Evolution, 57(1), 2003, pp. 191–195

QUANTITATIVE GENETICS OF GEOMETRIC SHAPE: HERITABILITY AND THE PITFALLS OF THE UNIVARIATE APPROACH

CHRISTIAN PETER KLINGENBERG
Department of Biology, University of Konstanz, 78457 Konstanz, Germany

Key words.—Breeders' equation, geometric morphometrics, heritability, quantitative genetics, selection, shape, size.

Received June 25, 2002. Accepted September 4, 2002.

There is considerable interest in the evolution of morphological traits, and morphometric studies in combination with the multivariate theory of quantitative genetics can provide a detailed understanding of the variation and evolutionary potential of these traits. For both morphometrics and quantitative genetics, new and improved techniques have been established recently (e.g., Dryden and Mardia 1998; Lynch and Walsh 1998). The combination of these two approaches makes it possible to study genetic variation with explicit reference to the geometry of the structure under investigation and to interpret the results in their anatomical context. Examples include studies using classical quantitative genetic designs (e.g., Arnqvist and Thornhill 1998; Currie et al. 2000; Klingenberg and Leamy 2001) and analyses of quantitative trait loci (Zimmerman et al. 2000; Klingenberg et al. 2001; Workman et al. 2002).

In a recent paper, Monteiro et al. (2002) proposed a univariate estimate of heritability for shape based on Procrustes distance, a measure of the extent of difference between pairs of landmark configurations. The method extracts a univariate heritability estimate from the inherently multidimensional shape data by assuming the model of isotropic variation (Goodall 1991), which presumes that there is an equal amount of nondirectional variation at each landmark and that the landmarks are independent of one another. Monteiro et al. (2002, pp. 565, 569) suggest that this univariate heritability estimate can be used to assess whether the relative amount of genetic versus phenotypic variation differs among populations in space and time, and to examine whether nonexisting shapes should be explained by selection or developmental constraints. They illustrate their method with a case study of shape variation in honeybee wings.

The assumptions of the isotropic model, on which the method of Monteiro et al. (2002) is based, are often unrealistic—even in the authors' own dataset. Moreover, these assumptions also have strong implications for the comparison of shape variation among populations and for the study of genetic constraints. Another difficulty is that the experimental design of the case study of Monteiro et al. (2002) is not large enough to characterize the genetic variation in all dimensions of shape space, and is therefore not sufficient to assess the method.

This comment provides a more explicit explanation of the heritability estimate, based on the multivariate theory of quantitative genetics (Lande 1979; Cheverud 1984), to highlight the assumptions inherent in the model used by Monteiro et al. (2002) and to demonstrate the implications that these assumptions have for the usefulness of the heritability estimate. A previous study (Klingenberg and Leamy 2001), which was also based on the combination of geometric morphometrics and quantitative genetics, but which used a fully multivariate approach, serves as a basis for comparison.

Quantitative Genetics of Shape and the Concept of Heritability

Shape is an inherently multivariate concept. Even relatively simple shapes such as configurations of only a few landmarks can vary in many ways by relative shifts of the landmarks against one another. In geometric morphometrics, each shape is represented by a point in a multidimensional shape tangent space (Rohlf 1999; Monteiro et al. 2000), which is amenable to analysis by the methods of multivariate statistics. The multivariate equivalent to the univariate breeders' equation for predicting the response to selection $\Delta\mu = h^2 s$ (e.g., Falconer and Mackay 1996) is the equation $\Delta\boldsymbol{\mu} = \mathbf{G}\mathbf{P}^{-1}\mathbf{s}$ (Lande 1979), where $\Delta\mu$ and $\Delta\boldsymbol{\mu}$ are the univariate and multivariate response to selection, h^2 is the heritability, $\mathbf{G}$ is the additive genetic covariance matrix, $\mathbf{P}$ the phenotypic covariance matrix, and s and $\mathbf{s}$ are the univariate and multivariate selection differentials, respectively. Whereas $\Delta\mu$ and s can be fully characterized by their magnitude, their multivariate equivalents have both a magnitude and direction. Moreover, the vectors $\mathbf{s}$ and $\Delta\boldsymbol{\mu}$ will not have the same direction unless $\mathbf{s}$ is an eigenvector of $\mathbf{G}\mathbf{P}^{-1}$ (for further details, see Klingenberg and Leamy 2001). Therefore, both the magnitudes and directions of $\mathbf{s}$ and $\Delta\boldsymbol{\mu}$ need to be taken into account.

