

Geographic variation in the **G** matrices of wild populations of the barn swallow

DA Roff¹, T Mousseau², AP Møller³, F de Lope⁴ and N Saino⁵

¹Department of Biology, University of California, Riverside, CA 92521, USA; ²Department of Biological Sciences, University of South Carolina, Columbia, SC 29208, USA; ³Laboratoire de Parasitologie Evolutive, CNRS UMR 7103, Université Pierre et Marie Curie, Bâtiment A, 7ème étage, 7 quai St Bernard, Case 237, FR-75252 Paris Cedex 05, France; ⁴Departamento de Biología Animal, Universidad de Extremadura, E-06071 Badajoz, Spain; ⁵Dipartimento di Scienze dell'Ambiente e del Territorio, Università di Milano-Bicocca, via Emanueli 15, I-20126 Milano, Italy

In this paper, we present an analysis of genetic variation in three wild populations of the barn swallow, *Hirundo rustica*. We estimated the **P**, **E**, and **G** matrices for six linear morphological measurements and tested for variation among populations using the Flury hierarchical method and the jackknife followed by MANOVA method. Because of nonpositive-definite matrices, we had to employ 'bending' to analyse the **G** and **E** matrices with the Flury method. Both statistical methods agree in finding

that the **P** and **G** matrices are significantly different but comparison between the analysis of the **P** matrices and pairwise analyses of the **P**, **E**, and **G** matrices suggests caution in interpreting the Flury results concerning differences in matrix structure. The significant variation among the populations in the **G** matrices appears to be due in large measure to the most geographically distant population.

Heredity (2004) 93, 8–14. doi:10.1038/sj.hdy.6800404

Keywords: **G** matrix; natural selection; genetic variance–covariance matrices; clinal variation

Introduction

Additive genetic variances and covariances play an important role in evolutionary change, but we are still surprisingly ignorant of how these genetic components themselves vary in natural populations. Whereas, we have abundant evidence that genotypic and phenotypic means vary in response to selection, we have little empirical information on the extent to which additive genetic variances and covariances are moulded by natural selection and/or genetic drift. Evolutionary change in quantitative traits can be described by the multivariate extension of the breeder's 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). The matrix combination $\mathbf{G}\mathbf{P}^{-1}$ can be viewed as the multivariate 'equivalent' of h^2 , with **G** replacing the additive genetic variance in the numerator and \mathbf{P}^{-1} replacing the phenotypic variance in the denominator. Note that although the nonadditive genetic variance components (dominance, epistasis) do not enter into the **G** matrix, they do impact on evolutionary change by their inclusion in the phenotypic variance.

Without further qualification, as stated, the multivariate evolutionary response equation assumes that the genetic and phenotypic matrices remain 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, 1980). 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). Thus, 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).

A number of studies have estimated and compared the **G** matrices from different populations of the same organism. In the cases of *Thamnophis ordinoides* (Brodie, 1993) and *Holcus lanatus* (Shaw and Billington, 1991), *Allonemobius fasciatus* and *A. socius* (Roff and Mousseau, 1999; Roff *et al.*, 1999) no differences were observed. Significant differences were found in populations of *Gammarus minus* (Fong, 1989; Jernigan *et al.*, 1994; Roff, 2002), *Thamnophis elegans* (Arnold and Phillips, 1999) and *Drosophila melanogaster* (van 'T Land *et al.*, 1999). With one exception, in all of the preceding cases and those involving higher taxa the offspring generation and typically the parental generation were raised in the laboratory. Arnold and Phillips (1999) brought in gravid females from the wild and estimated genetic (co)variances using the regression of mean offspring value of dam value. In this case it might reasonably be argued that the estimated **G** matrix is equivalent to that obtained for a wild population.

So far as we are aware, the Arnold and Phillips analysis is the only study of **G** matrix variation in free-ranging populations of animals, although several studies have estimated genetic components of individual wild populations. Estimates of morphological heritabilities for traits in free-ranging individuals are very similar to and not significantly different from those obtained under laboratory conditions ($\text{lab} = 0.50$, $\text{field} = 0.56$, see Table 1 in Weigensberg and Roff, 1996). Thus, we might expect that laboratory estimates of **G** also to be similar to those estimated from natural populations. Nevertheless, it is still preferable to estimate the **G** matrix within the environment in which the organism has evolved (Orengo and Prevosti, 1999; Blanckenhorn, 2002). On the other hand, this approach can potentially confound $\text{G} \times \text{E}$ interactions with genetic variation *per se*. Thus, in principle, an analysis would include a common garden experiment in which the **G** matrices were estimated under the same conditions. Such an experiment is not possible with the organism studied herein and hence conclusions must be tempered by considerations of environment on genetic expression. In the present paper, we present an analysis of variation in the phenotypic, environmental, and genetic (**P**, **E**, and **G**) matrices of morphological traits in three wild populations of the barn swallow, *Hirundo rustica*.

