Evolution, 53(5), 1999, pp. 1516–1527

HIERARCHICAL COMPARISON OF GENETIC VARIANCE-COVARIANCE MATRICES. II. COASTAL-INLAND DIVERGENCE IN THE GARTER SNAKE, *THAMNOPHIS ELEGANS*

STEVAN J. ARNOLD[1,2] AND PATRICK C. PHILLIPS[3]
[1]*Department of Ecology and Evolution, University of Chicago, Chicago, Illinois 60637*
[3]*Department of Biology, Box 19498, University of Texas at Arlington, Arlington, Texas 76019-0498*
E-mail: pphillips@uta.edu

Abstract.—The time-scale for the evolution of additive genetic variance-covariance matrices (*G*-matrices) is a crucial issue in evolutionary biology. If the evolution of *G*-matrices is slow enough, we can use standard multivariate equations to model drift and selection response on evolutionary time scales. We compared the *G*-matrices for meristic traits in two populations of garter snakes (*Thamnophis elegans*) with an apparent separation time of 2 million years. Despite considerable divergence in the meristic traits, foraging habits, and diet, these populations show conservation of structure in their *G*-matrices. Using Flury's hierarchial approach to matrix comparisons, we found that the populations have retained the principal components (eigenvectors) of their *G*-matrices, but their eigenvalues have diverged. In contrast, we were unable to reject the hypothesis of equal environmental matrices (*E*-matrices) for these populations. We propose that a conserved pattern of multivariate stabilizing selection may have contributed to conservation of *G*- and *E*-matrix structure during the divergence of these populations.

Key words.—Flury hierarchy, genetic correlation, matrix comparisons, quantitative genetics, snakes.

Received September 17, 1998. Accepted May 6, 1999.

The additive genetic variance-covariance matrix (*G*-matrix) plays a central role in evolutionary biology. This matrix enables us to predict how a whole set of continuously distributed traits will evolve by drift in finite populations or in response to selection in large populations (Lande 1979). The *G*-matrix also summarizes the pattern and strength of genetic constraints in a population (Lande 1979; Cheverud 1984; Maynard Smith et al. 1985; Charlesworth 1990; Arnold 1992a).

The constancy of the *G*-matrix has emerged as a critical issue in evolutionary biology. Constancy of the matrix in phylogenetic time would permit us to reconstruct historical patterns of selection (Lande 1979). This ability to reconstruct selection has attracted attention because it constitutes one the few bridges between the field of quantitative genetics and the discipline of systematics. Some have argued that we should abandon this bridge between time scales. The argument is that fluctuations in the *G*-matrix from generation to generation might invalidate the usual procedures for reconstructing selection (Turelli 1988). Nevertheless, it is not clear how much fluctuation in *G* would be required for us to throw up our hands. Fluctuations in *G* might be so minor that reconstruction of selection is unaffected or so profound that reconstruction is pointless. The comparison of *G*-matrices is crucial in deciding where we lie along this continuum. Despite their importance, however, only a few statistical comparisons of *G*-matrices have been accomplished (Lofsvold 1986; Billington et al. 1988; Kohn and Atchley 1988; Fong 1989; Wilkinson et al. 1990; Shaw and Billington 1991; Spitze et al. 1991; Platenkamp and Shaw 1992; Brodie 1993; Carr and Fenster 1994; Shaw et al. 1995; Paulsen 1996; Podolsky et al. 1997; Pfrender 1998; Roff and Mouseau 1999; Roff et al. 1999). The connection of the matrix to genetic constraints also motivates interest in *G*-matrix constancy. Comparison of *G*-matrices can be used to detect and localize evolutionary change in genetic constraints. Although change in constraints is often invoked to explain evolutionary patterns, such arguments are usually not bolstered—as they could be—by demonstrating a difference in *G*-matrices.

The comparison of matrices is a difficult statistical problem and many different approaches have been suggested over the last few decades. There has been some controversy regarding the appropriate statistic to use as well as the framework within which that statistic is tested (Shaw 1987, 1991, 1992; Cowley and Atchley 1992). Flury (1988) outlined a comprehensive and elegant solution to the matrix comparison problem. This solution rests on the realization that matrices can still share similarities in structure even if they are not strictly equal to one another. The companion paper (Phillips and Arnold 1999) outlines the application of this approach to quantitative genetic data in some detail. Here, we emphasize that this more general view of matrix comparison provides a detailed diagnosis of steps in *G*-matrix evolution.

Our goal in this paper is to compare the *G*-matrices of two populations of the garter snake, *Thamnophis elegans*, that are separated by a distance of nearly 300 km in northern California and have very different ecologies. Our coastal population is terrestrial and feeds primarily on slugs. Our inland population is aquatic and feeds primarily on aquatic prey (fish, anurans, and leeches; Arnold 1981a,b,c, 1992b). Analyses of the mitochondrial genome (which will be reported in a later paper) suggest that these populations may have diverged nearly 2 million years ago. Furthermore, two of the six meristic traits upon which our *G*-matrices are based show substantial divergence in mean (Arnold 1988). These and other results suggest that our populations have experienced divergent selection for a considerable period of time (Kelley et al. 1997). Thus, there are considerable grounds for expecting differences in *G*-matrices. In pursuit of our goal of geographic comparison, we first compare the *G*-matrices of males and females within our populations. We find detectable differences between the sexes in *G*-matrices and so proceed

[2] Present address: Department of Zoology, Oregon State University, Corvallis, Oregon 97331; E-mail: arnolds@bcc.orst.edu.

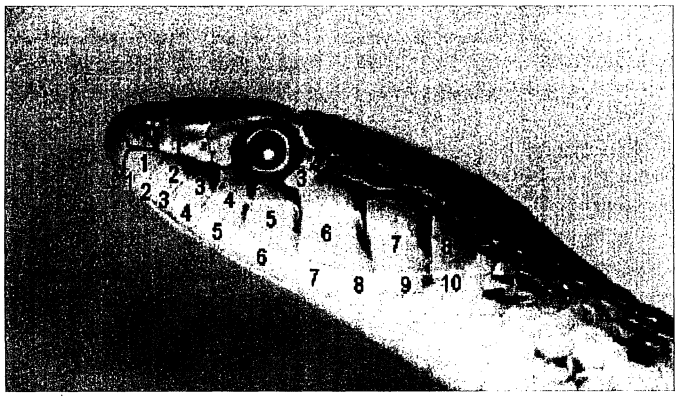

FIG. 1. Counting conventions for three meristic traits on the head of *Thamnophis elegans* (SJA 24300): number of scales on the upper lip (SLAB), number on the lower lip (ILAB), and number behind the eye (POST). The sum on counts on the right and left sides was used in all analyses.

by doing separate geographic comparisons of males and females. The end result of these comparisons suggests evolutionary conservation of the structure of both the genetic and environmental variance-covariance matrices.

MATERIALS AND METHODS

Samples of Snakes and Scoring of Traits

Samples.—Newborn snakes were obtained from pregnant females collected at two localities situated 282 km apart in northern California; at the mouth of the Mad River, Humboldt County (coastal), and Eagle Lake, Lassen County (inland; Arnold 1988). Females were obtained at two collecting stations at each of these localities, but, for simplicity of analysis, data were pooled for each locality after correcting for differences in station means. The justification for this pooling was that trait divergence between stations within localities was trivial compared with the divergence among localities. Thus, the two traits that differed most between stations (body and tail vertebral numbers) showed an order of magnitude less differentiation between stations within localities than between localities (Arnold 1988). Despite the small magnitude of divergence between sampling stations, the trait means were corrected for station differences so that those differences would not inflate within-population estimates of genetic variances and covariances.

After capture, females were transported to the laboratory, where they were maintained under uniform conditions until their litters were born (Arnold 1988). A total of 102 females, yielding 911 offspring, were sampled from the coastal population, whereas 156 females, yielding 1505 offspring, were sampled from the inland population.

