Evolution, 52(5), 1998, pp. 1363–1375

GEOMETRIC MORPHOMETRICS OF DEVELOPMENTAL INSTABILITY: ANALYZING PATTERNS OF FLUCTUATING ASYMMETRY WITH PROCRUSTES METHODS

CHRISTIAN PETER KLINGENBERG[1,2,3] AND GRANT S. MCINTYRE[1,4]
[1]*Department of Biological Sciences, University of Alberta, Edmonton, Alberta T6G 2E9, Canada*
[2]*Department of Zoology, Duke University, Durham, North Carolina 27708-0325*
[3]*E-mail: cpk@acpub.duke.edu*
[4]*E-mail: gmcintyr@gpu.srv.ualberta.ca*

Abstract.—Although fluctuating asymmetry has become popular as a measure of developmental instability, few studies have examined its developmental basis. We propose an approach to investigate the role of development for morphological asymmetry by means of morphometric methods. Our approach combines geometric morphometrics with the two-way ANOVA customary for conventional analyses of fluctuating asymmetry and can discover localized features of shape variation by examining the patterns of covariance among landmarks. This approach extends the notion of form used in studies of fluctuating asymmetry from collections of distances between morphological landmarks to an explicitly geometric concept of shape characterized by the configuration of landmarks. We demonstrate this approach with a study of asymmetry in the wings of tsetse flies (*Glossina palpalis gambiensis*). The analysis revealed significant fluctuating and directional asymmetry for shape as well as ample shape variation among individuals and between the offspring of young and old females. The morphological landmarks differed markedly in their degree of variability, but multivariate patterns of landmark covariation identified by principal component analysis were generally similar between fluctuating asymmetry (within-individual variability) and variation among individuals. Therefore, there is no evidence that special developmental processes control fluctuating asymmetry. We relate some of the morphometric patterns to processes known to be involved in the development of fly wings.

Key words.—Developmental stability, fluctuating asymmetry, geometric morphometrics, *Glossina palpalis*, Procrustes analysis, shape, tsetse flies, wing veins.

Received December 31, 1997. Accepted June 19, 1998.

Fluctuating asymmetry has been widely used as a measure of developmental instability and thus has become the focus of considerable attention (reviewed by Møller and Swaddle 1997). Despite the widespread use of fluctuating asymmetry for measuring developmental instability, its developmental origins are unclear (Markow 1995; Palmer 1996). A few studies have addressed this problem with experimental approaches (e.g., Smith and Palmer 1994; Klingenberg and Nijhout 1998), others followed growth of structures on both body sides (Chippindale and Palmer 1993; Møller 1996; Collin 1997; Swaddle and Witter 1997), whereas others have searched for quantitative trait loci (Leamy et al. 1997, 1998) or investigated specific candidate genes (Batterham et al. 1996; Davies et al. 1996) to uncover the processes underlying developmental instability.

Here we introduce a new approach that uses morphometric techniques. It is based on the fact that variation in developmental processes can generate distinctive patterns in the joint variation of multiple morphological traits affected by the processes. For instance, variation in the growth of a developmental precursor tends to generate positive correlations among all parts derived from it, whereas variable partitioning of precursor tissue tends to produce negative correlations between the resulting parts (e.g., Riska 1986). Therefore, it should be possible to derive information about developmental processes from the patterns of covariance among traits. Many studies have used morphometric techniques to infer the role of development for morphological variation from multivariate patterns of variation among individuals (e.g., Cheverud 1982a; Zelditch 1987; Cowley and Atchley 1990; Paulsen and Nijhout 1993; Paulsen 1994), but this approach has not been developed for fluctuating asymmetry (but see the early attempt by Sakai and Shimamoto 1965). Here, we examine both the variation among individuals and the within-individual asymmetry between body sides jointly.

Our analysis of asymmetry is based on the Procrustes technique, which is at the core of the linkage between geometric methods and conventional multivariate statistics that constitutes the recent "morphometric synthesis" (Bookstein 1996a). Previously, Bookstein (1991, p. 267–270) and Auffray et al. (1996) have introduced a Procrustes method for studies of asymmetry. Smith et al. (1997) revised the technique substantially, making it amenable for large sample sizes and standard software and linking it more closely to the core methods of geometric morphometrics (Bookstein 1996a). Here we extend the Procrustes approach of Smith et al. (1997) according to a two-factor ANOVA design (Leamy 1984; Palmer and Strobeck 1986; Palmer 1994) that allows us to quantify the different components of asymmetry and test them statistically. In addition, we use a multivariate approach to extract and analyze the patterns of covariation among landmarks to investigate the developmental basis of asymmetries.

We apply these methods in a case study of the wings of tsetse flies (*Glossina palpalis*). Fly wings are an excellent study system for our purpose, because they are essentially two dimensional and the wing veins provide many well-defined morphological landmarks. Moreover, the development of fly wings and their venation is fairly simple and known in considerable detail, mostly from studies of *Drosophila melanogaster* (e.g., Waddington 1940; Garcia-Bellido and de Celis 1992; Sturtevant and Bier 1995). Here we examine the claim that special processes govern developmental stability, such as a "localised, or left-right, signalling system which monitors and regulates morphogenesis" (Swaddle 1997, p.

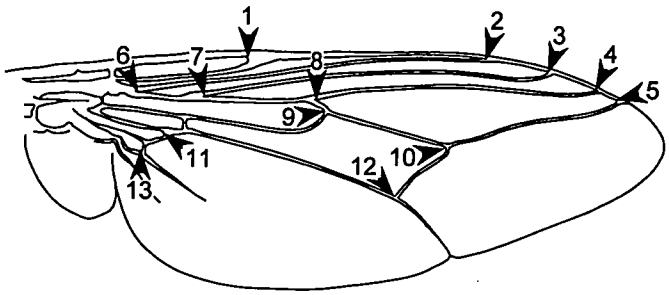

FIG. 1. The right wing of a tsetse fly. Arrowheads point to the landmarks used in this study.

59), which are distinct from the developmental processes that generate variation among individuals (see also Møller and Swaddle 1997). In this case, one would expect that fluctuating asymmetry shows patterns of covariance among landmarks that differ qualitatively from those of individual variation, whereas in the absence of special processes regulating developmental stability, fluctuating asymmetry and individual variation should have similar covariance structures. In this context, we demonstrate tests of specific hypotheses, and we relate the morphometric results to information about wing development. Finally, in the Appendix we offer some practical recommendations for users of these methods.

MATERIALS AND METHODS

Data

In this study, we analyze the wing shape of tsetse flies (*G. p. gambiensis* Vanderplank) from a line (GAMB-K) maintained in laboratory culture since 1972. Because the data used here are from a larger study of the effects of maternal age on offspring quality (McIntyre and Gooding 1998), the sample of 70 flies is composed of two subgroups, with 43 offspring of young females (< 40 d old) and 27 offspring of old females (> 65 d old). All the wings considered in this study are from male flies, and therefore sexual dimorphism does not contribute to morphological variation. To flatten the wings, they were mounted between microscope slides held together by elastic bands until the mountant (Euparal) had hardened.

The data consist of x and y coordinates of 13 morphological landmarks (Fig. 1). All landmarks are at intersections of wing veins, including landmarks 1–5, which are at the intersections of longitudinal veins with the costal vein that runs along the anterior edge of the wing. Therefore, they are easy to locate precisely and fulfill the criteria for Bookstein's (1991) "type 1" landmarks.

For each wing, one of us (GSM) digitized the coordinates of all landmarks using a dissecting microscope with a camera lucida and a Summasketch FX digitizing tablet. All wings were measured three times, which makes it possible to assess digitizing error. The flies in each maternal age group were digitized in sequence for each replicate (first replicate for all flies in a group, then second replicate, etc.). The two maternal age groups were measured in two separate sessions (first all three replicates for offspring of young flies, then for offspring of old flies). Because the wings were newly positioned under the microscope for each replicate and the raw landmark coordinates were recorded automatically in a computer file, the observer was blind with respect to the results of previous measurements.