Estimates of the matrix elements were made using offspring on female parent regression and the matrices then compared using two statistical approaches, the Flury approach (Phillips and Arnold, 1999) and the jackknife followed by MANOVA method (Roff, 2002). The first method examines the structure of the matrices (equality, proportionality, common principal components, etc), whereas the second tests for equality and correlations between the matrices and other variables such as sex, latitude of origin, etc.

While the **G** matrix is the focus of this paper, we present an analysis of the **P** matrix for two reasons: First, there is considerably less error in the estimation of the **P** matrix compared to the **G** matrix and hence the **P** matrix will be a much more sensitive indicator of variation, although it cannot itself tell us the source of the variation. Significant variation in the **P** matrix would imply variation in either or both the **G** and **E** matrices. The second reason for examining the **P** matrix is that, because the **G** matrix is part of the elements of the **P** matrix ($\text{P} = \text{G} + \text{E}$), it may be possible to use phenotypic (co)variances as surrogate measures of the genetic (co)variances (Cheverud, 1988; Roff, 1995, 1996; Koots and Gibson, 1996). Thus, we can ask two questions: 'Are the two matrices highly correlated?' and 'Do we get the same statistical answer using the **P** matrix as the **G** matrix?'

Materials and methods

Study sites

We studied barn swallows at Badajoz (38°50' N, 659' W), Southern Spain (since 1976), Milano (45°28' N, 9°10' E), Northern Italy (since 1993), and Kraghede (57°12' N, 10°00' E), Denmark (since 1970), as part of a long-term project. The study site at Badajoz consists of open farmland with pastures, cereals, and fruit plantations, and most barn swallows breed in barns. The study site at

Milano consists of open farmland with pastures, cereals, and hedges. Barn swallows are mainly restricted to dairy farms, where they breed inside buildings. The study site at Kraghede consists of open farmland with pastures, cereals, potatoes, and rape with mixed plantations, hedges, and ponds. As at the other two sites, barn swallows breed within barns and other buildings.

General field procedures

Barn swallows were captured at least weekly from the arrival in early spring until the end of the breeding season using mist nets at all entrances to the barns with breeding birds. Mark-recapture analyses of the data have revealed that the capture probability of birds exceeds 98% in Spain and Denmark (F de Lope, AP Møller and T Szép, unpublished manuscripts), with a presumably equally high capture probability in Milano (based on the presence of birds without individually numbered aluminium rings) (N Saino, unpublished data).

Measurements

Seven linear traits were measured: (1) beak length, BL, (2) beak width, BW, (3) beak height, BH, (4) mean wing length, W, (5) tarsus length, T, (6) mean outermost tail length, TL, (7) length of the shortest central tail feathers, TS. Not all traits were measured on all birds: in particular, beak width was often omitted and for this reason, we deleted this measurement from the analysis. Sample sizes for measurements on all six remaining traits for both female parent and offspring were 91 for Spain, 51 for Denmark, and 69 for Italy. In addition, there were five Spanish and 25 Danish families for which partial measurements were available.

Statistical methods

The genetic variance-covariance matrices were estimated using offspring on parent regression. Extra-pair paternity is common in barn swallows (Møller and Tegelström, 1997) but brood parasitism is not (none was found in the Italian sample, N Saino, unpublished data). The possibility that the male was not the father of all offspring precludes in principle the use of either offspring on midparent or offspring on father regressions. Estimates from mean offspring on mother are also confounded by the uncertain relationship between offspring (full or half siblings). Therefore, we took the most conservative approach and used the regression of a single male offspring on the mother. Cross-fostering experiments with song sparrows, blue tits, willow tits, collared flycatchers, pied flycatchers, starlings, and tree swallows found no evidence of maternal effects on offspring morphological traits (reviewed in Weigensberg and Roff, 1996). Møller (1994, pp 152–156), using a series of indirect measures of maternal effects, found in barn swallows no effect of the mother on tail length of her offspring. While we have no direct estimates from cross-fostering in the present experiments, the unanimous findings cited above indicate that the offspring on female parent approach can be expected to provide valid estimates.