Trait Scores.—The following scale counts were made on ethanol-preserved mothers and their offspring (Fig. 1): number of ventral scales (VENT); number of subcaudal scales (SUB); total number of infralabial (ILAB), supralabial (SLAB), postocular (POST) scales on the left and right sides; number of dorsal scale rows at midbody (MID). Conventions for making these counts are described by Ruthven (1908), Dowling (1951), and Peters (1964). The counts were made on stillborn animals (approximately 5% of the neonate sample), even when deformed, as well as on liveborn neonates. Counts were treated as missing values under the following circumstances: when the tail tip lacked the usual conical scale due to developmental abnormality or amputation (SUB), when eyes (POSTOC) or jaws (ILAB, SLAB) were missing due to developmental abnormality (POSTOC); when sutures between scales failed to form for at least half the length of the jaw (ILAB, SLAB). VENT and SUB correspond, respectively, to numbers of body and tail vertebrae (Alexander and Gans 1966; Voris 1975). None of these counts change during the postnatal life of an individual. Sex of neonates was determined at the time of preservation by eversion of hemipenes.

Parameter Estimation

Quantitative genetic parameters were estimated using a female-offspring regression of the female value on the litter average, with the number of offspring in the litter serving as a weighting factor (Falconer and Mackay 1996). Additive genetic variances (G_{ii}) were calculated as twice the parent-offspring covariance for the same trait across generations, whereas the additive genetic covariances (G_{ij}) were calculated as the average of the two cross-covariances across generations (i.e., trait 1 in parents on trait 2 in offspring and vice versa). This averaging of the two cross-covariances assumes that most loci affecting even sexually dimorphic characters are autosomal rather than sex linked, an assumption supported by empirical work (Lande 1980c, Cowley and Atchley 1988). Because of the possibility of natural selection in the parental generation and the much larger sample of individuals in the offspring generation, phenotypic variances and covariances were estimated using the offspring rather than the parents. The phenotypic variance components (P) were estimated as the sum of the within- and among-family variance components from a one-way analysis of variance of the full-sib families. The environmental variance and covariance components were then calculated as the difference between the phenotypic and additive genetic components ($E = P - G$). Standard errors on these values were calculated using a bootstrapping approach (Efron 1982; Efron and Tibshirani 1993) by repeatedly recalculating the quantitative genetic parameters after randomly resampling with replacement from the population of full-sib families. Errors were calculated as the variances of the empirically derived error distributions after resampling 10,000 times. All covariance component estimation calculations were carried out using the software package *H2boot* (Phillips 1998a).

Although covariance component analysis yields the correct parameters for quantitative genetic analysis, the covariance components themselves are often not statistically well behaved. In particular, the possibility of negative variances and correlations that are greater than 1.0 or less than -1.0 can make the analysis of variance-covariance matrices difficult (for a discussion of the issues, see Phillips and Arnold 1999). Some of these problems can be overcome by estimating the parameters directly from the individual values or family means rather than from the covariance components. For example, the phenotypic variances and covariances were also

estimated directly using the entire population of offspring (i.e., individual values). Similarly, the additive genetic variances and covariances can be estimated from the covariance among-family means. In both of these cases, the variances and covariances will be well behaved by the above criteria, but they are unfortunately biased estimates. The phenotypic estimates are biased by the among-family variance, whereas the genetic estimates are biased by the within-family variance (Via 1984; Lynch and Walsh 1998). The complications introduced by this bias have to be weighed against the fact that these parameters can be used in a well-established parametric analysis (Flury 1988), whereas the covariance components can only be tested using resampling methods (see below). In the analysis presented below, both the covariance component and the individual value estimates will be used where appropriate so that the strengths and weaknesses of each type of estimate can be used to balance against one another.

Multiple paternity does not complicate either approach to genetic parameter estimation. Multiple paternity of broods has been confirmed with molecular markers in several genera of snakes, including *Thamnophis* (Schwartz et al. 1989; Barry et al. 1992; Luiselli 1995). Such multiple paternity does not affect estimates based on offspring-mother regressions. Multiple paternity could affect our estimates based on litter means, which assume that the coefficient of relationship within litters is 0.5 (littermates full-sibs). This assumption does not appear to introduce any serious errors. In the one study of multiple paternity in garter snakes, the coefficient of relationship within litters was estimated to be 0.42–0.44, depending on whether the estimate is based on the confirmed or estimated incidence of multiple paternity (Schwartz et al. 1989). Thus, despite multiple paternity of 50–72% of litters, littermates were on the average nearly full-sibs. An error in either of our populations in the coefficient of relationship within litters would merely have a proportional affect on the *G*-matrix, and the Schwartz et al. (1989) study suggests that such errors would be slight. A potentially more serious problem could arise if our populations differed in the coefficient of relationship within litters. Even in this case, however, only proportional differences in *G*-matrices would result. Comparative studies of paternity in snake populations could evaluate this possibility, but have as yet not been conducted.

Matrix Comparisons

Matrices were compared using a hierarchical approach derived by Flury (1988) and applied to quantitative genetic analysis in Phillips and Arnold (1999). The essential point here is that matrix comparison is a multivariate problem with many degrees of similarity being possible between the extremes of equality and inequality. The hierarchy follows a progression from unrelated structure to partial common principal components (PCPC) to common principal components (CPC) to proportionality and finally to equality.

The testing procedure entails evaluating each pair of matrices (as well as all matrices simultaneously) at each level of the hierarchy. We use the jump-up approach (Phillips and Arnold 1999) to evaluate at which level in the hierarchy the two matrices could no longer be considered similar. Covariance component matrices were compared using a randomization approach (Phillips and Arnold 1999) with the program *CPCrand* (Phillips 1998b). Parametric tests (Flury 1988) were used to compare the product-moment-based matrices (e.g., family mean covariances) with the program *CPC* (Phillips 1998c). Because of the large number of tests involved, we do not present results for the entire hierarchy of comparisons, but instead report the highest point in the hierarchy for which the null model could not be rejected. An analysis of the coastal-inland matrix comparison for females detailing all statistical aspects of the comparisons is given in Phillips and Arnold (1999).

TABLE 1. Trait means for male and female offspring from coastal and inland populations of *Thamnophis elegans*. Population differentiation (Diff.) is expressed as the average geographic difference of males and females in units of average phenotypic standard deviation.

	Coastal		Inland		
Trait	Male n = 406[1]	Female n = 505[2]	Male n = 762[3]	Female n = 743[4]	Diff.
VENT	156.277	151.728	172.422	167.520	3.998
SUB	80.982	71.247	90.535	80.589	2.347
MID	18.776	18.920	20.151	20.486	1.967
ILAB	18.573	18.745	20.004	20.159	1.557
SLAB	14.665	15.080	15.879	15.876	1.181
POST	5.722	5.715	5.920	6.085	0.413

[1] Sample sizes for the six traits are, in order, 393, 389, 398, 405, 406, and 406.
[2] Sample sizes for the six traits are, in order, 481, 493, 498, 502, 502, and 505.
[3] Sample sizes for the six traits are, in order, 744, 739, 761, 760, 760, and 762.
[4] Sample sizes for the six traits are, in order, 712, 708, 739, 740, 735, and 743.

RESULTS

Means

Differences between the Sexes.—The six traits varied greatly in degree of sexual dimorphism (Table 1). SUB was the most sexually dimorphic trait, with males showing an average of 10 more subcaudal scales than females, a 12–13% difference in means. VENT also showed appreciable sexual dimorphism, with males showing an average of five more ventral scales than females, a 3% difference in means (Arnold 1988). Sexual dimorphism in the other traits was less substantial (0–3%) and usually in the same direction, with females averaging higher counts than males. MANOVA showed significant overall sexual dimorphism, and individual ANOVAs showed sexual dimorphism was significant for each of the six traits (Table 2). MID, SLAB, and POST showed a significant geographic difference in sexual dimorphism and all of the traits, except MID showed litter differences in degree of sexual dimorphism (Table 2).

Geographic Comparisons.—The six traits also differed in degree of differentiation between coastal and inland populations (Table 1). VENT and SUB showed the greatest differentiation, with the inland population means exceeding the coastal means by 4.0 and 2.3 phenotypic standard deviations, respectively. The inland population showed higher means for the other traits as well, with the degree of differentiation ranging from 0.4 to 2.0 phenotypic standard deviations.

