 October 2004; accepted 25 October 2004