A number of techniques for comparison of **G** matrices have been proposed (Roff, 2000; Steppan *et al*, 2002). The most widely used technique in the last few years is that known as the Flury model (Phillips and Arnold, 1999). This method uses a stepwise approach to examine the

matrices for equality, proportionality, common principal components, or variation in one or more of the principal components. To increase the robustness of the analysis, Phillips incorporated a randomization test rather than using the likelihood ratio directly. We, therefore, report the probability as ' P_{rand} '. Several problems arose when we implemented the Flury method: first, the method cannot readily accommodate missing data and second, because of nonpositive-definite matrices, we had to employ the 'bending' option in the Phillips program CPCrand (available at <<http://www.uta.edu/biology/philipps/software.html>>). 'Bending' makes the matrix positive definite by adjusting the eigenvalues of the matrix just enough to eliminate the negative eigenvalues (Hayes and Hill, 1981). Unfortunately, bending changes the error structure of the comparison matrix and hence may invalidate the statistical test (Phillips and Arnold, 1999).

An alternate approach is the 'jackknife followed by MANOVA' method of Roff (2002), hereafter referred to simply as the MANOVA method. The Flury and MANOVA methods bear strong statistical kinship and in the one case in which they have been compared reach the same statistical conclusion (Roff, 2002). An advantage of the MANOVA method for the present analysis is that it permits the incorporation of those offspring-parent combinations for which not all measurements were taken (see below) and does not suffer the problem arising from nonpositive-definite matrices. On the other hand, it does not provide the analysis of principal component structure that the Flury method does. We used both methods, the MANOVA method providing a check on the reliability of the initial comparison of equality given by the Flury method.

The 'jackknife followed by MANOVA' method is described in detail in Roff (2002). Briefly, the method is as follows. Taking each country separately, first the **G** matrix was estimated using the single offspring on female parent, \mathbf{G}_{obs} . Next, the first family was deleted and the **G** matrix estimated using the reduced data set, \mathbf{G}_{-1} . The first pseudo-value **G** matrix, ϕ_{-1} , was calculated as $\phi_{-1} = n\mathbf{G}_{\text{obs}} - (n-1)\mathbf{G}_{-1}$. The deleted family was added back to the data set, the next family deleted and the procedure repeated to obtain, ϕ_{-2} . The jackknife mean estimate of **G** is given by $\phi = \sum_{i=1}^n \phi_{-i} / n$. More importantly for the present analysis the set of pseudo-values have the same distribution as **G** and hence these may be used in a MANOVA analysis with population as an independent variable (Roff, 2002). Multivariate techniques such as MANOVA may be sensitive to outliers (Tabachnick and Fidell, 2001), and so we additionally estimated the probability for this test using the following randomization procedure. To understand the process it is necessary to understand the data layout: the data set consists of $n_1 + n_2 + n_3$ rows, where n_i is the number of families in the i th population and each row is the set of (co)variance pseudo-values for a single deletion. The last column in the data set contains the country designator. After initial calculation of Wilks' λ , which we denote as λ_{obs} , we randomized the country designations and recalculated Wilks' λ , which we denote as λ_{rand} . This procedure was repeated 999 times and the probability of obtaining a Wilks' λ as small or smaller than observed (note that in contrast to most statistical measures variance accounted for increases with a decrease in

Wilks' λ) estimated as $(1+N)/1000$, where N is the number of cases in which $\lambda_{\text{rand}} < \lambda_{\text{obs}}$ (Manly, 1997). We designate this estimate of the probability as P_λ : in no case did the two probability values differ significantly, confirming that our data satisfied the assumptions of MANOVA.

Because of the varying number of families per trait in some offspring-parent combinations the total number of pseudo-values differed among the variances and covariances. We retained all families in the initial estimation of the pseudo-values as this minimizes the standard error, but we reduced this set of pseudo-values such that the set used in the MANOVA had equal numbers per (co)variance (MANOVA requires that there be no missing variables). We achieved this by random deletion of pseudo-value sets (ie random ϕ_{-i}). We also analysed the data using only that set used in the Flury analysis. In general, both sets of MANOVA analyses gave the same result and for ease of comparison between the MANOVA and Flury method, unless otherwise indicated, we present the results using the same data set (ie complete measures per family).

Because the **P** matrix is guaranteed to be positive definite, the analysis of the phenotypic matrix does not present the same problem as the **G** matrix. For assurance, we used both methods of analysis. Similarly, we also analysed the environmental covariance matrix, which, like the **G** matrix also required bending for the Flury method.