TABLE 2. Multivariate analysis of variance testing for geographic differences (pop), litter differences (litter), and sexual dimorphism (sex) in six traits of *Thamnophis elegans* offspring. Tabulated results give the ANOVA *F*-values for signal traits and an *F*-ratio based on Wilk's lambda for the MANOVA with all traits combined.

Source	df	VENT	SUB	MID	ILAB	SLAB	POST	MANOVA
Pop[1]	1	1457.5 ****	522.3 ****	453.1 ****	183.3 ****	96.2 ****	24.4 ****	340.4 ****
Litter(pop)[2]	257	3.8 ****	3.4 ****	3.0 ****	3.3 ****	4.1 ****	3.0 ****	3.3 ****
Sex[2]	1	556.1 ****	2116.3 ****	21.9 ****	10.0 **	10.7 **	13.4 ***	403.0 ****
Pop*sex[2]	1	0.7 ns	0.0 ns	11.3 ***	0.0 ns	23.1 ****	7.9 **	7.4 ****
Litter*sex(pop)	246	1.2 **	1.5 ****	1.1 ns	1.3 **	1.5 ****	1.2 *	1.3 ****
Residual	1745							

**** $P < 0.0001$; *** $P < 0.001$; ** $P < 0.01$; * $P < 0.05$; ns, nonsignificant.
[1] Using Type III mean squares for litter(pop) as denominator in *F*-test.
[2] Using Type III mean squares for litter*sex(pop) as denominator in *F*-test.

MANOVA showed that significant geographic differences for all the traits, as well as significant differences among families within populations (Table 2).

Variance-Covariance Matrices and Heritabilities

The phenotypic matrices (Table 3) were nearly diagonal in structure, with average correlations in the range of 0.04–0.10. The genetic matrices (Table 4) showed more off-diagonal structure, with average correlations in the range of 0.15–0.45. In contrast, the environmental matrices (Table 5) are almost diagonal in structure, with average correlations in the range of −0.06–0.12.

All of the traits show some evidence of heritable variation (Tables 2, 6). VENT and SUB show the highest heritabilities (h^2 = 0.41–0.79) and POST shows the lowest (h^2 = 0.00–0.62). Geographic differences are apparent for some traits (e.g., regression estimates for SLAB), but sexual differences appear to be present for other traits (e.g., regression estimates for POST). To evaluate these apparent differences and to take account of all the traits at once, we will use a hierarchical approach.

Matrix Comparisons

In general the *G*-matrices show common structure in all of their principal components (Table 7). The exceptions to this rule involve the *G*-matrix for the coastal males. The regression estimates (covariance components) for coastal males shares one to six principal components with the other matrices, whereas the family-mean estimates (product-moment components) share three to six principal components. The differences among these estimates are undoubtedly caused by unusually high genetic correlations among some of the trait combinations involving MID and ILAB. Here the covariance component analyses, which should be more sensitive to ill-behaved matrices, suggest less similarity than the parametric, family-mean-based analyses, which should have fewer difficulties with problematic matrix structure. It is currently not known which estimation procedure is best in this case, because each has its own problems. The overall result of some shared similarity is preserved across both analyses, however, and in comparisons not involving the coastal males, both methods yield the same answer.

In contrast to the results for the *G*-matrices, we cannot reject the hypothesis of equality for any pair of *E*-matrices (Table 7). Thus, not only do the *E*-matrices show common principal components, as do the *G*-matrices, they also show common eigenvalues. At least part of this similarity result is caused by the sometimes large errors on the *E*-matrix elements (Table 5), so although we cannot rule out equality in these data, it is likely that there are at least some true differences among the matrices in particular elements. These differences are small relative to the overall pattern of similarity, however. Further, matrices that are nearly diagonal are much more likely to be similar to one another than those displaying more complex structure.

The results for the *P*-matrices are generally in-between those for the *G*- and *E*-matrices, as might be expected. The *P*-matrix is the sum of the *G*- and *E*-matrix and so we expect the result of a *P*-matrix comparison to be close to the average of *G*- and *E*-matrix comparisons. Thus, the comparisons show either six common principal components (full CPC) or equality (covariance component estimates) or two to six common principal components (individual estimates). In all six comparisons, those based on the covariance component estimates indicated more similarity in matrices than comparisons based on the individual estimates. This difference undoubtedly reflects the greater statistical power of the individual estimates because within each sex there were approximately five times as many individuals as families.

A comparison of the results in Table 7 with calculations of matrix correlation reveals how much structural differentiation can be concealed by the latter approach. For phenotypic matrices, the product moment matrix correlations for the six comparisons shown in Table 7 ranged from 0.96 to 1.00. Clearly, the phenotypic matrices are highly correlated across populations, but the simple correlation coefficients give no indication that the matrices are neither equal nor proportional or that they differ in extent of shared principal components. Likewise, the matrix correlations for *G* vary

TABLE 3. Phenotypic variance-covariance matrices (± SE) for male and female offspring from coastal and inland populations of *Thamnophis elegans*. Indicated sample sizes are numbers of litters, which vary from element to element in the matrix because of missing values.

	Coastal		Inland	
Trait(s)	Male n = 75–102	Female n = 75–102	Male n = 116–156	Female n = 116–156
VENT	11.095 ± 1.360	15.706 ± 2.137	22.967 ± 5.555	15.140 ± 1.342
VENT, SUB	0.999 ± 1.883	4.440 ± 3.019	4.875 ± 6.431	3.849 ± 2.219
VENT, MID	0.020 ± 0.997	0.073 ± 1.454	0.079 ± 3.940	−0.163 ± 0.943
VENT, ILAB	−0.168 ± 0.967	0.286 ± 1.501	0.283 ± 3.986	0.129 ± 0.953
VENT, SLAB	0.507 ± 1.106	0.274 ± 1.575	0.084 ± 4.065	0.197 ± 1.005
VENT, POST	0.079 ± 0.947	0.238 ± 1.556	0.163 ± 4.254	−0.049 ± 0.971
SUB	13.919 ± 1.609	13.272 ± 1.552	21.174 ± 2.181	17.051 ± 1.744
SUB, MID	0.044 ± 1.163	−0.296 ± 1.038	−0.110 ± 1.629	0.230 ± 1.230
SUB, ILAB	0.400 ± 1.199	0.403 ± 1.134	0.085 ± 1.549	0.133 ± 1.248
SUB, SLAB	−0.049 ± 1.154	0.043 ± 1.150	−0.010 ± 1.554	−0.263 ± 1.227
SUB, POST	0.367 ± 1.271	0.057 ± 1.119	0.109 ± 1.566	0.101 ± 1.234
MID	0.382 ± 0.058	0.218 ± 0.039	1.023 ± 0.042	0.799 ± 0.060
MID, ILAB	−0.002 ± 0.112	0.107 ± 0.092	0.009 ± 0.055	0.014 ± 0.073
MID, SLAB	0.085 ± 0.086	0.051 ± 0.061	−0.018 ± 0.083	0.039 ± 0.154
MID, POST	0.023 ± 0.074	0.016 ± 0.058	−0.038 ± 0.054	−0.055 ± 0.063
ILAB	1.472 ± 0.115	1.186 ± 0.077	0.380 ± 0.046	0.542 ± 0.070
ILAB, SLAB	0.290 ± 0.129	0.216 ± 0.110	0.057 ± 0.074	0.031 ± 0.129
ILAB, POST	0.046 ± 0.122	0.042 ± 0.087	0.054 ± 0.058	0.023 ± 0.069
SLAB	0.948 ± 0.082	0.913 ± 0.058	0.370 ± 0.089	0.751 ± 0.177
SLAB, POST	0.044 ± 0.090	0.060 ± 0.082	0.033 ± 0.075	0.034 ± 0.128
POST	0.452 ± 0.070	0.445 ± 0.061	0.450 ± 0.054	0.546 ± 0.061
Average correlation (r_p)	0.082	0.100	0.087	0.044

from 0.93 to 0.98 and those for E vary from 0.73 to 0.95. Coefficients of matrix correlation bear no relationship to position in the Flury hierarchy.