The difference between two shapes can be quantified by their Procrustes distance (e.g., Dryden and Mardia 1998). It is important to note, however, that Procrustes distance only measures the magnitude of shape differences, but ignores their direction, because it does not consider which

[1] Present address: School of Biological Sciences, University of Manchester, 3.164 Stopford Building, Oxford Road, Manchester M13 9PT, United Kingdom; E-mail: cpk@man.ac.uk.

landmarks shift against each other or in which anatomical directions those landmarks move. In practice, it is convenient to use the Euclidean distance in shape tangent space, which is a very close approximation of Procrustes distance for the small amounts of shape variation typically found in intraspecific studies. The lengths of the vectors $\mathbf{s}$ and $\Delta\boldsymbol{\mu}$ can also be expressed in units of Procrustes distance in this manner, e.g., $\|\mathbf{s}\| = (\mathbf{s}^T\mathbf{s})^{0.5}$, where the superscript ''T'' denotes the vector transpose. Because $\mathbf{s}$ and $\Delta\boldsymbol{\mu}$ each have a specific direction in shape space, the lengths of these vectors cannot fully characterize the selection differential or the response to selection.

Monteiro et al. (2002) use Procrustes distance as the basis for their univariate method for assessing genetic and phenotypic variation in shape, and therefore ignore the directionality in shape tangent space. The justification for this is the isotropic model of Goodall (1991), which assumes that there is the same amount of variation around the average shape at each landmark, that the variation at each landmark is equal in all directions, and that the variation is independent among landmarks. This model implies that there is an equal amount of variation in each direction of the shape tangent space. Therefore, the genetic and phenotypic covariance matrices can be written as $\mathbf{G} = \sigma_A^2\mathbf{I}$ and $\mathbf{P} = \sigma_P^2\mathbf{I}$, where $\mathbf{I}$ is an identity matrix and σ_A^2 and σ_P^2 are scalar constants corresponding to the additive genetic and phenotypic variances that are equal for all dimensions. In this case, the ratio $h^2 = \sigma_A^2/\sigma_P^2$ can be defined as the shape heritability as proposed by Monteiro et al. It follows that $\Delta\boldsymbol{\mu} = h^2\mathbf{s}$, and therefore also $\|\Delta\boldsymbol{\mu}\| = h^2\|\mathbf{s}\|$ (Monteiro et al. 2002, p. 569). In this model, all landmark coordinates or dimensions of shape space are equivalent, and it is possible simply to sum up variation to estimate h^2 as a scalar quantity. There is no need to keep track of the directions of variation because the isotropic model specifically assumes that there is no directionality.

The isotropic model is not the only situation where the use of Procrustes distance would be justified. The univariate approach is also applicable if the additive genetic and phenotypic covariance matrices of the coordinates in shape tangent space are proportional, that is, if $\mathbf{G} = c\mathbf{P}$ for some constant c ($0 < c < 1$; the isotropic model is a special case of this situation). In this case, the heritability can be defined as $h^2 = c$ for all shape variables, and would also be correctly estimated with the procedures outlined by Monteiro et al. But even though phenotypic and genetic covariance matrices often are related to some degree (e.g., Cheverud 1988; Roff 1997), exact proportionality must be considered the exception.

Apart from these special cases, however, the ratio $\|\Delta\boldsymbol{\mu}\|/\|\mathbf{s}\|$ will differ according to the direction of $\mathbf{s}$ in shape tangent space (e.g., Klingenberg and Leamy 2001, Fig. 6). Therefore, contrary to the recommendation of Monteiro et al. (2002, p. 569), the univariate breeders' equation can normally not be used to predict the magnitude of the selection response for shape. This means that selection for each scalar shape variable will, in general, have a different heritability. Therefore, no single value can represent the overall heritability of shape as a whole with the same predictive utility that heritability has in the univariate context.

The Case Study of Bee Wings

Monteiro et al. (2002) illustrate their method with a case study on honeybee wings, which estimates the $\mathbf{G}$ matrix from the variation among the means of the 21 bee colonies, each represented by 10 bees (it is assumed here than nonadditive genetic and common-environment effects are negligible, as the authors argue on p. 567).