Results

Variation in phenotypic means

There is highly significant multivariate variation among countries in phenotypic means (Wilks' $\lambda = 0.421$, $F_{14,442} = 17.088$, $P < 0.0001$). Previous work has established that morphological trait means increase with latitude (Møller, 1994). The present data also show this pattern with a general size increase northwards (Figure 1), but there are too few data points to statistically test for clinal variation. Variation in trait means does not imply variation in trait variances or covariances, but it does suggest the action of natural selection, genetic drift, and/or environmental effects, all of which could affect the **G** matrix (Roff, 1997, 2000).

Variation in the **P**, **E**, and **G** matrices

There is highly significant variation among countries in the phenotypic (co)variances ($P_{\text{rand}} = 0.012$), but the matrices do not differ with respect to proportionality ($P_{\text{rand}} = 0.06$, Table 1). This result suggests that the variation among countries might be a scaling phenomenon. The typical relationship between means and variances is allometric, which can be removed by a logarithmic transformation. Such a transformation makes no qualitative difference to the statistical conclusion (for equality $P_{\text{rand}} = 0.014$, for proportionality $P_{\text{rand}} = 0.061$), indicating that the variation is not due to allometric scaling. The MANOVA analysis also indicates highly significant variation among populations (Wilks' $\lambda = 0.633$, $F_{42,376} = 2.299$, $P < 0.001$, $P_\lambda < 0.001$).

The environmental matrices differ with respect to equality ($P_{\text{rand}} < 0.0001$), proportionality ($P_{\text{rand}} < 0.0001$), common principal components ($P_{\text{rand}} = 0.0167$), but not

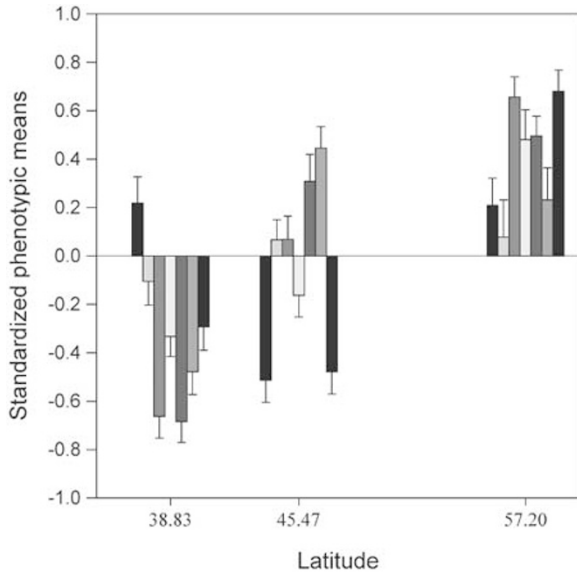


Figure 1 Variation (mean ± SE) in morphological traits in three populations of barn swallows as a function of latitude. Traits are displayed sequentially in the order as presented in the methods. For visual purposes, each trait was standardized to a mean of zero and a SD 1 using the combined sample (ie all three countries). Statistical analysis was done using the original data.

Table 1 Results of the Flury hierarchical test on the P, E, and G matrices of three barn swallow populations

Test	P_{rand}		
	P	E	G
Equality	0.012	<0.001	<0.001
Proportionality	0.060	<0.001	0.007
Full CPC	0.235	0.016	<0.001
CPC(4) ^a	0.234	0.067	<0.001
CPC(3)	0.348	0.042	<0.001
CPC(2)	0.667	0.006	<0.001
CPC(1)	0.830	0.519	0.007

^aCPC(1): a partial common principal component model with 1 components in common.

with respect to partial principal components (Table 1). Using the log-transformed data made little difference except that the hierarchy now becomes significant at CPC(4), the partial common principal component model with four components shared in common. However, the MANOVA method did not show a significant difference among countries (Wilks' $\lambda = 0.758$, $F_{42,376} = 1.329$, $P = 0.090$, $P_{\lambda} = 0.073$).

There is considerable variation in the estimated components of the G matrices (Table 2). This variation is reflected in significant statistical differences in the G matrices (Table 1). Both the Flury and MANOVA methods agree in finding significant variation with respect to equality (Flury, $P_{rand} = <0.001$, MANOVA, Wilks' $\lambda = 0.732$, $F_{42,376} = 1.511$, $P = 0.025$, $P_{\lambda} = 0.022$ for complete data set Wilks' $\lambda = 0.718$, $F_{42,376} = 1.613$, $P = 0.012$, $P_{\lambda} = 0.009$). The Flury method suggests that there are no shared principal components (Table 1). Analysis of the log-transformed data show statistically significant differences with respect to equality ($P_{rand} < 0.001$)

Table 2 Genetic variances and covariances of six morphological traits in three wild populations of barn swallow