The principal component (eigenvector) structure of the matrices yields some insight into what CPC structure is maintained between the two populations. Table 8 shows the combined common principal component solution for the G-matrices of both sexes from the coastal and inland populations (the principal components of the original matrices show similar structure). (Results here are slightly different from those

TABLE 4. Genetic variance-covariance matrices (± SE) for male and female offspring from coastal and inland populations of *Thamnophis elegans*. Indicated sample sizes are numbers of litters, which vary from element to element in the matrix because of missing values. Note: Pooled estimates in appendix 1 of Arnold (1988) involving VENT and SUB are in error by a factor of about four and should be ignored.

	Coastal		Inland	
Trait(s)	Male n = 75–102	Female n = 75–102	Male n = 116–156	Female n = 116–156
VENT	5.020 ± 1.845	6.877 ± 1.970	9.055 ± 1.905	8.172 ± 1.699
VENT, SUB	1.464 ± 1.270	0.189 ± 1.658	2.112 ± 1.559	3.779 ± 1.668
VENT, MID	0.089 ± 0.199	−0.035 ± 0.252	0.382 ± 0.368	0.088 ± 0.293
VENT, ILAB	0.419 ± 0.380	1.147 ± 0.466	0.410 ± 0.324	0.652 ± 0.212
VENT, SLAB	0.745 ± 0.334	0.296 ± 0.354	0.105 ± 0.161	0.066 ± 0.175
VENT, POST	0.346 ± 0.210	−0.037 ± 0.197	−0.247 ± 0.162	0.115 ± 0.196
SUB	9.201 ± 2.326	7.800 ± 2.655	9.500 ± 2.095	8.159 ± 1.727
SUB, MID	−0.298 ± 0.334	−0.022 ± 0.300	0.211 ± 0.393	0.284 ± 0.404
SUB, ILAB	0.848 ± 0.646	0.032 ± 0.479	0.199 ± 0.197	0.238 ± 0.309
SUB, SLAB	0.767 ± 0.441	0.117 ± 0.348	−0.066 ± 0.192	0.225 ± 0.214
SUB, POST	−0.118 ± 0.235	−0.113 ± 0.252	−0.048 ± 0.200	0.173 ± 0.234
MID	0.044 ± 0.056	0.007 ± 0.033	0.510 ± 0.104	0.271 ± 0.097
MID, ILAB	0.078 ± 0.094	0.044 ± 0.061	−0.025 ± 0.055	0.079 ± 0.047
MID, SLAB	0.149 ± 0.059	0.105 ± 0.058	−0.004 ± 0.047	0.046 ± 0.061
MID, POST	−0.049 ± 0.033	0.026 ± 0.039	−0.076 ± 0.049	−0.097 ± 0.049
ILAB	0.264 ± 0.193	0.359 ± 0.135	0.049 ± 0.038	0.025 ± 0.047
ILAB, SLAB	0.512 ± 0.124	0.468 ± 0.110	−0.031 ± 0.034	−0.016 ± 0.038
ILAB, POST	0.085 ± 0.073	−0.000 ± 0.065	0.036 ± 0.040	0.030 ± 0.049
SLAB	0.427 ± 0.131	0.366 ± 0.118	0.027 ± 0.019	0.024 ± 0.027
SLAB, POST	0.046 ± 0.075	0.049 ± 0.055	0.027 ± 0.026	0.021 ± 0.029
POST	−0.031 ± 0.052	0.098 ± 0.045	−0.015 ± 0.032	0.088 ± 0.047
Average correlation (r_G)	0.455	0.372	0.152	0.384

TABLE 5. Environmental variance-covariance matrices (± SE) for male and female offspring from coastal and inland populations of *Thamnophis elegans*. Indicated sample sizes are numbers of litters, which vary from element to element in the matrix because of missing values.

	Coastal		Inland	
Trait(s)	Male n = 75–102	Female n = 75–102	Male n = 116–156	Female n = 116–156
VENT	6.076 ± 2.140	8.829 ± 2.752	13.912 ± 5.475	6.968 ± 1.457
VENT, SUB	−0.465 ± 1.932	4.251 ± 3.479	2.763 ± 6.489	0.070 ± 2.509
VENT, MID	−0.069 ± 0.988	0.108 ± 1.503	−0.303 ± 3.949	−0.251 ± 0.971
VENT, ILAB	−0.587 ± 1.033	−0.861 ± 1.586	−0.127 ± 4.021	−0.524 ± 0.949
VENT, SLAB	−0.238 ± 1.096	−0.022 ± 1.643	−0.021 ± 4.057	0.131 ± 1.014
VENT, POST	−0.266 ± 0.980	0.275 ± 1.594	0.410 ± 4.302	−0.165 ± 0.996
SUB	4.718 ± 2.434	5.473 ± 2.456	11.673 ± 2.629	8.891 ± 2.040
SUB, MID	0.342 ± 1.255	−0.274 ± 1.086	−0.321 ± 1.717	−0.054 ± 1.276
SUB, ILAB	−0.447 ± 1.284	0.371 ± 1.220	−0.114 ± 1.560	−0.106 ± 1.269
SUB, SLAB	−0.816 ± 1.164	−0.074 ± 1.182	0.056 ± 1.603	−0.488 ± 1.252
SUB, POST	0.485 ± 1.271	0.169 ± 1.114	0.157 ± 1.601	−0.072 ± 1.201
MID	0.339 ± 0.083	0.210 ± 0.050	0.513 ± 0.108	0.528 ± 0.106
MID, ILAB	−0.080 ± 0.141	0.062 ± 0.107	0.034 ± 0.071	−0.065 ± 0.079
MID, SLAB	−0.064 ± 0.095	−0.054 ± 0.080	−0.014 ± 0.082	−0.007 ± 0.141
MID, POST	0.072 ± 0.086	−0.009 ± 0.063	0.038 ± 0.077	0.042 ± 0.075
ILAB	1.209 ± 0.216	0.827 ± 0.149	0.331 ± 0.059	0.516 ± 0.086
ILAB, SLAB	−0.222 ± 0.146	−0.252 ± 0.128	0.088 ± 0.085	0.047 ± 0.146
ILAB, POST	−0.039 ± 0.144	0.042 ± 0.100	0.017 ± 0.075	−0.003 ± 0.083
SLAB	0.521 ± 0.135	0.574 ± 0.121	0.343 ± 0.092	0.727 ± 0.188
SLAB, POST	−0.001 ± 0.098	0.010 ± 0.087	0.006 ± 0.075	0.013 ± 0.127
POST	0.484 ± 0.089	0.347 ± 0.071	0.465 ± 0.059	0.458 ± 0.070
Average correlation (r_E)	−0.055	0.050	0.117	−0.029

presented in Phillips and Arnold [1999] because all four matrices are evaluated simultaneously, rather than just the female matrices.) The first two principal components are the correlated submatrix for the traits VENT and SUB. The third and sixth principal components describe a correlated submatrix for the traits ILAB and SLAB. The fourth and fifth principal components describe a correlated submatrix for the traits MID and POST. The main difference between these coastal and inland *G*-matrices lies in their eigenvalues, which are often strikingly different (Table 8). The *E*-matrices for coastal and inland females share a principal component structure that is similar to the *G*-matrix structure. Unlike the *G*-matrices, we cannot show that the eigenvalues of the *E*-matrix are different.

The shared principal component structure can perhaps be best visualized by graphically displaying the pattern of covariance structure using 95% confidence ellipses. Focusing on the first two components, those dominated by the VENT/SUB submatrix, it is clear that the *G*-matrices from all of the populations and sexes have an overall similar orientation in two-dimensional space, although the amount of variance along any particular axis (the length and width of the ellipse) varies from matrix to matrix (Fig. 2). The similarity of orientation (eigenvectors) yields the CPC results, whereas the difference in variances (eigenvalues) precludes the possibility of this similarity leading to proportionality or equality. It is also clear that the pattern of divergence between populations is oriented along the primary axis of genetic variation (Schluter 1996).