Variation and Asymmetry of Wing Size

For most of this study, we focus on wing shape, because it can be characterized by the geometric configuration of morphological landmarks. Variation in wing size was eliminated before these analyses by scaling specimens to unit centroid size (the square root of the sum of squared distances from a set of landmarks to their centroid; Slice et al. 1996). Nevertheless, individual variation and asymmetry of wing size also should be considered in their own right, complementing the analyses of shape. Asymmetry in the overall size of a structure reflects positive correlations among left-right differences of interlandmark distances, which have been studied by other authors as the "individual asymmetry parameter" (e.g., Leamy 1993; Leamy et al. 1997).

To assess asymmetry of total wing size, we used an ANOVA with centroid size as the dependent variable, body side as a fixed effect, and individuals as a random effect (Palmer and Strobeck 1986; Palmer 1994). In this analysis, the main effect of flies stands for individual variation in size, the main effect of body sides is for directional asymmetry (one side is systematically larger than the other), and the interaction term is a measure of fluctuating asymmetry (the *variation* in left-right differences among individuals). Because the flies in our sample were from two subgroups, we used maternal age (young/old) as an additional fixed effect in the ANOVA (this factor did not take part in any interactions); the effect of individuals was nested within the maternal age classes, and therefore the among-fly mean square was used as the error term (denominator of the F-ratio) for the maternal age effect. The fly × side interaction was used as the error term to test significance of the main effects of flies and sides and the measurement error for the fly × side interaction effect.

A Procrustes Method for Quantifying Asymmetry of Shape

Procrustes methods analyze shape by superimposing configurations of landmarks in two or more specimens to achieve an overall best fit (Rohlf and Slice 1990). In studies of asymmetry, the Procrustes method proceeds in several steps (Fig. 2; Bookstein 1991; Auffray et al. 1996; Smith et al. 1997): (1) reflect the landmark configuration of one body side to its mirror image to align corresponding landmarks of both sides; (2) scale the configurations to unit centroid size; (3) superimpose the left and right configurations so that they have the same centroid (the point of mean x and y coordinates for each configuration is shifted usually to the coordinates [0, 0]); and (4) rotate the configurations against each other around their centroid to achieve an optimal fit of corresponding landmarks. Asymmetry can then be measured as the deviations between the pairs of corresponding landmarks.

Our analysis used a single Procrustes superimposition (steps 3 and 4) to align simultaneously all the landmark configurations (both wings of all specimens and their replicate measurements). A single overall consensus configuration is computed as the mean coordinates of corresponding land-

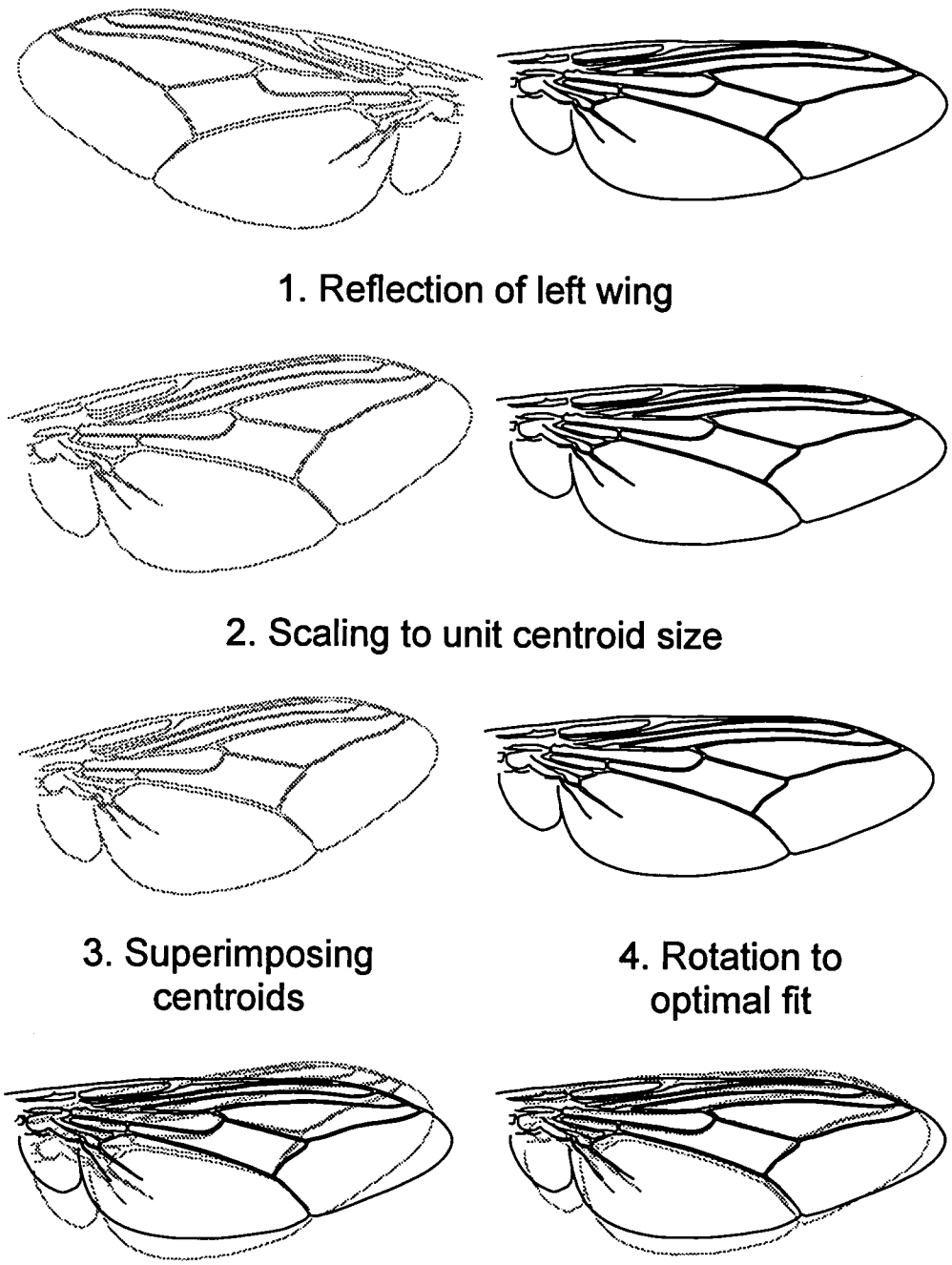

FIG. 2. Procrustes procedure for analyzing asymmetry of shape. First, landmark configurations of the one body side are reflected to mirror images (left wing in the example). Second, configurations from both left and right body sides are scaled to have the same overall size (centroid size = 1). Third, configurations are superimposed so that the centroid (center of gravity of the landmarks) of each has coordinates (0, 0). Finally, configurations are rotated around the common centroid to achieve an overall best fit between corresponding landmarks.

marks in the aligned configurations (Rohlf and Slice 1990; Bookstein 1996a). The coordinates of the aligned configurations constitute a new set of variables that contains the complete shape information. Further analyses focus on these coordinates, using standard methods of multivariate statistics. This version of the procedure, which was introduced by Smith et al. (1997), is considerably simpler to implement than separate left-right comparisons for each pair of wings, the procedure originally described by Bookstein (1991) and Auffray et al. (1996). Moreover, it relates to the methodological core of the "morphometric synthesis" because the coordinates of the aligned configurations correspond to points in a common shape space (e.g., Bookstein 1996a; Small 1996).

Details of Computation.—The Procrustes procedure of this study used the least-squares criterion to find an optimal alignment in step 4 of the procedure (generalized orthogonal least-squares fit; Rohlf and Slice 1990). The procedure iteratively minimizes the sum of the squared distances between the land-

marks of all objects in the sample and the corresponding landmarks of the consensus configuration. The square root of the sum of squared distances between corresponding landmarks of two optimally aligned configurations is an approximation of Procrustes distance (Slice et al. 1996). Procrustes distance plays a central role in the theory of shape analysis (Small 1996), and it is the metric that ties together the collection of methods for analysis of shape variation that constitute the new "morphometric synthesis" (see Bookstein 1996a).