Trait(s) ^a	Denmark		Italy		Spain	
	(Co)variance	SE	(Co)variance	SE	(Co)variance	SE
BL	221.44	238.27	342.13	173.66	214.89	288.72
BL, BH	425.32 ^b	155.97	22.06	59.87	52.05	112.04
BL, W	7.01	12.35	-3.24	10.78	19.02	12.40
BL, T	331.73	301.88	-13.39	155.25	-77.98	174.51
BL, TL	8.86	27.80	36.18	21.56	-27.67	21.78
BL, TS	2.30	8.73	6.89	4.45	1.72	5.86
BH	495.04	169.15	10.95	71.26	-15.49	64.94
BH, W	18.44	9.10	1.07	5.13	7.70	5.67
BH, T	279.40	210.29	96.53	63.35	84.09	73.91
BH, TL	10.23	20.13	9.61	13.50	13.25	10.28
BH, TS	12.78	6.16	-1.42	2.58	8.01	3.48
W	2.65	0.94	2.61	0.96	2.28	0.83
W, T	25.72	16.63	2.85	9.69	0.29	9.23
W, TL	1.11	1.38	2.29	1.89	4.04	1.24
W, TS	0.05	0.52	0.33	0.36	1.27	0.33
T	472.79	561.60	397.50	206.09	310.57	191.90
T, TL	-4.65	41.90	13.47	18.04	-12.88	18.67
T, TS	20.26	11.07	2.79	4.35	3.69	5.32
TL	4.31	4.48	10.24	4.67	8.95	3.35
TL, TS	1.20	1.19	0.82	0.84	1.82	0.85
TS	0.14	0.48	0.30	0.23	0.70	0.29

Standard errors estimated by jackknife.

BL, beak length; BH, beak height; W, mean wing length; T, tarsus length; TL, mean outermost tail length; TS, length of the shortest central tail feathers.

^aBeak length, beak height, and tarsus length were measured in mm × 100, other three traits measured in mm.

^bEstimates in italics significantly different from zero.

Table 3 Summary of Flury hierarchical analysis using pairwise contrasts

Matrix	Denmark	Denmark	Spain
	vs Spain	vs Italy	vs Italy
	Proportional	Proportional	Equal
P			
E	CPC(1)	CPC	CPC
G	CPC(1)	CPC	CPC
Probabilities associated with equality of matrices			
P Flury	0.023	0.040	0.114
P MANOVA ^a	0.003, 0.006	0.001, 0.001	0.088, 0.035
	0.003, 0.003	0.001, 0.001	0.075, 0.022
G Flury	<0.001	<0.001	<0.001
G MANOVA ^a	0.053, 0.006	0.103, 0.062	0.290, 0.259
	0.055, 0.008	0.095, 0.047	0.281, 0.275

^aTop line: First probability is that using the restricted data set (the same as used for the Flury method), while the second shows that using the complete data set. Bottom line gives equivalent probabilities estimated using the randomization method.

and proportionality ($P_{rand} < 0.001$) but not with respect to common principal components ($P_{rand} = 0.072$).

Table 3 summarizes pairwise analyses of the three matrices and Figure 2 shows the standardized (co)variances as a function of latitude. The two latitudinally closest pair of countries, Spain and Italy, do not statistically differ in their P matrices, but the other two comparisons show a statistically significant proportional difference, as in the full analysis. Both the E and G matrices differ in equality and proportionality in all

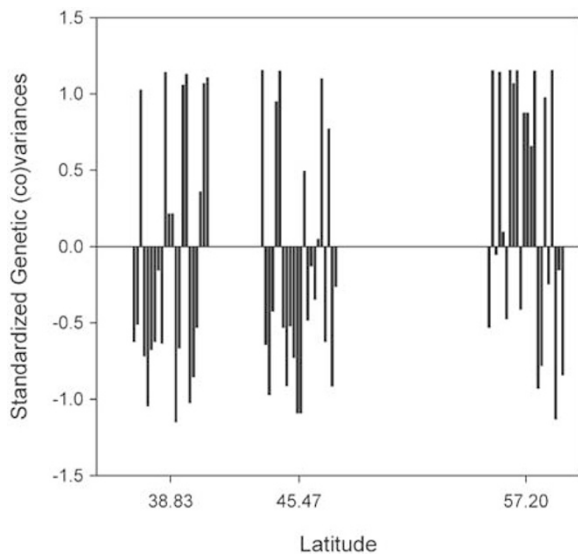


Figure 2 Standardized genetic (co)variances of six morphological traits in three populations of barn swallow. For visual purposes each trait (co)variance was standardized to a mean of zero and a SD of 1 using the combined sample (ie all three countries). Statistical analysis was done using the original data.