DISCUSSION

Parameters Estimated

Some of the technical aspects of parameter estimation should be kept in mind in later discussions of sexual and geographic comparisons. Because the traits are generally sexual dimorphic, the estimates based on son-mother regressions are across-sex covariances and will represent within-sex

TABLE 6. Heritability estimates (± SE) for male and female offspring from coastal and inland populations of *Thamnophis elegans* based on mother-offspring regressions. Indicated sample sizes are numbers of litters, which vary within populations because of missing trait values. Significance levels are as given in Table 2.

	Coastal		Inland	
Trait(s)	Male n = 75–102	Female n = 75–102	Male n = 116–156	Female n = 116–156
VENT	0.45 ± 0.17**	0.44 ± 0.14***	0.41 ± 0.11***	0.54 ± 0.10***
SUB	0.66 ± 0.16***	0.59 ± 0.18***	0.45 ± 0.10***	0.48 ± 0.10***
MID	0.12 ± 0.15 ns	0.04 ± 0.15 ns	0.50 ± 0.10***	0.34 ± 0.12***
ILAB	0.18 ± 0.13 ns	0.30 ± 0.12**	0.13 ± 0.10 ns	0.05 ± 0.09 ns
SLAB	0.45 ± 0.13***	0.40 ± 0.13***	0.08 ± 0.06 ns	0.04 ± 0.04 ns
POST	−0.07 ± 0.12 ns	0.22 ± 0.10**	−0.04 ± 0.07 ns	0.16 ± 0.08*

TABLE 7. Results of hierarchical comparisons of matrices from male and female offspring from coastal and inland *Thamnophis elegans* populations. Two estimates of phenotypic matrices were compared: an estimate derived from the within- and among-litter components of variance and covariance (cov. comp.) and an estimate derived using each individual offspring as a datapoint (individual). Likewise, two estimates of genetic matrices were compared: an estimate derived from regression of litter means on maternal values (regression) and an estimate derived from the variances and covariances among litter means (family-mean). The covariance component matrices were compared using a resampling approach, whereas the point-estimate matrices were compared using a parametric approach. The environmental matrix was estimated as the difference between the cov. comp. estimate of the **P**-matrix and the regression estimate of the **G**-matrix. The first six columns show the results of all possible pairwise comparisons. The last column shows the results of simultaneous comparison of all four matrices. Table entries give the highest point in the hierarchy at which the listed null hypothesis could not be rejected.

	Coastal male vs. coastal female	Coastal male vs. inland male	Coastal male vs. inland female	Coastal female vs. inland male	Coastal female vs. inland female	Inland male vs. inland female	All together
Phenotypic:							
Cov. comp.	Equal	Full CPC	Full CPC	Full CPC	Full CPC	Full CPC	Full CPC
Individual	Full CPC	CPC(2)	CPC(4)	CPC(4)	CPC(3)	CPC(4)	CPC(2)[1]
Genetic:							
Regression	CPC(2)	CPC(1)	Full CPC	Full CPC	Full CPC	Full CPC	CPC(2)[2]
Family-mean	Full CPC	CPC(3)	CPC(4)	Full CPC	Full CPC	Full CPC	CPC(4)[2]
Environmental:	Equal	Equal	Equal	Equal	Equal	Equal	Equal

[1] CPC(4) when coastal males are excluded.
[2] Full CPC when coastal males are excluded.

(male) variances and covariances to the extent that genes are expressed similarly in males and females. Estimates based on daughter-mother regressions represent within-sex (female) variances and covariances (Lynch and Walsh 1998). For simplicity we refer to the former as "male" estimates and the latter as "female" estimates.

Common Principal Components in the G-Matrices

One of our most important findings is that the *G*-matrix shows conservation of structure in the coastal-inland comparison. The theme that emerges from our geographic comparisons is that principal components of the *G*-matrix are conserved but eigenvalues of the matrix vary (Table 7). Conservation of principal components can apply to only the first principal component or the full set of principal components, depending on which matrices are compared and on the method of estimating *G*. Thus, both methods of estimation agree that the first principal component of the *G*-matrix is conserved during the coastal-inland divergence. Comparison of sequences for two mitochondrial genes (about 1600 base pairs from cytochrome-*b* and ND2) revealed a sequence divergence of 4.2% for our coastal and inland populations (M. Pfrender, M. Alfaro, and S. Arnold, in prep.). Using an expected evolutionary rate of about 2% sequence divergence per million years based on mammalian mtDNA (Brown 1983), our populations might have separated as long ago as 2 million years. Such an estimate of expected divergence time is predicated on multiple assumptions and is subject to a large error of estimation (Hillis et al. 1996). Nevertheless, it is fair to say that principal components were conserved on a time scale much longer than local population divergence and approaching—if not equaling—species-level divergence (note also the large divergence in means for most of the traits; Table 1). Thus, our geographic comparisons suggest conservation of principal components for a considerable period of time. In contrast, the eigenvalues of the *G*-matrix differ between males and females and geographically. Thus, the eigenvalues of the *G*-matrix seem prone to relatively rapid evolution. Furthermore, differentiation in eigenvalues is not of the simplest type. The eigenvalues for our ensemble of traits do not differ by a single constant of proportionality. Instead, eigenvalues show a chaotic pattern in geographic comparisons; some are larger in the inland population, others are larger in the coastal population (Table 8). What might be the causes of conservation of eigenstructure on the one hand and of disruption of eigenvalues on the other?

Conserved patterns of selection and mutation are both capable of explaining phylogenetic persistence in *G*-eigenstructure. The *G*-matrix evolves in response to both mutation and selection and the matrix at equilibrium is a compromise between the patterns imposed by those two forces (Lande 1980a). The easiest explanation for phylogenetic persistence of *G*-eigenstructure is phylogenetic persistence in eigenstructures of the matrices describing multivariate stabilizing selection and mutation-recombination. Alternatively, one of these matrices might be conserved and swamp out fluctuations in the other, or, the least likely explanation, both matrices might vary but with fluctuations that cancel out each other. Geographic divergence in the means of traits implies

TABLE 8. Common principal components for the genetic matrices of coastal and inland males and females. The eigenvectors represent the solution for a model constrained to have common principal components.

	Eigenvectors					
Trait	1	2	3	4	5	6
VENT	0.615	0.777	−0.058	0.009	−0.055	−0.102
SUB	0.785	−0.618	−0.014	0.017	0.038	0.015
MID	0.012	0.077	0.165	0.692	0.635	0.288
ILAB	0.070	0.072	0.590	−0.248	−0.223	0.727
SLAB	0.002	−0.032	0.774	0.166	−0.105	−0.601
POST	0.016	0.041	0.145	−0.657	0.728	−0.124
Eigenvalues:						
Coastal males	7.609	3.423	0.880	0.325	0.081	0.126
Inland males	10.164	6.512	0.469	0.683	0.591	0.522
Coastal females	5.726	5.244	1.027	0.273	0.446	0.031
Inland females	5.850	5.271	1.005	0.373	0.249	0.001

that some aspects of selection have changed since descent from a common ancestor. A plausible model is that the location of adaptive peaks has changed, but the shape of the selection surface (especially its eigenstructure) has remained the same. Conservation of the eigenstructure of selection in turn produces conservation of *G*-matrix eigenstructure.