Shape spaces are curved, non-Euclidean spaces (e.g., Bookstein 1991; Rohlf 1996; Small 1996). To apply the usual methods of statistics, it is therefore advantageous to project shape space onto a linear, Euclidean space, just as it is useful to have flat maps of the curved surface of the Earth. And just as maps are based on projections of the Earth's curved surface onto tangent planes or cylinders, it is convenient to use a linear tangent space that touches the curved shape space at the location of the consensus configuration (but recall that both shape space and tangent space are multidimensional). To ensure proper alignment of configurations in tangent space, we scaled the coordinates of the consensus configuration to ensure that it has unit centroid size, and we rescaled the final coordinates of each configuration (corresponding to the scaling option "1/Cos(rho)" in Rohlf 1997). The effects of these scaling steps are extremely subtle, in the order of 0.001% of the shape variation in our dataset; for studies of small amounts of shape variation (such as in this study of adults from a single species), the results of analyses are expected to be effectively identical regardless of the choice of scaling option.

For this study, we used SAS for all statistical analyses, including a routine for the generalized least-squares algorithm for Procrustes analysis (adapted from Rohlf and Slice 1990) written in the SAS/IML language. An alternative procedure for users of standard software is described in the Appendix.

Procrustes ANOVA

Because calculation of Procrustes coordinates is based on the algebra of sums of squares, shape deviations from the consensus, or mean, configuration can be partitioned in a way analogous to the deviations from a grand mean in conventional ANOVA (Goodall 1991; Smith et al. 1997). Therefore, they are amenable to the two-factor ANOVA recommended by Palmer and Strobeck (1986) and Palmer (1994) for studies of asymmetry. This ANOVA uses the coordinates of the Procrustes-aligned configurations (all three replicates of each wing) as the data.

The identity of specimens entered the model as a random effect and body sides as a fixed effect. The among-specimen main effect stands for individual shape variation (e.g., some flies may have more rounded wings than others). The main effect for the sides expresses directional asymmetry in shape (e.g., left wings may tend to be consistently narrower than right wings). The side × specimen interaction serves as a measure of fluctuating asymmetry; it is the deviation of each individual's asymmetry from the overall average of asymmetry in shape. Finally, the residual term, variability among replicates, is measurement error. As in the analyses of centroid size, we used maternal age (young/old) as an additional fixed effect in the analysis. We used the side × individual interaction as the error term for tests of the main effects of side and individuals, and the among-individuals effect for tests of the maternal age effect.

There are more degrees of freedom in Procrustes ANOVA than in conventional ANOVA (e.g., Goodall 1991), because the squared deviations are summed over all the landmark coordinates (instead of a single sum of squares in conventional ANOVA). Therefore, the number of degrees of freedom is that for ordinary ANOVA multiplied by the shape dimension, which is, for two-dimensional coordinate data, twice the number of landmarks minus four (the number of coordinates minus two dimensions for translation and one each for scaling and rotation).

Statistical Inference for Procrustes ANOVA

For testing the statistical significance of ANOVA effects, we used permutation tests (Good 1994; Edgington 1995). We chose this nonparametric approach to avoid making assumptions about the specific distribution of shapes around the mean landmark configuration. Permutation tests also avoid the rather stringent statistical constraints of the covariance structure described by Goodall (1991). The assumptions of the analysis are thus that the effects are additive and that residuals are independent, random, and have homogeneous variance.

As recommended by Edgington (1995), we performed random permutations for each main effect separately (e.g., exchanging wing configurations across individuals within sides or across sides within individuals). This means that each test had a separate null hypothesis that postulated only the absence of the particular effect tested. For tests of interaction effects, the data had to be adjusted to eliminate main effects (subtract both side and fly means from each value and add the grand mean) before permuting observations across both sides and individuals (Good 1994). Each test used 10,000 random permutations of the observations.

Localization of Effects

To assess how much of the shape variation was due to each landmark, we decomposed the Procrustes mean squares for each effect according to the landmarks. That means, in this analysis we summed x- and y-mean squares of each landmark separately, but did not sum across landmarks. From these mean squares, we computed variance components for each of the effects according to the expected mean squares (Palmer and Strobeck 1986; Sokal and Rohlf 1995).

This procedure attempts to localize the shape variation for the various factors considered in the ANOVA. However, such localized shape features must be interpreted cautiously because overall shape variation is measured by generalized least-squares superimposition, which is based on a global fit of overall shape. There is a difficulty with this global method if one or a few landmarks are much more variable than all the others (the "Pinocchio effect" of Chapman 1990). The least-squares algorithm tends to spread variation from these variable landmarks to the others (the sum of many small

squared distances tends to be less than the sum of one or a few large squared distances). This property of least-squares Procrustes fits is well known; several authors have advocated the use of resistant-fit methods in this situation (e.g., Chapman 1990; Rohlf and Slice 1990). Resistant-fit methods, however, do not produce residuals compatible with the Procrustes metric used in subsequent statistical analyses, and thus should not be used in this context (Bookstein 1996b; F. J. Rohlf, pers. comm.). The method used here, analysis of least-squares Procrustes residuals by landmarks, tends to underestimate the differences among landmarks in their degree of variability. Any Pinocchio effect that this method finds, however, reflects true differences in variability among landmarks. This procedure therefore is a valid, although somewhat conservative, method to discover localized variation.

Comparison of Covariance Matrices

The analyses quantifying shape variation in a sample, attributing it to different sources and partitioning it according to landmarks, extract only a part of the information contained in morphometric data. In addition, the same data can be used to infer relationships among landmarks from the patterns of covariances of their positions. For instance, if one landmark is displaced distally, do its neighboring landmarks tend to move in the same or in different directions or is there no association at all among landmark displacements? (We use terms like "movement" to visualize the relative displacement of landmarks, e.g., in a comparison of the left and right wings.)

To characterize these patterns of joint displacements of landmarks, we analyzed the covariance matrices of the coordinates of superimposed landmarks (equivalent to covariance matrices of Procrustes residuals). We used covariance matrices because they preserve the Procrustes metric, which underlies the entire study; the use of correlation matrices would eliminate this common scale for shape variation. We computed covariance matrices for the between-fly effect from individual means (both sides and all replicates averaged for each fly), for fluctuating asymmetry (fly × side interaction) from individual left-right differences (wing averages from all replicates), and for measurement error from the residual variation of the replicate measurements about the average for each wing. After computing matrices of sums of squares and cross-products (SSCP) and dividing by the appropriate degrees of freedom, we separated effects according to the expected mean squares (see Palmer and Strobeck 1986: table 3) by subtracting the fly × side covariance matrix from the among-fly covariance matrix, and the measurement covariance matrix from the fly × side covariance matrix.

As a first step to address the principal question of our case study, whether the developmental basis of fluctuating asymmetry is the same or qualitatively distinct from that of individual variation, we performed an overall comparison of covariance matrices. If the same developmental processes that are responsible for variation among individuals also generate fluctuating asymmetry, the respective covariance matrices should be similar. In contrast, measurement error does not have a corresponding cause, and therefore is expected to differ from the other two in a random manner. We used a Mantel test of matrix correlations for this purpose (Mantel 1967; Cheverud et al. 1989; Manly 1991; Sokal and Rohlf 1995). The null hypothesis in this test is that two matrices are completely dissimilar, that is, that corresponding entries of the matrices are uncorrelated.

As the test statistic, we calculated the matrix correlation, the correlation between corresponding entries of the covariance matrices. Because variances and covariances jointly define patterns of variation, we included the diagonal along with the off-diagonal entries of the matrices in the calculation of matrix correlations. Because covariance matrices are symmetric, each off-diagonal element appears twice; therefore, we calculated matrix correlations from the upper triangular part of each matrix (entries on or above the diagonal).