comparisons. The MANOVA results for the pairwise matrix comparisons are not the same as the Flury results (Table 3). The MANOVA analysis of the **P** matrices give all pairwise comparisons to be significant, although the Spain *vs* Italy comparison is marginal (nonsignificant for the restricted data set). The MANOVA analysis of the **G** matrices gives results most similar to those of the Flury analysis of the **P** matrices (Table 3): pairwise comparison of countries reveals a significant difference between the geographically most separate swallow populations of Denmark and Spain, but there is no significant difference between Denmark and Italy or Spain and Italy. Similar discrepancies exist for the **E** matrix, with the Flury method giving highly significant pairwise differences ($P < 0.001$ in all cases), whereas the MANOVA analysis of the same data set gives none to be significant ($P = 0.077$, $P_\lambda = 0.073$ for Denmark *vs* Spain, $P = 0.368$, $P_\lambda = 0.374$ for Denmark *vs* Italy, and $P = 0.375$, $P_\lambda = 0.355$ for Spain *vs* Italy).

The phenotypic (co)variances are made up of the genetic and environmental (co)variances. Morphological traits typically have high heritabilities and genetic correlations (Mousseau and Roff, 1987; Roff, 1996). Consequently, we would expect that the phenotypic and genetic (co)variances will also be highly correlated. Comparison both within and between countries indicates that the **P** and **G** matrices are indeed highly correlated (Table 4). Because of the problem of part-whole correlation the statistical tests associated with these correlations are unreliable, but they are sufficiently high to indicate statistical significance. More importantly, the high correlation between the **P** and **G** matrix elements suggests that the former might be used as a surrogate for the latter, at least as a first approximation.

Discussion

Selection on morphological components has been demonstrated to occur in a wide range of organisms

Table 4 Pearson product moment correlations (below diagonal, probability^a above the diagonal) between the **G** and **P** matrices of morphological traits in the barn swallow

	<i>G</i> matrix for			<i>P</i> matrix for		
	Denmark	Spain	Italy	Denmark	Spain	Italy
G matrices						
Denmark		0.033	0.020	0.001	0.011	0.003
Spain	0.468		<0.0001	<0.0001	<0.0001	<0.0001
Italy	0.502	0.957		<0.0001	<0.0001	<0.0001
P matrices						
Denmark	0.688	0.889	0.903		<0.0001	<0.0001
Spain	0.540	0.866	0.933	0.862		<0.0001
Italy	0.621	0.894	0.960	0.950	0.953	

^aProbabilities only approximate, because correlation is part-whole

(Endler, 1986; Fairbairn and Reeve, 2001; Kingsolver *et al*, 2001) and in various species of birds (Holland and Yalden, 1995; Brown and Brown, 1998; Larsson *et al*, 1998; Merila *et al*, 1999; Balmford *et al*, 2000; Barbraud, 2000; Nowakowski, 2000; Przybylo *et al*, 2000), including barn swallows (Møller, 1993; Møller and Tegelstrom, 1997; Møller *et al*, 1998; Brown and Brown, 1999). In Nebraska, a period of severe weather resulted in size-related mortality of barn swallows (Brown and Brown, 1999) indicating that, as with other bird species natural selection acts upon at least some morphological components. Of particular interest would be an analysis of the pattern of genetic variances and covariances in relation to estimates of selection on the traits. The present data are not suitable for such an analysis. However, a quantitative genetic analysis of morphometric traits in the water strider *Aquarius remigis* suggests that patterns of selection may be reflected in patterns of variation in the **G** matrix (Preziosi and Roff, 1998). Thus, a fruitful avenue of empirical investigation is the relationship between the structure of the **G** matrix and the structure of selection (Arnold *et al*, 2001).

Using data on three geographically distinct populations of barn swallow, we have shown using the Flury method that the populations differ in mean phenotype and in their phenotypic, environmental, and genetic variance-covariances matrices. These results differ from those obtained by Arnold and Phillips (1999) for two populations of garter snakes: in their analysis, Arnold and Phillips observed equality of the environmental matrices and differences in the **P** and **G** matrices. Further, Arnold and Phillips found that, in general, the results from the phenotypic analysis were the same as those from the genetic analysis. In contrast, in the present study using the Flury method, we observed significant variation in the environmental matrices and obtained different conclusions from the phenotypic and genetic analyses (Table 1). There is generally far more statistical power in the analysis of phenotypic than genetic variation and hence the observation of only proportional differences in the **P** matrices whereas the **G** matrices differed in their principal components suggests that in this case, the two matrices are quite different. However, the correlations between the elements of the two matrices are high (Table 4) as typically found for morphological traits (Roff, 1996). It was necessary to apply a bending coefficient to the **G** and **E** matrices to carry out the Flury hierarchical test: it is possible that this may produce