Test of the model assumptions

I first examine the assumption of isotropic shape variation for that dataset. Monteiro et al. (2002, p. 568) report that the first relative warp of the $\mathbf{P}$ matrix, that is, the first principal component (PC) in shape tangent space, takes up 16.5% of the total shape variation. I simulated the corresponding isotropic distribution in tangent space, generating 10,000 random datasets, each consisting of 210 observations with spherical variation in 36 dimensions (multivariate normal distribution with means zero for all variables and with an identity matrix as the covariance matrix). On average, the first principal component accounted for 5.3% of the total variance, and in none of these simulation runs for as much as 16.5%. Therefore, the dataset shows a highly significant deviation of the $\mathbf{P}$ matrix from the isotropic model because variation is overly concentrated in some dimensions of shape space. This is also consistent with the scatter of Procrustes-superimposed configurations around the overall consensus (Monteiro et al. 2002, Fig. 1B), where some landmarks appear more variable than others, and some show clearly directional scatter. The misfit of the isotropic model in this case is not unusual, because patterns of shape variation encountered in morphometric studies are typically nonisotropic. These deviations from isotropy reflect directional variation and interrelationships among landmarks, and thus may be of biological interest on their own (Klingenberg and McIntyre 1998; Badyaev and Foresman 2000; Debat et al. 2000; Klingenberg and Zaklan 2000; Klingenberg 2002).

It is more difficult to assess whether the $\mathbf{P}$ and $\mathbf{G}$ matrices may be proportional. Monteiro et al. (2002) found that these matrices show significant similarities with a matrix correlation of 0.57, but this value is also sufficiently far from 1.0 to indicate it is unlikely that the matrices are exactly proportional. Therefore, the conditions for using the univariate approach based on Procrustes distance appear not to be fulfilled.

Sample size and dimensionality

The dataset also illustrates a further pitfall. Monteiro et al. (2002) analyzed configurations of 20 landmarks, for which the shape tangent space has 36 dimensions. They estimated genetic variation from the variation among the shape averages of 21 bee colonies, yielding 20 degrees of freedom among the means. The variation among colonies therefore defines a 20-dimensional subspace at most, and the estimated $\mathbf{G}$ matrix can therefore only span just over half of the dimensionality of the shape tangent space. In particular, it will be impossible to use such a dataset to

assess whether there are genetic constraints; that is, to determine whether dimensions without associated variation are truly invariant or lack variation merely due to inadequate sampling. The quantitative genetic analysis of shape is inherently a very ambitious project, and accordingly, such studies require sufficiently large experimental designs.

A Different Example: Mouse Mandibles

It is instructive to contrast the approach of Monteiro et al. (2002) with a fully multivariate approach using the results of a quantitative genetic study of geometric shape variation in the mouse mandible, where genetic parameters were estimated from an unbalanced parent-offspring design with restricted maximum-likelihood methods (Klingenberg and Leamy 2001). The total sample consists of 1241 specimens and considers 11 landmarks, so that the shape tangent space has 18 dimensions, of which only 15 could be considered due to software limitations. The first principal component of the **P** matrix takes up 17.8% of the total shape variation. In simulations of spherical variation for this dimensionality and sample size, the PC1 accounts for 8.0% of the variation on average. This 2.2-fold excess of the observed value is somewhat less than the corresponding 3.1-fold excess for the bee wing data, so that the deviation from the isotropic model at least does not appear greater for the mouse mandibles than for the bee wings. Moreover, the matrix correlation between the **P** and **G** matrices was 0.88 (Klingenberg and Leamy 2001, p. 2346) for the mouse mandibles as compared to 0.57 for the bee wings (Monteiro et al. 2002, p. 569). These matrices do not seem any less proportional for the mouse mandibles than for the bee wings.

The univariate heritability for shape of the mouse mandibles, following the approach of Monteiro et al., can be estimated as the ratio of the total variance of the **G** matrix (2.35×10^{-4}, in units of squared Procrustes distance) to that of the **P** matrix (8.05×10^{-4}) and amounts to 0.29. In reality, however, the heritabilities of possible shape variables range from zero to 0.73 (Klingenberg and Leamy 2001, Fig. 7). The univariate estimate averages over the dimensions of shape tangent space and so suggests a moderate heritability. It thereby obscures the fact that shape variables with high heritability exist side by side with others that show strong genetic constraints. This example illustrates clearly that the application of the univariate breeders' equation to Procrustes distance (Monteiro et al. 2002, p. 569) can over- or underestimate the magnitude of the response to selection by several times.

Suggested Applications of the Heritability Estimate

As possible applications for the univariate estimate of shape heritability, Monteiro et al. (2002, pp. 565, 569) specifically suggest the detection of constraints, comparisons of genetic variation in different populations, and rate tests for the reconstruction of past selection. In all these contexts, however, there are significant difficulties with the method because it ignores the directionality of variation.