The Mantel test simulates the distribution of matrix correlations under the null hypothesis by randomly permuting rows and columns of one of the matrices. In the context of a morphometric study, however, it is necessary to take into account the geometry of the landmarks as well. Therefore, we modified the null hypothesis for geometric morphometrics: The displacements of different landmarks are uncorrelated among matrices (but because the scatter around the mean location of each landmark can be directed, the modified null hypothesis permits correlations between the x and y coordinates of each landmark). Because we are interested in the associations among landmark movements, the x and y coordinates of each landmark are not independent even under the null hypothesis and are certainly not interchangeable. We therefore interchanged the landmarks, that is, we permuted pairs of rows and columns of the covariance matrices; this procedure maintained the association between the x and y coordinates of each landmark. This permutation procedure was carried out 10,000 times. For each iteration, the landmarks were permuted for one matrix and its matrix correlation to the other matrix was computed. The resulting null distribution was then compared to the matrix correlation calculated for the original pair of matrices.

Identifying Patterns of Variation: Principal Components

For a more detailed investigation of the patterns of joint displacements of landmarks, we used principal component analysis (PCA). In many applications of PCA, the first few principal components (PCs) account for most of the total variation contained in a dataset and therefore can summarize multidimensional variation effectively in far fewer dimensions than originally included in the analysis. Such data reduction is especially helpful in the current context, because the original number of dimensions is twice the number of landmarks (26 in this study). At least four of these dimensions contain no variation (degrees of freedom lost for scaling, translation, and rotation in the Procrustes fit), and many more can be expected to account only for trivial amounts of variability.

Besides data reduction, PCA also analyzes and display patterns of variation so that they can be interpreted biologically. The PCs can be viewed as features of shape variation that are mutually uncorrelated, and therefore can be examined one by one. For instance, the PCs allow an assessment of relations between the patterns of covariation corresponding

TABLE 1. Asymmetry of overall wing size. Centroid size is the dependent variable in an analysis of variance according to the model recommended by Palmer and Strobeck (1986), which includes maternal age as an additional effect. The denominator used to calculate F-values for the effect of maternal age is the fly mean square, for the main effects of fly and side it is the mean square of the fly × side interaction, and for the fly × side interaction it is the mean square of measurement error.

Source	df	SS	MS	F
Maternal age	1	0.233	0.233	0.88
Fly	68	17.976	0.264	102.07***
Side	1	0.0723	0.0723	27.91***
Fly × side	69	0.179	0.00259	3.13***
Measurement	280	0.232	0.000828	

*** $P < 0.001$.

TABLE 2. Asymmetry of shape. Analysis of variance used Procrustes sums of squares as a measure of overall variation in shape. The denominator used to calculate F-values for the effect of maternal age is the fly mean square, for the main effects of fly and side it is the mean square of the fly × side interaction, and for the fly × side interaction it is the mean square of measurement error. We used a separate permutation test to determine statistical significance of each effect.

Source	df	SS	MS	F
Maternal age	22	0.032399	0.0014727	14.87***
Fly	1496	0.148200	0.0000991	4.96***
Side	22	0.003246	0.0001475	7.38**
Fly × side	1518	0.030331	0.0000200	3.51***
Measurement	6160	0.035036	0.0000057	

** $P < 0.01$; *** $P < 0.001$.

to the factors in the ANOVA. Principal components coefficients are usually presented in tabular form, but in the context of geometric morphometrics, these coefficients can be presented graphically in direct relation to the landmark positions on the fly wing. Therefore, we displayed PC coefficients directly as movements of landmarks. As an alternative form of presentation, transformation grids (e.g., Bookstein 1991; Rohlf 1997) could be used for the same purpose.

For our case study of the sources of morphometric variation of fly wings, PCA can be used to refine the comparisons of covariance structures: Whereas the Mantel test (above) considers the overall similarity of covariance matrices, PCA can isolate specific features of variation for comparison. We tested the opposite expectations of agreement (individual variation and fluctuating asymmetry) or random difference (either of these vs. measurement error) among PCs in two different ways. The first test examined the null hypothesis that the PCs for the different effects were no more similar than pairs of random vectors (corresponding to the null hypothesis of the Mantel test). We tested this using Monte Carlo (MC) simulation of random vectors (Cheverud 1982b; Klingenberg and Zimmermann 1992). The procedure generated 100,000 pairs of random vectors as points on a 22-dimensional unit sphere and computed the absolute angles between these vectors. This null distribution of angles was then compared to the angles obtained from the original sample.

The second test examined the opposite null hypothesis, that the PCs are the same for the different effects and differ only by sampling error. We implemented this with a bootstrap test for the angles between PCs (Klingenberg 1996). Observations were resampled (with replacement) from the distribution of within-sample PC scores to generate a null distribution with parallel PC axes. The test then examined how often, given this null distribution, the absolute angles between PCs for different effects exceeded the ones in the original sample. Resampling of observations was done at the levels suggested by the hierarchical design of our study (i.e., individual averages for between-fly variation, individual differences of left and right wings for fly × side interaction, and residuals of replicates for measurement error). We used 10,000 bootstrap iterations for this test.

Antisymmetry and Allometric Effects

To check the data for antisymmetry, we visually examined scatter plots of vectors of left-right differences for each landmark after superimposition by the Procrustes algorithm. There was no evidence for clustering of these vectors (as the equivalent to bimodal distributions of left-right differences) that would have suggested antisymmetry. We also examined whether size affects the asymmetry of wing size or shape (Palmer 1994). The regressions of signed and unsigned asymmetry of centroid size against mean centroid size were both nonsignificant (P-values > 0.3; $r^2 < 0.015$).

To test for size effects on shape asymmetry, we used multivariate regression (e.g., Jobson 1992) of vectors of shape asymmetry onto mean centroid size. We defined a multivariate equivalent to univariate unsigned left-right differences, where the right-left difference is used whenever the left-right difference is negative. Likewise, we changed the signs of all coordinate differences (from left-right to right-left) whenever the inner product of a left-right difference vector with the vector of mean left-right difference was negative. The regressions of both signed and "unsigned" shape asymmetries were not related to size (P-values > 0.3; each regression accounted for less than 1.5% of total Procrustes sums of squares). These regression analyses provided no evidence of size effects on asymmetry, nor did visual inspection of plots of asymmetry in the PC scores. Therefore, no size corrections are necessary, and the additive model underlying the two-factor ANOVA is appropriate for both the size and shape data.

RESULTS

Variation and Asymmetry in Size

All effects in the ANOVA for centroid size were significant, except for maternal age (Table 1). Size variation among individuals takes up the largest part of the variation. In addition, there was subtle but significant directional asymmetry of wing size (mean centroid size of the left wing, 6.802 mm; right wing, 6.829 mm), as well as a small amount of fluctuating asymmetry.

Quantifying Variation and Asymmetry in Shape

The Procrustes ANOVA of shape variation showed that all effects of the model were statistically significant (Table 2). A more differentiated pattern emerged when variance components from the Procrustes ANOVA were apportioned by landmarks (Table 3).

TABLE 3. Variance components for the effects in the Procrustes ANOVA listed by landmark. This is an approximate measure of the variability around the mean location of each landmark. These are minimal estimates of differences among landmarks, because the least-squares superimposition tends to underestimate the degree to which shape variation is localized. All entries have been multiplied by 10^8 to make them more readable.