The pattern of mutational input into the system should also have an influence on the pattern of genetic covariance (Lande 1980a). Mutations with pleiotropic effects contribute to the maintenance of genetic correlations between traits. If the variance introduced by new mutations were constrained to follow the same pathway of influence by, for example, the organisms developmental system, then the pattern of genetic correlation might evolve to reflect these pathways (Cheverud 1984; Riska 1986; Slatkin 1987; Houle 1991). Alternatively,

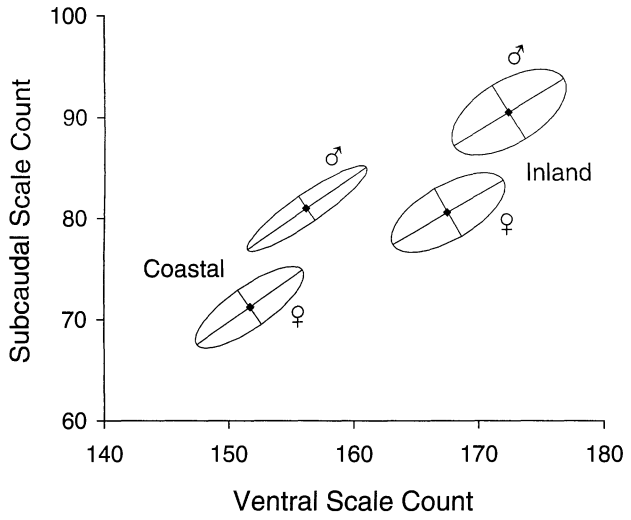

FIG. 2. Variance-covariance ellipses representing the first two principal components (and associated eigenvalues) for the genetic matrices of male and female offspring from coastal and inland populations of *Thamnophis elegans*. Under the assumption of normality, the scale of the ellipses encompasses 95% of the underlying genetic variation in the population. The figure shows the projections of the first two principal components onto the VENT and SUB axes. Because these are projections from a higher-dimensional space, the actual regions are will not be completely elliptical, but are presented as so here for clarity. Note the common principal component orientation despite differences in variance along each axis.

changes in population size might explain eigenvalue fluctuations. To model the effects of drift, we can consider a set of populations that are all of the same finite size and descended from the same ancestral population. Although the modal population in this set should show proportionality in eigenvalues compared to the ancestral population (Lande 1979; Lofsvold 1988; Roff and Mousseau 1999; Roff et al. 1999), the variation about this mode is likely to be huge. Consequently, almost any pattern of eigenvalue differences could be produced by drift (P. C. Phillips, M. C. Whitlock, and K. Fowler, in prep.). However, drift cannot explain the gender differences that we observed in *G*-eigenvalues within populations. Assuming that our populations are at equilibrium under selection, the simplest explanation for differences in eigenvalues is that males and females and our two populations experience different strengths of multivariate stabilizing selection. Continuing the model developed in the preceding paragraph, the selection surface keeps its principal components intact even though its curvature and the location of its peaks change.

Equality of E-Matrices

Equality of *E*-matrices was an unanticipated but intriguing result. Relative constancy of *E* might reflect phylogenetic conservatism in both patterns of multivariate stabilizing selection and polygenic mutation. Stabilizing selection should favor genotypes near the optimum that have small environmental deviations from their genotypic means. Over time, this process should build up a zone of canalization in the vicinity of the selective optimum (Fisher 1930; Schmalhausen 1949; Waddington 1957; Lande 1980b). If the selection surface is a multivariate hill (concave downward in all dimensions), as seems likely for the traits considered here, an evolutionarily persistent surface should foster a corresponding multivariate pattern of canalization (Wagner et al. 1997; Rice 1998). Here we imagine that selection keeps its orientation and curvature, even though the location of the optimum in character space may change in evolutionary time, as the character mean tracks the position of the optimum. *E* also reflects the nonadditive effects of new mutations each generation, so constancy of *E* implies persistence in the multivariate pattern of mutation, perhaps because of conservatism in developmental programs. Whether constancy of *E*-

matrices arises from conservatism in multivariate stabilizing selection or in developmental programs, tests for constancy should be made in other taxa and for other character sets.

Steppan (1997a,b) has made the important point that to understand matrix evolution we must move beyond pairwise comparisons and take a phylogenetic perspective. The best data for such comparative work are *G*-matrices estimated for many populations for which there is a well-resolved phylogeny. Amassing such a dataset is a formidable undertaking. In most taxa, however, *P* is substantially easier to estimate than *G* because only data on individuals rather than on families are required. Thus, it will often be possible to obtain a relatively large sample of *P*-matrices, even though *G* is difficult to estimate (e.g., Steppan 1997a). An observation of constancy of *E* suggests a method for estimating *G*-matrices in this situation. Suppose that a few estimates of *G* and *E* are available and yield the result, as in our study, that *E* is constant, although *G* varies. In that circumstance, the data from all populations could be pooled to give a best estimate of *E*, which could then be used to estimate *G* for each population in the larger set using the relationship $G = P - E$. In our study a pooled estimate of *E* could be made across both sexes using over 250 families. Such estimates of *G* should take account of sampling variation in both *P* and *G*, but that accounting will be straightforward. The method just described seems preferable to using *P* as a simple surrogate for *G* in comparative studies of the *G*-matrix (Cheverud 1988, 1995; Willis et al. 1991; Roff 1995, 1996; Lynch and Walsh 1998).

How general is the constancy of *E*? This question is difficult to answer from the available literature because the empirical focus has been on the *G*-matrix. Kohn and Atchley (1988) found that the *E*-matrices for two genera of murine rodents were completely dissimilar. Our results suggest, however, that comparisons of *E*-matrices might be profitable.

Comparison of P-Matrices

Steppan (1997a,b) has made a compelling case for comparison of *P*-matrices. The present results reinforce some of his suggestions. Steppan argues that the *P*-matrix comparison is a valid enterprise for three reasons. First, studies of phenotypic variation should give insights into the evolution of genetic variation. Because *P*-matrices can be estimated more readily than *G*-matrices, dense phylogenetic sampling is feasible. Such sampling may reveal patterns of matrix evolution that apply to the evolution of *G*-matrices. A common pattern has been detected in three comparative studies of *P*-matrices: modest divergence among taxa within terminal clades, but little or no higher-level patterning (Riska 1985; Goodin and Johnson 1992; Steppan 1997a). (The present results vote against one of the explanations that have put forward for this pattern: variation in *E*-matrices.) A similar pattern has been detected in the two studies that achieved relatively dense sampling of *G*-matrices (Podalsky et al. 1997; Pfrender 1998). Thus, the phenotypic comparisons lend credence to what may be an emerging empirical generalization about the evolution of covariance matrices: decoupling of matrix divergence from divergence in mean. The pattern might not have been appreciated if we relied solely on *G*-matrix comparisons.

Second, the evolution of *P*-matrices is an important issue in its own right, regardless of whether patterns of *P*-matrix evolution turn out to mirror patterns of *G*-matrix evolution. The methodological transition from univariate to multivariate analyses of geographic variation was made relatively recently (Gould and Johnston 1972). It seems likely that new kinds of phenomena will be uncovered as comparisons proceed from the first to the second moments of phenotypic distributions.

Finally, the *P*-matrix plays a separate role from the *G*-matrix in dynamic equations for evolutionary change (Lande 1979). Thus, the *P*-matrix can be viewed as a transformation that reveals targets of phenotypic selection (Lande and Arnold 1983). The evolution of the *P*-matrix transformation can tell us about the evolution of one feature of multivariate selection. For all these reasons, the evolution of the *P*-matrix is an important issue that deserves more study.

Morphological Integration

The result that phenotypic correlations tend to be lower than genetic correlations (Tables 3, 4) has been found in many other studies (Cheverud 1982). Olson and Miller (1958) argued that developmental and functional dependence among characters would be reflected in their phenotypic correlations. In their terminology, high correlation is a sign of morphological integration. Lande (1980a) has shown that the equilibrium *G*-matrix will be a compromise between the prevailing patterns of multivariate stabilizing selection and mutation. Thus, ridges in the selection surface will enhance genetic correlation and promote morphological integration (Sokal 1978; Cheverud 1982, 1984). Because of its more direct relation to functional interactions and selection, the *G*-matrix is probably a better guide to morphological integration than the *P*-matrix. The genetic correlations we observed for our scalation traits were generally higher than genetic correlations observed among functionally interacting traits in a primate cranium (Cheverud 1982), but we have too few traits to make meaningful comparisons within and among functional sets. The best test for morphological integration would be to compare the structure of the *G*-matrix with direct measures of multivariate stabilizing selection, a goal that has eluded most workers thus far (but see Brodie 1992, 1993).

Implications for Systematics

The observation that scale counts are heritable and genetically correlated in two populations of *T. elegans* has implications for the conduct of systematics. Only a few other studies have documented heritable variation in snake scalation (Beatson 1976; Arnold 1988; Dohm and Garland 1993; King 1997). The common observation in these studies and the present one of genetic correlation between different scale count traits suggests that these traits should not be treated as independent entities in studies of evolutionary process or in reconstructions of phylogeny. Our results with *T. elegans* identify two correlation pleiades (Berg 1960), one involving VENT and SUB and the other involving ILAB and SLAB. These two sets of traits appear to be almost genetically in-

TABLE 9. Statistical comparisons of **G**-matrices and their results.