spurious results. However, both the Flury method and the MANOVA method indicate significant differences among the **G** matrices and the latter method does not require the adjustment used in the Flury method. Further, the randomization procedure applied to the MANOVA method provided confirmation of the estimated probabilities. We recommend use of this randomization procedure as a general approach to the MANOVA analysis. Given the correspondence between the Flury and MANOVA methods, the inequality among the **G** matrices is a robust result, and agrees with the **P** matrix analysis. Because of the possible problem associated with 'bending' differences found by the Flury method at a deeper level of structure should be regarded with caution. Similarly, the significant differences found by the Flury method for the **E** matrix but not for the MANOVA method may be a result of bending the **E** matrices for the former analysis. The general findings of the MANOVA method in the present analysis do agree with the qualitative findings of Arnold and Phillips (1999), that is, differences in the **P** and **G** matrices but not the **E** matrices. More work on the statistical properties of the several methods of matrix analysis are required.

While the present analysis does indicate overall variation in the **G** matrices, much of this variation appears to come from differences between the population of Denmark *vs* the other two (Figure 2, Table 3). Because the birds were free ranging, we cannot say if this variation is a consequence of genotype by environment interaction or differences in the **G** matrices that would be found under a common garden rearing regime. Regardless of the source of variation, the differences do indicate that evolutionary trajectories in the field populations will be different under the same selection regime. We do not suggest that the present results demonstrate clinal variation – there are too few populations and the geographic distribution insufficiently great – but note that it is consistent with the observed clinal variation in mean morphological trait values and hence the hypothesis of clinal variation is worth investigating further.

References

- Arnold SJ (1992). Constraints on phenotypic evolution. *Am Nat* **140**(Suppl): s85–s107.
- Arnold SJ, Pfrender ME, Jones AG (2001). The adaptive landscape as a conceptual bridge between micro- and macroevolution. *Genetica (Dordrecht)* **112–113**: 9–32.
- Arnold SJ, Phillips PC (1999). Hierarchical comparison of genetics variance–covariance matrices. II. Coastal–inland divergence in the garter snake *Thamnophis elegans*. *Evolution* **53**: 1516–1527.
- Balmford A, Lewis MJ, Brooke MDL, Thomas ALR, Johnson CN (2000). Experimental analyses of sexual and natural selection on short tails in a polygynous warbler. *Proc Roy Soc Biol Sci Series B* **267**: 1121–1128.
- Barbraud C (2000). Natural selection on body size traits in a long-lived bird, the snow petrel *Pagodroma nivea*. *J Evol Biol* **13**: 81–88.
- Blanckenhorn WU (2002). The consistency of quantitative genetic estimates in field and laboratory in the yellow dung fly. *Genetica* **114**: 171–182.
- Brodie EDI (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 CR, Brown MB (1998). Intense natural selection on body size and wing and tail asymmetry in cliff swallows during severe weather. *Evolution* **52**: 1461–1475.
- Brown CR, Brown MB (1999). Natural selection on tail and bill morphology in barn swallows *Hirundo rustica* during severe weather. *Ibis* **141**: 652–659.
- Cheverud JM (1988). A comparison of genetic and phenotypic correlations. *Evolution* **42**: 958–968.
- Endler JA (1986). *Natural Selection in the Wild*. Princeton University Press: Princeton, NJ.
- Fairbairn DJF, Reeve J (2001). Natural selection. In: Fox CW, Roff DA, Fairbairn DJ (eds) *Evolutionary Ecology: Concepts and Case Studies*. Oxford University Press: Oxford, pp 29–43.
- Fong DW (1989). Morphological evolution of the amphipod *Gammarus minus* in caves: quantitative genetic analysis. *Am Midl Nat* **121**: 361–378.
- Hayes JF, Hill WG (1981). Modification of estimates of parameters in the construction of genetic selection indices ('Bending'). *Biometrics* **37**: 483–493.
- Holland PK, Yalden DW (1995). Who lives and who dies? The impact of severe April weather on breeding common sandpipers *Actitis hypoleucos*. *Ringing Migration* **16**: 121–123.
- Jernigan RW, Culver DC, Fong DW (1994). The dual role of selection and evolutionary history as reflected in genetic correlations. *Evolution* **48**: 587–596.
- Kingsolver JG, Hoekstra HE, Hoekstra JM, Berrigan D, Vignieri SN, Hill CE *et al* (2001). The strength of phenotypic selection in natural populations. *Am Nat* **157**: 245–261.
- Koots KR, Gibson JP (1996). Realized sampling variances of estimates of genetic parameters and the difference between genetic and phenotypic correlations. *Genetics* **143**: 1409–1416.
- 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.
- Larsson K, Der Jeugd HP, Der Veen IT, Forslund P (1998). Body size declines despite positive directional selection on heritable size traits in a barnacle goose population. *Evolution* **52**: 169–1184.
- Lynch M (1990). The rate of morphological evolution in mammals from the standpoint of the neutral expectation. *Am Nat* **136**: 727–741.
- Manly BFJ (1997). *Randomization, Bootstrap and Monte Carlo Methods in Biology*. Chapman & Hall: New York.
- Merila J, Przybylo R, Sheldon BC (1999). Genetic variation and natural selection on blue tit body condition in different environments. *Genet Res* **73**: 165–176.
- Møller AP (1993). Sexual selection in the barn swallow *Hirundo rustica*: III. Female Tail Ornaments *Evol* **47**: 417–431.
- Møller AP (1994). *Sexual Selection and the Barn Swallow*. Oxford University Press: Oxford, UK.
- Møller AP, Barbosa A, Cuervo J, De Lope F, Merino S, Saino N (1998). Sexual selection and tail streamers in the barn swallow. *Proc Roy Soc Lond Ser B Biol Sci* **265**: 409–414.
- Møller AP, Tegelstrom H (1997). Extra-pair paternity and tail ornamentation in the barn swallow *Hirundo rustica*. *Behav Ecol Sociobiol* **41**: 353–360.
- Mousseau TA, Roff DA (1987). Natural selection and the heritability of fitness components. *Heredity* **59**: 181–198.
- Nowakowski JJ (2000). Long-term variability of wing length in a population of the Reed warbler *Acrocephalus scirpaceus*. *Acta Ornith (Warsaw)* **35**: 173–182.
- Orengo DJ, Prevosti A (1999). Wing-size heritability in a natural population of *Drosophila subobscura*. *Heredity* **82**: 100–106.
- Phillips PC, Arnold SJ (1999). Hierarchical comparison of genetic variance–covariance matrices. I. Using the Flury hierarchy. *Evolution* **53**: 1506–1515.