	Effect				
Landmark	Maternal age	Fly	Side	Fly × side	Measurement
1	226	86	0	18	36
2	349	493	4	285	319
3	8	74	0	25	20
4	12	43	3	12	21
5	23	67	1	17	22
6	12	42	7	29	30
7	24	31	2	10	18
8	14	111	1	14	14
9	10	98	3	18	14
10	2	96	12	15	21
11	4	19	0	6	18
12	2	109	21	16	21
13	3	50	7	10	16

Landmark 2 dominated for all effects in the analysis, except for directional asymmetry. Therefore, it is clear that this landmark caused a large Pinocchio effect at most levels of variation included in the ANOVA, which was most extreme for the fly × side interaction (fluctuating asymmetry) and for measurement error (Table 3). In contrast, landmark 11 had consistently low amounts of variability. Others with mostly small amounts of variation were landmarks 7 and 13, which are located near landmark 11 on the basal part of the wing blade.

The effect of maternal age on shape was highly statistically significant and affected mostly the landmarks at the anterior wing margin, where landmarks 1 and 2 took up disproportionate shares of the variance (Table 3). In contrast to the mainly local effects of maternal age, the effects of individual variation were distributed more evenly among landmarks. Apart from the Pinocchio effect caused by landmark 2, the largest share of the among-fly variation is located at the landmarks defining anterior and posterior crossveins (8, 9, 10, and 12). For directional asymmetry, the largest left-right differences were found at the posterior crossvein (10, 12), in the proximalmost landmarks (6, 13), and at the anterior wing margin (2). For fluctuating asymmetry, the landmarks that take up the largest shares of variation are located in the anterior part of the wing (2, 3, and 6). In summary, the relative amounts of variation at each landmark vary considerably among the factors included in the analysis.

Overall Similarity of Covariance Matrices

The matrix correlations between all three covariance matrices were high and statistically significant (individual variation and fluctuating asymmetry: MC = 0.75, $P = 0.005$; individual variation and measurement error: MC = 0.77, $P = 0.001$; fluctuating asymmetry and measurement error: MC = 0.98, $P < 0.0001$). They all remain significant after correction for multiple tests with the sequential Bonferroni method. The high correlations are partly due to the extremely high variability in the x coordinate of landmark 2, which is much more variable than any of the other landmarks (see Table 3). Apart from this effect, it is not clear whether the differences between the matrix correlations for the three comparisons provide meaningful information. But further investigation of the patterns of covariation among landmarks is clearly warranted.

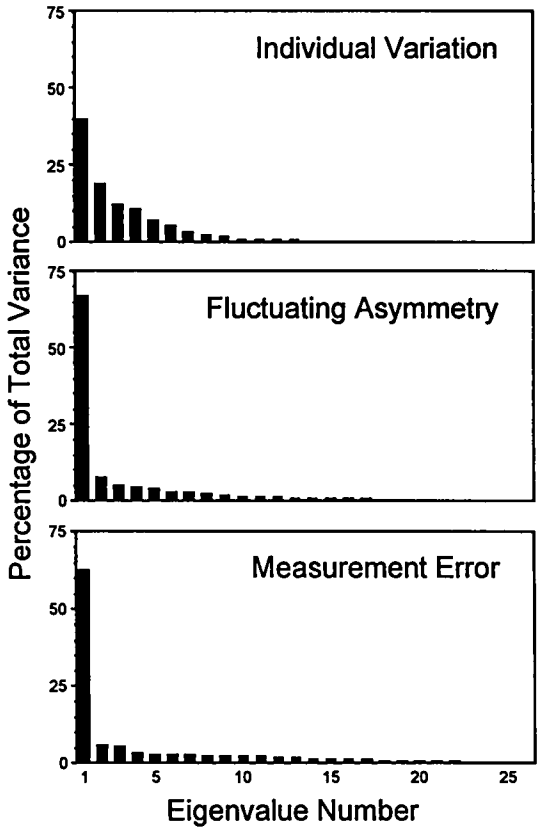

FIG. 3. Percentages of total shape variation taken up by the principal components for the covariance matrices of individual variation, fluctuating asymmetry, and measurement error.

Patterns of Integration among Landmarks

Principal component analyses showed that most variation was concentrated in just a few dimensions. The PC1 accounted for more than twice the amount of variation taken up by any of the other PCs at all levels of the analysis and dominated most extremely for fluctuating asymmetry and measurement error (Fig. 3). For individual variation, several PCs accounted for relatively large amounts of variability. In contrast, for fluctuating asymmetry and measurement error, the values dropped sharply from the PC1 to the PC2 and tapered off very slowly in subsequent PCs.

We displayed the features of variation associated with the dominant PCs graphically as plots of the PC coefficients superimposed onto a drawing of the wing (Figs. 4–6). The dominance of the PC1 was linked to the large variability of landmark 2 for individual variation, fluctuating asymmetry, and measurement error. The PC1 coefficients of this landmark were by far the largest of all landmarks for all three sources of variation, and this large amount of variation was directed

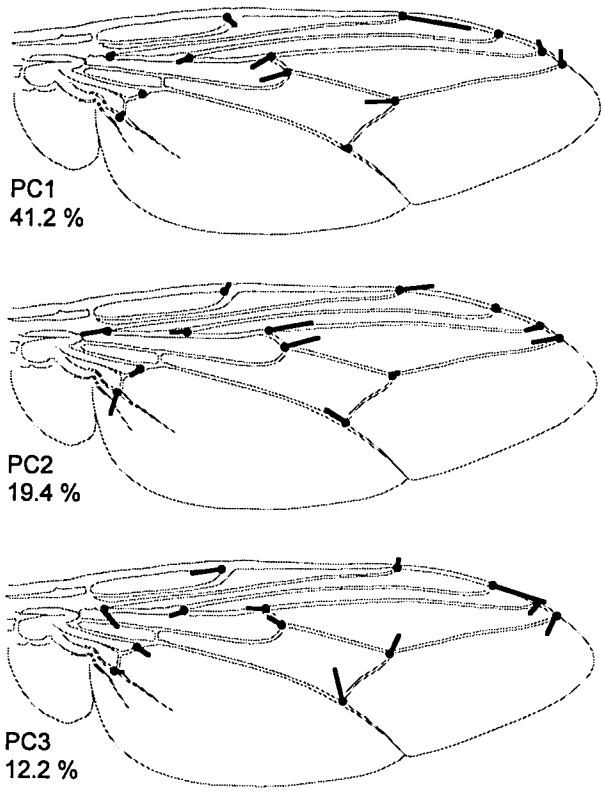

FIG. 4. Principal component analysis (PCA) of joint variation in landmark positions for individual variability. The PCA used the among-individual covariance matrix corrected for between-sides variation and measurement error. The diagrams visualize the PC coefficients of each landmark in x and y directions by a line originating at the mean location of the landmark (dots) and ending at the locations to which the landmarks would move for an imaginary wing with an arbitrary PC score of + 0.15 Procrustes units. This is a very large shape change, and an exaggeration of the variation in the dataset: For the PC1 it is 6.5 standard deviations (SDs) from the mean configuration, for the PC2 9.4 SDs, and for the PC3 11.9 SDs. The percentage in the lower left corner of each box is the proportion of total among-individual Procrustes mean squares for which the respective PC accounts.