Reference	Taxa	Traits	Statistic	Result
Lofsvold 1986	mice	morphometrics	matrix correlation	r_M = 0.16–0.58
Billington et al. 1988; Shaw and Billington 1991	grasses	morphometrics/meristics	element × element[1]	equal
Kohn and Atchley 1988	murines	morphometrics	matrix correlation	r_M = 0.01
Fong 1989	amphipods	morphometrics	matrix correlation	r_M = 0.25–0.77
Wilkinson et al. 1990; Shaw et al. 1995	flies	morphometrics	maximum likelihood	equal/nonequal[2]
Spitze et al. 1991	cladocerans	life history	matrix correlation	equal
Platenkamp and Shaw 1992	grasses	growth	maximum likelihood	equal
Brodie 1993	snakes	behavior/color	element × element[3]	equal
Carr and Fenster 1994	flowers	morphometrics	matrix regression	r_M = 0.66–0.92
Paulson 1996	butterflies	morphometrics	likelihood-ratio test[4]	equal
Podolsky et al. 1998	flowers	morphometric/discrete	likelihood-ratio test	equal
Pfrender 1998	cladocerans	life history	Flury hierarchy	CPC/nonequal[5]
Roff and Mousseau 1999	crickets	morphometrics	element × element	proportional
Roff et al. 1999	crickets	behavior (calling)	element × element, regression	proportional/nonequal[6]
This study	snakes	meristics	Flury hierarchy	CPC

[1] Main diagonal only.
[2] Result depends on which pair of populations is compared.
[3] All elements.
[4] Element × element tests reveal many differences.
[5] CPC for comparisons within clades, nonequal for comparisons between clades.
[6] Equal for comparisons within populations, proportional for comparisons between species, nonequal for comparisons with hybrids.

dependent of one another and also appear to be genetically independent of MID and POST. Thus, in a systematics study one could approach the ideal of genetically independent traits by using the sum of VENT and SUB as one trait and the sum (or average of ILAB and SLAB) as a second trait.

Other Statistical Comparisons of G-Matrices

It is hard to escape the conclusion that *G* can keep its structure, at least on a microevolutionary time scale, when one surveys the existing comparative results (Roff 1997; Roff et al. 1999; Table 9). That is not to say that *G*-matrices have been found equal in all comparisons. Most recent studies employing multivariate tests have found equal matrices or common principal components. The striking differences in matrices documented in some of the earliest studies may reflect the fact that they involved comparisons between genera (e.g., Kohn and Atchley 1988), while recent studies have compared congeneric species (Carr and Fenster 1994; Paulson 1996; Roff and Mouseau 1999; Roff et al. 1999) or conspecific populations (Billington et al. 1988; Fong 1989; Wilkinson et al. 1990; Spitze et al. 1991; Platenkamp and Shaw 1992; Brodie 1993; Podolsky et al. 1997; Pfrender 1998; Roff et al. 1999). One can argue that the constancy of *G* in the comparative studies of natural populations is an illusion produced by the lack of power in data and tests (Shaw et al. 1995). However, this argument should not send us in relentless pursuit of larger samples and more powerful tests. We expect natural populations to differ in virtually all of their statistical parameters, and with sufficient sample sizes even trivial differences can be shown to be statistically significant. The issue is not whether *G*-matrices differ—because they are bound to—but rather how they differ and whether the differences matter for evolutionary inferences. One conspicuous new issue raised by our results is the question of whether the *G*-matrix differences in other studies reflect mainly differences in eigenvalues, as in *T. elegans*, or whether they also involve eigenstructure. We need more theoretical and empirical investigations into the evolutionary forces that influence genetic covariance structure.

ACKNOWLEDGMENTS

The work reported here could not have been done without the untiring devotion of J. Gladstone to her laboratory work. L. Houck, C. R. Peterson, D. Kephart, and many other snake campers helped with the fieldwork. We thank J. Cheverud, R. Lande, M. Pfrender, R. Shaw, and G. Wagner for helpful discussions and two anonymous reviewers for comments on the manuscript. This work was supported by National Science Foundation grants BSR-8111489, BSR-8918581, and BSR-9119588 to SJA and BIR-9612469 and DBI-9722921 to PCP.

LITERATURE CITED

Alexander, A. A., and C. Gans. 1966. The pattern of dermal-vertebral correlation in snakes and amphisbaenians. Zool. Mededelingen. Leiden 41:171–190.

Arnold, S. J. 1981a. The microevolution of feeding behavior. Pp. 409–453 *in* A. Kamil and T. Sargent, eds. Foraging behavior: ecological, ethological and psychological approaches. Garland Press, New York.

———. 1981b. Behavioral variation in natural populations. I. Phenotypic, genetic and environmental correlations between chemoreceptive responses to prey in the garter snake, *Thamnophis elegans*. Evolution 35:489–509.

———. 1981c. Behavioral variation in natural populations. II. The inheritance of feeding response in crosses between geographic races of the garter snake, *Thamnophis elegans*. Evolution 35: 510–515.

———. 1988. Quantitative genetics and selection in natural populations: microevolution of vertebral numbers in the garter snake *Thamnophis elegans*. Pp. 619–636 *in* B. S. Weir, E. J. Eisen, M. M. Goodman, and G. Namkoong, eds. Proceedings of the second international conference on quantitative genetics. Sinauer, Sunderland, MA.

———. 1992a. Constraints on phenotypic evolution. Am. Nat. 140: S85–S107.

———. 1992b. Behavioural variation in natural populations. VI.