For determining the potential and constraints for the evolution of shape, the direction of variation in shape space is critical, because genetic constraints are identified as directions in shape space for which there is no genetic variation (e.g., Cheverud 1984; Maynard Smith et al. 1985; Kirkpatrick and Lofsvold 1992; Schluter 1996). The only ''constraint'' that can be accommodated by the model of isotropic shape variation is the complete lack of any genetic variation. Otherwise, the isotropic model explicitly assumes the absence of constraints, because genetic and phenotypic variations are distributed equally over all dimensions of shape space. In most cases where constraints exist, they will not be due to the complete lack of any genetic variation, but to the absence of variation in specific aspects of shape. It is the very nature of the isotropic model that it practically excludes constraints a priori, and therefore empirical studies searching for constraints should not assume that model. The procedure of averaging over the dimensions of shape space is not an effective way to test for constraints, because it would systematically conceal real constraints in some shape dimensions with variation from other dimensions.

While the suggested approach of Monteiro et al. (2002, pp. 565, 569) to base population comparisons in space and time on a scalar measure of variation would be convenient, unfortunately there is no scalar measure that can characterize multivariate variation sufficiently. Just as shape averages in geometric morphometrics are compared among populations by using multivariate methods (e.g., Monteiro and Abe 1999; Duarte et al. 2000), comparisons of variation within populations should also consider the multidimensional nature of shape. Methods for characterizing and comparing phenotypic or genetic covariance matrices have been established (e.g., Cheverud 1988; Kohn and Atchley 1988; Shaw 1991; Phillips and Arnold 1999; Roff et al. 1999) and can be used for geometric morphometric data (e.g., Klingenberg and McIntyre 1998; Debat et al. 2000; Klingenberg and Leamy 2001).

Retrospective analyses of phenotypic evolution, such as rate tests for distinguishing selection from random drift (Turelli et al. 1988; Spicer 1993), are also inherently multivariate problems when they are to be applied to multidimensional features such as shape (Lande 1979). They are in many ways analogous to discriminant analysis (e.g., discussion in Marroig and Cheverud 2001), and therefore must take into account the covariance structure of the shape variables. The recommendation of Monteiro et al. (2002, p. 569) to use their heritability estimate for shape in such tests is therefore justified only when the isotropic model of shape variation applies. All the suggested uses of the shape heritability estimate therefore make strong assumptions that are rarely fulfilled in empirical data.

Conclusions

Approaches based on Procrustes distance have a useful role in morphometrics because they provide summary statistics about the amount of shape variation. For instance, in multivariate regression analyses, the relative amounts of explained and residual variation can be assessed (Monteiro 1999), or the approach can provide a measure of the relative

magnitude of individual variation, fluctuating asymmetry, and measurement error (Klingenberg and McIntyre 1998). In such situations, the specific patterns of variation are not the focus of the analysis, and Procrustes distance measures are informative as summary statistics even in cases where the isotropic model does not hold. Even there, however, methods based on Procrustes distance may have less statistical power than fully multivariate methods that consider all the available shape information (e.g., Klingenberg et al. 2002).

For quantitative genetic studies of shape, however, a fully multivariate approach is quintessential. Heritability, a univariate concept, has not played a role in the multivariate theory of quantitative genetics; if one insists on having a multivariate equivalent for heritability, then it is the matrix $\mathbf{GP}^{-1}$ (Roff 2000; Klingenberg and Leamy 2001). Numerous morphometric studies have shown that shape variation is usually localized and richly structured. The very strength of geometric morphometrics is that the analyses can account explicitly for the spatial heterogeneity that is associated with the anatomy and the ontogenetic origins of biological structures. Although it is mathematically possible to compute a ''global'' heritability estimate by averaging across all dimensions of shape space, such an overall measure must ignore the spatial structure of variation, and is therefore fundamentally at odds with the goals of geometric morphometrics.

Acknowledgments

I am grateful to L. J. Leamy, I. Somorjai, and two anonymous reviewers for helpful comments on earlier versions of the manuscript.

Literature Cited

Arnqvist, G., and R. Thornhill. 1998. Evolution of animal genitalia: patterns of phenotypic and genotypic variation and condition dependence of genital and non-genital morphology in water strider (Heteroptera: Gerridae: Insecta). Genet. Res. 71: 193–212.

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

Cheverud, J. M. 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.

Currie, A. J., S. Ganeshanandam, D. A. Noiton, D. Garrick, C. J. A. Shelbourne, and N. Oraguzie. 2000. Quantitative evaluation of apple (*Malus* × *domestica* Borkh.) fruit shape by principal component analysis of Fourier descriptors. Euphytica 111: 219–227.

Debat, V., P. Alibert, P. David, E. Paradis, and J.-C. Auffray. 2000. Independence between developmental stability and canalization in the skull of the house mouse. Proc. R. Soc. Lond. B 267: 423–430.