- Preziosi RF, Roff DA (1998). Evidence of genetic isolation between sexually monomorphic and sexually dimorphic traits in the water strider *Aquarius remigis*. *Heredity* **81**: 92–99.
- Przybylo R, Sheldon BC, Merila J (2000). Patterns of natural selection on morphology of male and female collared flycatchers (*Ficedula albicollis*). *Biol J Linn Soc* **69**: 213–232.
- Roff DA (1995). The estimation of genetic correlations from phenotypic correlations: a test of Cheverud's conjecture. *Heredity* **74**: 481–490.
- Roff DA (1996). The evolution of genetic correlations: an analysis of patterns. *Evolution* **50**: 1392–1403.
- Roff DA (1997). *Evolutionary Quantitative Genetics*. Chapman & Hall: New York.
- Roff DA (2000). The evolution of the G matrix: selection or drift? *Heredity* **84**: 135–142.
- Roff DA (2002). Comparing Gmatrices: a MANOVA method. *Evolution* **56**: 1286–1291.
- Roff DA, Mousseau TA (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 DA, Mousseau TA, Howard DJ (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.
- Shaw RG, Billington HL (1991). Comparison of variance components between two populations of *Holcus lanatus*: a reanalysis. *Evolution* **45**: 1287–1289.
- Steppan SJ, Phillips PC, Houle D (2002). Comparative quantitative genetics: evolution of the Gmatrix. *TREE* **17**: 320–327.
- Tabachnick BG, Fidell LS (2001). *Using Multivariate Statistics*. Allyn and Bacon: Boston.
- Turelli M (1988). Phenotypic evolution, constant covariance, and the maintenance of additive variance. *Evolution* **42**: 1342–1347.
- van 'T Land J, van Putten P, Zwaan BJ, Kamping A, van Delden W (1999). Latitudinal variation in wild populations of *Drosophila melanogaster*: heritabilities and reaction norms. *J Evol Biol* **12**: 222–232.
- Weigensberg I, Roff DA (1996). Natural heritabilities: can they be reliably estimated in the laboratory. *Evolution* **50**: 2149–2157.