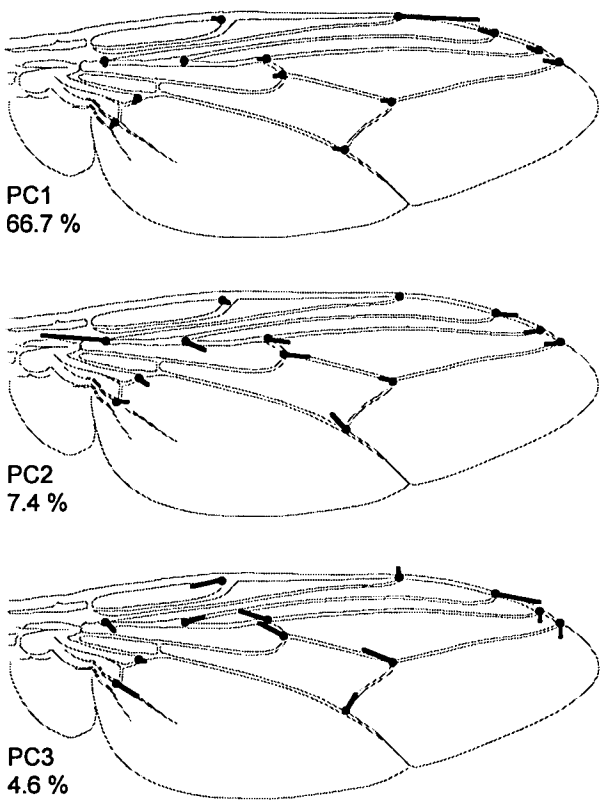

FIG. 5. Principal component analysis (PCA) of joint variation in landmark positions for fluctuating asymmetry. The PCA used the side × fly covariance matrix corrected for measurement error. The end points of the lines are at the location of the landmarks for a wing with a PC score of + 0.15 Procrustes units: For the PC1, this is 6.7 standard deviations (SDs) from the mean configuration, for the PC2 20.1 SDs, and for the PC3 25.6 SDs. (For further explanations of the diagrams, see Fig. 4.)

along the wing margin (Figs. 4–6). In contrast, the coefficients of landmark 2 were small for the PC2 and all subsequent PCs, as most of the variability of this landmark was already taken up by the PC1. For all three sources of variation, the PC2 and PC3 of a complete analysis were very similar to the PC1 and PC2 of an analysis excluding landmark 2. This suggests that the strong Pinocchio effect caused by landmark 2, despite the disproportionate amount of variation associated with it, did not compromise the analysis of patterns of covariation among the other landmarks.

The angles between the three PC1s also reflect the similarity due to this Pinocchio effect. The PC1 of individual variation, with the least extreme Pinocchio effect, was fairly distinct from both the PC1 of fluctuating asymmetry (34.3°) and the PC1 of measurement error (32.7°), but the latter two were more similar to each other (8.4°). All of these associations were highly significant—the observed angles were smaller than any of the 100,000 angles between pairs of random vectors (recall that these are vectors in 22-dimensional space). Moreover, the bootstrap test was consistent with the null hypothesis that the covariances matrices share the same PC1s (individual variation and fluctuating asymmetry, $P = 0.998$; individual variation and measurement error, $P = 0.999$; fluctuating asymmetry and measurement error, $P = 0.63$). Considered together, the results of these tests clearly support the similarity of the three PC1s, which reflects the dominance of landmark 2 for all sources of variation.

The landmark displacements pertaining to the PC2s of individual variation and fluctuating asymmetry were similar in a number of ways (Figs. 4, 5; the following passage describes the PC coefficient vectors as joint displacements from the mean configuration, moving in the direction that corresponds to an increase of the PC scores). Landmarks 8 and 9 moved distally together, reflecting shifts of the anterior crossvein along the adjoining longitudinal veins. Landmark 12, which is a part of the posterior crossvein, moved along the longitudinal vein toward the wing base; however, landmark 10, at the other end of the same crossvein, did not move in close association. The landmarks of the wing tip (2–5) realigned to form a blunter tip (landmarks 4 and 5 shift proximally) and a more curved anterior wing margin (distal shift of landmark 2 for individual variation or landmark 3 for fluctuating asymmetry). Finally, landmark 6 moved toward the wing base for both PC2s. In agreement with these similarities, the angle

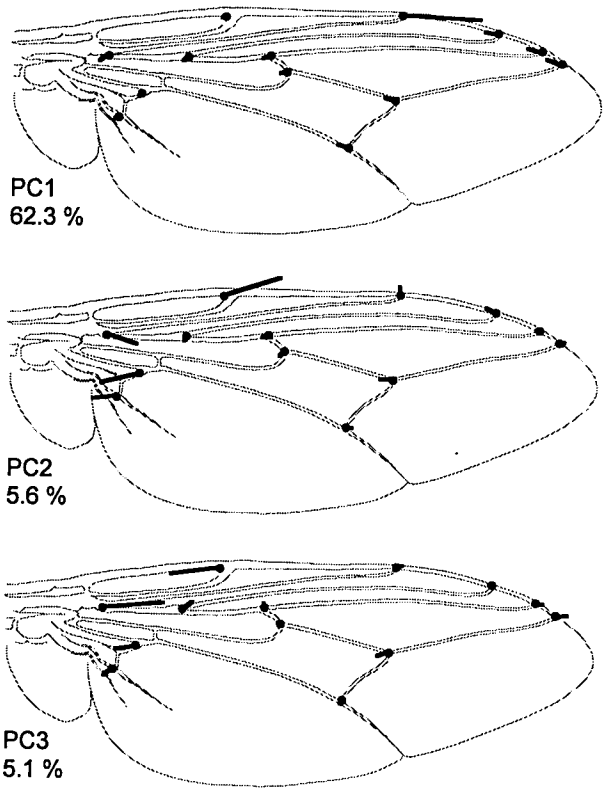

FIG. 6. Principal component analysis (PCA) of joint variation in landmark positions for measurement error. The PCA used the covariance matrix among replicate measurements. The end points of the lines are at the location of the landmarks for a wing with a PC score of + 0.15 Procrustes units: for the PC1, this is 17.0 standard deviations (SDs) from the mean configuration, for the PC2 56.5 SDs, and for the PC3 59.2 SDs. (For further explanations of the diagrams, see Fig. 4.)

between the PC2s for individual variation and fluctuating asymmetry (57.1°) was significantly smaller than between random vectors ($P = 0.003$), and the bootstrap test suggests the PC2s were not statistically distinguishable from each other ($P = 0.925$).

The similarities between the PC3s of individual variation and fluctuating asymmetry were even closer than those of the PC2s. The main differences between PC3s were the directions of the movements of landmarks 7 and 10 and some relatively minor differences in the relative magnitudes of movements at different landmarks. The angle between the two PC3s was 44.6°, which is significantly smaller than an angle between random directions ($P = 0.00006$), and the bootstrap test did not distinguish the two PC3s ($P = 0.98$).

In contrast to the second and third PCs of individual variation and fluctuating asymmetry, those of measurement error showed patterns of variation with little discernible coordination (Fig. 6). Landmarks tended to shift independently of each other and neighboring landmarks often moved in opposite directions (landmark 1 separate, and 6 and 13 vs. 11 for PC2; 1 vs. 6 for PC3). The angles between these two PCs and the corresponding PCs for individual variation or fluctuating asymmetry were not smaller than expected for random vectors (P-values in the Monte Carlo test ranging from 0.08 to 0.79).

DISCUSSION

We have presented new methodology for studies of morphological asymmetry by combining the method of quantifying individual variation and asymmetry by ANOVA (Leamy 1984; Palmer and Strobeck 1986; Palmer 1994) and the analysis of shape as configurations of landmarks (Bookstein 1991, 1996a; Small 1996). We have extended earlier studies of asymmetry that used similar methods (Bookstein 1991; Auffray et al. 1996; Smith et al. 1997) by providing a procedure to test the significance of observed effects statistically. Moreover, we have introduced new methodology for studying the patterns of covariance among landmarks for the various causal effects in the ANOVA. This source of information has not been considered by other studies, which have focused almost exclusively on the amount of fluctuating asymmetry rather than on patterns of localized landmark variation. A multivariate analysis of these patterns, however, can reveal new insight about fluctuating asymmetry.

We demonstrate this methodology with a case study of fluctuating asymmetry and individual variation in the wings of tsetse flies. Here we attempt to interpret the results of this case study. Given the relatively small sample used in this study and the lack of similar studies for comparison, the biological conclusions will necessarily be somewhat tentative and await confirmation from independent datasets (e.g., C. P. Klingenberg and S. D. Zaklan, unpubl.). Nevertheless, they clearly illustrate the potential of this approach to the study of individual variation and asymmetry.