Prey responses by two species of garter snakes in three areas of sympatry. Anim. Behav. 44:705–719.
Barry, F. E., P. J. Weatherhead, and D. P. Phillipp. 1992. Multiple paternity in a wild population of northern water snakes, *Nerodia sipedon*. Behav. Ecol. Sociobiol. 30:193–199.
Beatson, R. R. 1976. Environmental and genetic correlates of disruptive coloration in the water snake, *Natrix s. sipedon*. Evolution 30:241–252.
Berg, R. 1960. The ecological significance of correlation pleiades. Evolution 14:171–180.
Billington, H. L., A. M. Mortimer, and T. McNeilly. 1988. Divergence and genetic structure in adjacent grass populations. I. Quantitative genetics. Evolution 42:1267–1277.
Brodie, E. D., III. 1992. Correlational selection for color pattern and antipredator behavior in the garter snake *Thamnophis ordinoides*. Evolution 46:1284–1298.
———. 1993. Homogeneity of the genetic variance-covariance matrix for antipredator traits in two natural populations of the garter snake *Thamnophis ordinoides*. Evolution 47:844–854.
Brown, W. M. 1983. Evolution of mitochondrial DNA. Pp. 62–88 *in* M. Nei and R. K. Koehn, eds. Evolution of genes and proteins. Sinauer, Sunderland, MA.
Carr, D. E., and C. B. Fenster. 1994. Levels of genetic variation and covariation for *Mimulus* (Scrophulariaceae) floral traits. Heredity 72:606–618.
Charlesworth, B. 1990. Optimization models, quantitative genetics, and mutation. Evolution 44:520–538.
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.
———. 1995. Morphological integration in the saddle-back tamarin (*Sanguinus fuscicollis*) cranium. Am. Nat. 145:63–89.
Cowley, D. E., and W. R. Atchley. 1988. Quantitative genetics of *Drosophila melanogaster*. II. Heritabilities and genetic correlations between the sexes for head and thorax traits. Genetics 119: 421–433.
———. 1992. Comparison of quantitative genetic parameters. Evolution 46:1965–1967.
Dohm, M. R., and T. J. Garland Jr. 1993. Quantitative genetics of scale counts in the garter snake *Thamnophis sirtalis*. Copeia 1993:987–1002.
Dowling, H. G. 1951. A proposed standard system of counting ventrals in snakes. Brit. J. Herpetol. 1:97–99.
Efron, B. 1982. The jackknife, the bootstrap and other resampling plans. Society for Industrial and Applied Mathematics, Philadelphia, PA.
Efron, B., and R. J. Tibshirani. 1993. An introduction to the bootstrap. Chapman and Hall, New York.
Falconer, D. S., and T. F. C. Mackay. 1996. Introduction to quantitative genetics. 4th ed. Longman, Harlow, UK.
Fisher, R. A. 1930. The genetical theory of natural selection. Oxford Univ. Press, Oxford.
Flury, B. 1988. Common principal components and related multivariate models. Wiley, New York.
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.
Gould, S. J., and R. F. Johnston. 1972. Geographic variation. Annu. Rev. Ecol. Syst. 3:457–498.
Hillis, D. M., B. K. Mable, and C. Moritz. 1996. Applications of molecular systematics: the state of the field and a look to the future. Pp. 515–543 *in* D. M. Hillis, C. Moritz, and B. K. Mable, eds. Molecular sytematics. 2d ed. Sinauer, Sunderland, MA.
Houle, D. 1991. Genetic covariance of fitness correlates: what genetic correlations are made of and why it matters. Evolution 45:630–648.
Kelley, K. C., S. J. Arnold, and J. Gladstone. 1997. The effects of substrate and vertebral number on locomotion in the garter snake *Thamnophis elegans*. Funct. Ecol. 11:189–198.
King, R. B. 1997. Variation in brown snake (*Storeria dekayi*) morphology and scalation: sex, family, and microgeographic differences. J. Herpetol. 31:335–346.
Kohn, L. A. P., and W. R. Atchley. 1988. How similar are genetic correlations structures? Data from mice and rats. Evolution 42: 467–481.
Lande, R. 1979. Quantitative genetic analysis of multivariate evolution, applied to brain:body size allometry. Evolution 33:402–416.
———. 1980a. The genetic covariance between characters maintained by pleiotropic mutations. Genetics 94:203–215.
———. 1980b. Genetic variation and phenotypic evolution during allopatric speciation. Am. Nat. 116:463–479.
———. 1980c. Sexual dimorphism, sexual selection, and adaptation in polygenic characters. Evolution 34:292–305.
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. Test of the homogeneity of genetic covariance structure among species and subspecies. Evolution 40:559–573.
———. 1988. Quantitative genetics of morphological differentiation in *Peromyscus*. II. Analysis of selection and drift. Evolution 42:54–67.
Luiselli, L. 1995. The mating strategy of the European adder, *Vipera berus*. Acta Oecol. 16:375–388.
Lynch, M., and B. Walsh. 1998. Genetics and analysis of quantitative traits. Sinauer, Sunderland, MA.
Maynard Smith, J., R. Burian, S. Kauffman, P. Albrech, J. Campbell, B. Goodwin, R. Lande, D. Raup, and L. Wolpert. 1985. Developmental constraints and evolution. Q. Rev. Biol. 60:265–287.
Olson, E., and R. Miller. 1958. Morphological integration. Univ. of Chicago Press, Chicago.
Paulsen, S. M. 1996. Quantitative genetics of wing color pattern in the buckeye butterfly (*Precis coenia* and *Precis evarete*): evidence against the constancy of *G*. Evolution 50:1585–1597.
Peters, J. A. 1964. Dictionary of herpetology. Hafner, New York.
Pfrender, M. E. 1998. Evolutionary dynamics of molecular and quantitative genetic variation in ephemeral pond populations of *Daphnia pulex*. Ph.D. diss. University of Oregon, Eugene, OR.
Phillips, P. C. 1998a. H2boot: bootstrap estimates and tests of quantitative genetic data. Univ. of Texas at Arlington. Software available at www.uta.edu/biology/phillips/software.
———. 1998b. CPCrand: randomization test of the CPC hierarchy. Univ. of Texas at Arlington. Software available at www.uta.edu/biology/phillips/software.
———. 1998c. CPC: common principal components analysis. Univ. of Texas at Arlington. Software available at www.uta.edu/biology/phillips/software.
Phillips, P. C., and S. J. Arnold. 1999. Hierarchical comparison of genetic variance-covariance matrices. I. Using the Flury hierarchy. Evolution 53:1506–1515.
Platenkamp, G. A. J., and R. G. Shaw. 1992. Constraints on adaptive population differentiation in *Anthoxanthum odoratum*. Evolution 46:341–352.
Podolsky, R. H., R. G. Shaw, and F. H. Shaw. 1997. Population structure of morphological traits in *Clarkia dudleyana*. II. Constancy of within-population genetic variance. Evolution 51: 1785–1796.
Rice, S. H. 1998. The evolution of canalization and the breaking of von Baer's laws: modeling the evolution of development with epistasis. Evolution 52:647–656.
Riska, B. 1985. Group size factors and geographic variation of morphometric correlation. Evolution 39:792–803.
———. 1986. Some models for development, growth, and morphometric correlation. Evolution 40:1303–1311.
Roff, D. A. 1995. The estimation of genetic correlations from phenotypic corrrelations: a test of Cheverud's conjecture. Heredity 74:481–490.

———. 1996. The evolution of genetic correlations: an analysis of patterns. Evolution 50:1392–1403.
———. 1997. Evolutionary quantitative genetics. Chapman and Hall, New York.
Roff, D. A., and T. A. Mousseau. 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.
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.
Ruthven, A. G. 1908. Variation and genetic relationships of the garter snakes. Bulletin of the U.S. Natural History Museum, no. 61.
Schluter, D. 1996. Adaptive radiation along genetic lines of least resistance. Evolution 50:1766–1774.
Schmalhausen, I. I. 1949. Factors of evolution, the theory of stabilizing selection. Univ. of Chicago Press, Chicago.
Schwartz, J. M., G. F. McCraGcken, and G. M. Burghardt. 1989. Multiple paternity in wild populations of the garter snake *Thamnophis sirtalis*. Behav. Ecol. Sociobiol. 25:269–274.
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.
Shaw, R. G. 1987. Maximum-likelihood approaches applied to quantitative genetics of natural populations. Evolution 41:812–826.
———. 1991. The comparison of quantitative genetic parameters between populations. Evolution 45:143–151.
———. 1992. Comparison of quantitative genetic parameters: reply to Cowley and Atchley. Evolution 46:1967–1969.
Shaw, R. G., and H. L. Billington. 1991. Comparison of variance components between two populations of *Holcus lanatus*: a reanalysis. Evolution 45:1287–1289.
Slatkin, M. 1987. Quantitative genetics of heterochrony. Evolution 41:799–811.
Sokal, R. R. 1978. Population differentiation: something new or more of the same? Pp. 215–239 *in* P. F. Brussard, ed. Ecological genetics: the interface. Springer, Berlin.
Spitze, K., J. Burnson, and M. Lynch. 1991. The covariance structure of life-history characters in *Daphnia pulex*. Evolution 45: 1081–1090.
Steppan, S. J. 1997a. Phylogenetic analysis of covariance structure. I. Contrasting results from matrix correlations and common principal component analyses. Evolution 51:571–586.
———. 1997b. Phylogenetic analysis of covariance structure. II. Reconstructing matrix evolution. Evolution 51:587–594.
Turelli, M. 1988. Phenotypic evolution, constant covariances, and the maintenance of additive genetic variance. Evolution 42: 1342–1347.
Via, S. 1984. The quantitative genetics of polyphagy in an insect herbivore. I. Genotype-environment interaction in larval performance on different host plant species. Evolution 38:881–895.
Voris, H. K. 1975. Dermal scale-vertebra relationships in sea snakes (Hydrophidae). Copeia 1975:746–755.
Waddington, C. H. 1957. The strategy of the genes. Allen and Unwin, London.
Wagner, G. P., G. Booth, and H. Bagheri-Chaichian. 1997. A population genetic theory of canalization. Evolution 51:329–347.
Wilkinson, G. S., K. Fowler, and L. Partridge. 1990. Resistance of genetic correlation structure to directional selection in *Drosophila melanogaster*. Evolution 44:1990–2003.
Willis, J. H., J. A. Coyne, and M. Kirkpatrick. 1991. Can one predict the evolution of quantitative characters without genetics? Evolution 45:441–444.

Corresponding Editor: L. Leamy