Dryden, I. L., and K. V. Mardia. 1998. Statistical analysis of shape. Wiley, Chichester, U.K.

Duarte, L. C., L. R. Monteiro, F. J. Von Zuben, and S. F. Dos Reis. 2000. Variation in mandible shape in *Trichomys apereoides* (Mammalia: Rodentia): geometric analysis of a complex morphological structure. Syst. Biol. 49:563–578.

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

Goodall, C. R. 1991. Procrustes methods in the statistical analysis of shape. J. R. Stat. Soc. B 53:285–339.

Kirkpatrick, M., and D. Lofsvold. 1992. Measuring selection and constraint in the evolution of growth. Evolution 46:954–971.

Klingenberg, C. P. 2002. Developmental instability as a research tool: using patterns of fluctuating asymmetry to infer the developmental origins of morphological integration. Pp. 427–442 *in* M. Polak, ed. Developmental instability: causes and consequences. Oxford Univ. Press, New York.

Klingenberg, C. P., and L. J. Leamy. 2001. Quantitative genetics of geometric shape in the mouse mandible. Evolution 55: 2342–2352.

Klingenberg, C. P., and G. S. McIntyre. 1998. Geometric morphometrics of developmental instability: analyzing patterns of fluctuating asymmetry with Procrustes methods. Evolution 52: 1363–1375.

Klingenberg, C. P., and S. D. Zaklan. 2000. Morphological integration between developmental compartments in the *Drosophila* wing. Evolution 54:1273–1285.

Klingenberg, C. P., L. J. Leamy, E. J. Routman, and J. M. Cheverud. 2001. Genetic architecture of mandible shape in mice: effects of quantitative trait loci analyzed by geometric morphometrics. Genetics 157:785–802.

Klingenberg, C. P., M. Barluenga, and A. Meyer. 2002. Shape analysis of symmetric structures: quantifying variation among individuals and asymmetry. Evolution 56:1909–1920.

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

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

Lynch, M., and B. Walsh. 1998. Genetics and analysis of quantitative traits. Sinauer, Sunderland, MA.

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

Maynard Smith, J., R. Burian, S. Kauffman, P. Alberch, J. Campbell, B. Goodwin, R. Lande, D. Raup, and L. Wolpert. 1985. Developmental constraints and evolution. Q. Rev. Biol. 60: 265–287.

Monteiro, L. R. 1999. Multivariate regression models and geometric morphometrics: the search for causal factors in the analysis of shape. Syst. Biol. 48:192–199.

Monteiro, L. R., and A. S. Abe. 1999. Functional and historical determinants of shape in the scapula of Xenarthran mammals: evolution of a complex morphological structure. J. Morphol. 241:251–263.

Monteiro, L. R., B. Bordin, and S. F. dos Reis. 2000. Shape distances, shape spaces and the comparison of morphometric methods. Trends Ecol. Evol. 15:217–220.

Monteiro, L. R., J. A. F. Diniz-Filho, S. F. dos Reis, and E. D. Araújo. 2002. Geometric estimates of heritability in biological shape. Evolution 56:563–572.

Phillips, P. C., and S. J. Arnold. 1999. Hierarchical comparison of genetic variance-covariance matrices. I. Using the Flury hierarchy. Evolution 53:1506–1515.

Roff, D. A. 1997. Evolutionary quantitative genetics. Chapman and Hall, New York.

———. 2000. The evolution of the **G** matrix: selection or drift? Heredity 84:135–142.

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.

Rohlf, F. J. 1999. Shape statistics: Procrustes superimpositions and tangent spaces. J. Classif. 16:197–223.

Schluter, D. 1996. Adaptive radiation along genetic lines of least resistance. Evolution 50:1766–1774.

Shaw, R. G. 1991. The comparison of quantitative genetic parameters between populations. Evolution 45:143–151.
Spicer, G. S. 1993. Morphological evolution of the *Drosophila virilis* species group as assessed by rate tests for natural selection on quantitative characters. Evolution 47:1240–1254.
Turelli, M., J. H. Gillespie, and R. Lande. 1988. Rate tests for selection on quantitative characters during macroevolution and microevolution. Evolution 42:1085–1089.
Workman, M. S., L. J. Leamy, E. J. Routman, and J. M. Cheverud. 2002. Analysis of quantitative trait locus effects on the size and shape of mandibular molars in mice. Genetics 160: 1573–1586.
Zimmerman, E., A. Palsson, and G. Gibson. 2000. Quantitative trait loci affecting components of wing shape in *Drosophila melanogaster*. Genetics 155:671–683.

Corresponding Editor: M. Zelditch