Procrustes ANOVA

Before patterns of morphological asymmetry can be analyzed in detail, the statistical significance of the causal effects should be tested by ANOVA to ensure that the variance components to be analyzed are not simply due to sampling or measurement error (e.g., Palmer 1994). Moreover, and perhaps more importantly, ANOVA provides a means to estimate the components of variance that correspond to the various causal effects (Leamy 1984; Palmer and Strobeck 1986; Palmer 1994).

The Procrustes method can be used for ANOVA (e.g., Goodall 1991) because it uses the same algebra of sums of squares. Our data illustrate a practical difficulty, however, because the distribution of shapes around the mean configuration is complex. The amount and direction of variation differs from one landmark to another and it reflects the local geometry of wing veins. Three observations from our case study illustrate this complexity, which may pose problems for applying statistical constraints to simplify models of landmark covariance (see Goodall 1991: section 5). The first and most obvious is the Pinocchio effect at landmark 2: Variability at this landmark exceeds variation at any other landmark severalfold. This effect exists for most of the factors of the ANOVA design including measurement error (Table 3) and indicates that landmark 2 is not only biologically variable but also difficult to locate for human observers. Second, both biological variability and measurement error depend critically on the local geometric and anatomical structure. For our data, where the landmarks are at the intersections of wing veins, the dominant components of variation (first

few PCs) are often oriented along one of the adjoining wing veins (Figs. 4–6). Third, landmark displacements can be highly coordinated depending on their relative positions (e.g., landmarks 8 and 9 at the anterior crossvein; Figs. 4, 5) and on the geometry of the structures involved (e.g., wing margin); variation of different landmarks clearly is not mutually independent. These three observations suggest that, to be realistic, statistical models of landmark covariation may need to include fairly intricate details of biological processes and of the geometry of the structure studied.

To avoid these statistical difficulties altogether, it is possible to resort to permutation (Good 1994; Edgington 1995) and bootstrap (Efron and Tibshirani 1993) methods for statistical tests. These are nonparametric methods based on resampling of the observations, and thus make no assumptions about the distribution or covariance structure. Interestingly, we found that in this dataset *P*-values from permutation tests differed little from those derived from the *F*-distribution as for conventional ANOVA (data not presented). Nevertheless, it is not clear whether the parametric methods will be similarly robust in other cases. We therefore recommend the use of resampling methods whenever the multivariate distribution of Procrustes residuals appears complex.

In our case study of tsetse flies, the results of analyses of centroid size closely agree with the overall picture of the univariate analyses of interlandmark distances (McIntyre and Gooding 1998). The effect of maternal age on size (0.7% difference) is not statistically significant, but there is subtle, but significant directional asymmetry (0.4% difference). In addition, there is ample individual size variation and fluctuating asymmetry.

A similar pattern holds for the analyses of shape, but all effects included in the ANOVA are statistically significant, including maternal age. Yet, the effects of both maternal age and directional asymmetry are very subtle, and the biological relevance for the effect of maternal age is unclear. Similar subtle but significant directional asymmetry of wing shape has also been found in other flies (Klingenberg et al. 1998) as well as in honey bees (Smith et al. 1997; without statistical testing). Furthermore, the ANOVA showed that the data contain statistically significant fluctuating asymmetry and individual variation; a more detailed investigation of their relationship is therefore warranted.

Overall, the Procrustes ANOVA is a simple and powerful alternative to a collection of separate tests of linear measurements. This is especially important because morphometric measurements usually are intercorrelated and even Bonferroni adjustment for multiple tests (e.g., Sokal and Rohlf 1995) would therefore be problematic. Because Procrustes ANOVA provides a single overall test for effects on shape, it circumvents this difficulty entirely. However, the main advantage of analyses using the methods of geometric morphometrics is their ability to pinpoint the location and direction of specific features of variation.

Variation and Covariation among Landmarks

The results of our analyses indicate that morphological variation tends to affect different parts of the wing differently. The decomposition of variance components according to landmarks showed that the landmarks differ in the amount of variation, and the causal factors included in this analysis tend to partition variation among landmarks in different proportions (Table 3). The factor that especially stands out in this respect is directional asymmetry. It appears that this is not simply a random outcome linked to the subtlety of this effect, however, because similar patterns of directional asymmetry have also been found in two different species of flies (Klingenberg et al. 1998).

Analyses of the covariation among landmarks by PCA also revealed distinct patterns at the different levels examined. The landmark displacements are mostly oriented in the direction of some structure, for instance, along a wing vein or parallel to the wing margin (Figs. 4, 5). Moreover, there are coordinated displacements of several landmarks, for instance, those at the ends of the crossveins and those near the wing tip.

In the context of fluctuating asymmetry, the fact that there are identifiable patterns of variation is remarkable in itself. Such patterns have been sought for some time, but usually have been difficult to demonstrate (Leamy 1993; Palmer 1994; Leamy et al. 1997; Møller and Swaddle 1997). The significant component for fluctuating asymmetry of overall wing size further suggests coordination throughout the wing. Although they used different methods, multivariate studies of fluctuating asymmetry in the mouse mandible (Leamy 1993; Leamy et al. 1997) and our study concur in that both systems show that (signed) asymmetries covary among traits that are related developmentally.

Such covariation is to be expected among traits that share a common developmental precursor on each body side. For instance, a left-right difference in the growth of wing imaginal discs may be reflected in multiple traits that depend on the size of the developing wing. Morphometric associations among traits can reflect developmental relationships in left-right asymmetry just as in variation among individuals (e.g., Riska 1986; Cowley and Atchley 1990; Atchley and Hall 1991). Random differences between structures on either body side can thus also *create*, not just obscure, patterns of covariance among the traits of these structures (Sakai and Shimamoto 1965; Palmer 1994, p. 360). This developmental relationship is a phenomenon often omitted from theoretical models of fluctuating asymmetry (e.g., Whitlock 1996), but can substantially influence the patterns of covariance among asymmetries of multiple traits.

Fluctuating Asymmetry and Individual Variation

In addition to demonstrating the methods presented in this paper, our case study addressed the question whether there is evidence of special developmental processes involved in the control of morphological asymmetry (e.g., Møller and Swaddle 1997). For this purpose, we examined whether fluctuating asymmetry and individual variation showed similar patterns of covariation among landmarks.

Comparisons between the patterns of landmark covariation for individual variation and fluctuating asymmetry showed considerable correspondence. The matrix correlations among covariance matrices for the different causal factors were high and statistically significant in a Mantel test, indicating overall

similarity among matrices. A second test, using the angles between corresponding PCs, confirmed the similarity between the matrices for fluctuating asymmetry and individual variation. This test also established, as was expected, that the similarity did not extend to the matrix of measurement error.

Given the correspondence between individual variation and fluctuating asymmetry, there is no reason to suspect that developmental processes other than those influencing the mean shape across body sides specifically affect asymmetry, or buffer against it. Yet, each pattern of variation accounted for different amounts of variability for fluctuating asymmetry and individual variation, thus indicating that there is not complete congruence. It is possible that the same developmental processes respond in a different manner to environmental or genetic differences between individuals than they respond to random differences between body sides. This agrees with results from a theoretical model in which most of the properties of fluctuating asymmetry (e.g., correlations with trait size and heterozygosity) could be generated without assuming any developmental processes specifically affecting asymmetry (C. P. Klingenberg and H. F. Nijhout, unpubl.). Clearly, this question needs to be addressed in more detail by studies aimed more directly at the developmental mechanisms involved.

Morphometric Variation and Developmental Mechanisms

Some of the patterns of landmark variation can be interpreted in terms of developmental processes. Such interpretation, while necessarily tentative in the absence of experimental data, is crucial for a full understanding of the developmental basis of morphometric variation.

The large amount of variation at landmark 2, which is clearly directed along the wing margin, may be a joint consequence of the geometry of the wing and the processes of wing vein development. More specifically, this pattern could result from the combination of the narrow angle at which the longitudinal vein intersects the wing margin with lateral signaling (lateral inhibition), which plays an important role in the development of wing veins (e.g., Garcia-Bellido and de Celis 1992; Sturtevant and Bier 1995). Wing-vein development involves several steps: first a series of "coordinate genes" establish longitudinal domains whose boundaries define the location of vein formation; second, the prospective veins are narrowed to stripes seven to eight cell diameters across through the joint action of genes that promote vein formation, genes that suppress veins in the intervein regions, and genes involved in neurogenesis; third, vein differentiation occurs after lateral signaling further reduces the width of the prospective vein to two to three cell diameters (Sturtevant and Bier 1995). A small lateral displacement or thickening of a longitudinal vein by variability in these processes will move the landmark an equally small distance along the wing margin (and costal vein) if they meet at a right angle; but if they abut at a narrow angle, the same initial sideways displacement of the vein will produce a much larger shift of the landmark. Of all longitudinal veins, the second, which defines landmark 2, is the one that meets the wing margin at the narrowest angle (Fig. 1). Therefore, for a given amount of variation in vein width or position, this landmark should experience the most extreme landmark displacement. This scenario illustrates how morphological variation can result from the way in which a developmental process and local tissue geometry interact.

The possible role of lateral signaling as the developmental origin of morphological variability is especially interesting in conjunction with morphological asymmetry. One of the molecular mechanisms responsible for lateral signaling in various developmental contexts is the Delta–Notch signaling pathway of *Drosophila* (e.g., Artavanis-Tsakonas et al. 1995; Collier et al. 1996; Gerhart and Kirschner 1997; Kimble and Simpson 1997). This pathway is involved prominently in wing-vein formation, where it mediates the narrowing of the veins to their final width (Garcia-Bellido and de Celis 1992; Sturtevant and Bier 1995; de Celis et al. 1997; Huppert et al. 1997). In the context of asymmetry, this is especially intriguing because a homologue of the *Notch* gene has been implicated as a modifier gene of the asymmetry of bristle counts in the blowfly *Lucilia cuprina* (Batterham et al. 1996; Davies et al. 1996). Although there are no morphometric data in this case, it is conceivable that there are also effects on wing shape.

The landmarks of the anterior crossvein (8, 9) tend to move in unison, both for individual variation and fluctuating asymmetry (Figs. 4, 5). These landmarks may move together because crossveins form after the longitudinal veins that they connect (in *Drosophila*; Waddington, 1940; Garcia-Bellido and de Celis, 1992); therefore, most of the variation may occur as shifts along the longitudinal veins. The geometry of veins in the tsetse flies is more complex than in *Drosophila* because of the curvature of the fourth longitudinal vein (Media 1 + 2) at landmarks 9 and 10. Thus, the points where the longitudinal vein bends must also, simultaneously, determine the crossveins. Perhaps this and the divergence of adjoining longitudinal veins account for the less coordinated variation of the landmarks (10, 12) at the posterior crossvein.

Although these interpretations are necessarily hypothetical, this study provides a promising new approach to address questions about the nature of developmental stability in a rigorous quantitative manner. Further studies using similar methods in different study systems, together with experimental approaches, should greatly improve our understanding of the developmental basis of fluctuating asymmetry.

Acknowledgments

We thank F. L. Bookstein and A. R. Palmer for extensive discussion before and during preparation of the manuscript. F. L. Bookstein also kindly provided a preprint of the paper by Smith et al. (1997). F. L. Bookstein, J. H. Graham, L. J. Leamy, A. R. Palmer, F. J. Rohlf, M. L. Zelditch, and two anonymous reviewers commented on earlier versions of the manuscript. Financial support was provided by operating grants from the Natural Sciences and Engineering Research Council of Canada to R. H. Gooding and A. R. Palmer, and by a postdoctoral fellowship from the Swiss National Science Foundation to CPK.

Literature Cited

Artavanis-Tsakonas, S., K. Matsuno, and M. E. Fortini. 1995. Notch signaling. Science 268:225–232.

ATCHLEY, W. R., AND B. K. HALL. 1991. A model for development and evolution of complex morphological structures. Biol. Rev. 66:101–157.

AUFFRAY, J.-C., P. ALIBERT, S. RENAUD, A. ORTH, AND F. BONHOMME. 1996. Fluctuating asymmetry in *Mus musculus* subspecific hybridization: traditional and Procrustes comparative approach. Pp. 275–283 *in* L. F. Marcus, M. Corti, A. Loy, G. J. P. Naylor, and D. E. Slice, eds. Advances in morphometrics. Plenum Press, New York.

BATTERHAM, P., A. G. DAVIES, A. Y. GAME, AND J. A. MCKENZIE. 1996. Asymmetry—where evolutionary and developmental genetics meet. BioEssays 18:841–845.

BOOKSTEIN, F. L. 1991. Morphometric tools for landmark data: geometry and biology. Cambridge Univ. Press, Cambridge, MA.

———. 1996a. Biometrics, biomathematics and the morphometric synthesis. Bull. Math. Biol. 58:313–365.

———. 1996b. Combining the tools of geometric morphometrics. Pp. 131–151 *in* L. F. Marcus, M. Corti, A. Loy, G. J. P. Naylor, and D. E. Slice, eds. Advances in morphometrics. Plenum Press, New York.

CHAPMAN, R. E. 1990. Conventional Procrustes approaches. Pp. 251–267 *in* F. J. Rohlf and F. L. Bookstein, eds. Proceedings of the Michigan morphometrics workshop. Spec. Publ. no. 2. Univ. of Michigan Museum of Zoology, Ann Arbor, MI.

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

———. 1982b. Relationships among ontogenetic, static, and evolutionary allometry. Am. J. Phys. Anthropol. 59:139–149.

CHEVERUD, J. M., G. P. WAGNER, AND M. M. DOW. 1989. Methods for the comparative analysis of variation patterns. Syst. Zool. 38:201–213.

CHIPPINDALE, A., AND A. R. PALMER. 1993. Persistence of subtle departures from symmetry over multiple molts in individual brachyuran crabs: relevance to developmental stability. Genetica 89:185–199.

COLLIER, J. R., N. A. M. MONK, P. K. MAINI, AND J. H. LEWIS. 1996. Pattern formation by lateral inhibition with feedback: a mathematical model of Delta-Notch intercellular signalling. J. Theor. Biol. 183:429–446.

COLLIN, R. 1997. Ontogeny of subtle skeletal asymmetries in individual larvae of the sand dollar *Dendraster eccentricus*. Evolution 51:999–1005.

COWLEY, D. E., AND W. R. ATCHLEY. 1990. Development and quantitative genetics of correlation structure among body parts of *Drosophila melanogaster*. Am. Nat. 135:242–268.

DAVIES, A. G., A. Y. GAME, Z. CHEN, T. J. WILLIAMS, S. GOODALL, J. L. YEN, J. A. MCKENZIE, AND P. BATTERHAM. 1996. *Scalloped wings* is the *Lucilia cuprina Notch* homologue and a candidate for the *Modifier* of fitness and asymmetry of diazinon resistance. Genetics 143:1321–1337.

DE CELIS, J. F., S. BRAY, AND A. GARCIA-BELLIDO. 1997. Notch signalling regulates *veinlet* expression and establishes boundaries between veins and interveins in the *Drosophila* wing. Development 124:1919–1928.

EDGINGTON, E. S. 1995. Randomization tests. 3d ed. Marcel Dekker, New York.

EFRON, B., AND R. J. TIBSHIRANI. 1993. An introduction to the bootstrap. Chapman and Hall, New York.

GARCIA-BELLIDO, A., AND J. F. DE CELIS. 1992. Developmental genetics of the venation pattern of *Drosophila*. Annu. Rev. Genet. 26:277–304.

GERHART, J., AND M. KIRSCHNER. 1997. Cells, embryos, and evolution. Blackwell Science, Malden, MA.

GOOD, P. 1994. Permutation tests: a practical guide to resampling methods for testing hypotheses. Springer-Verlag, New York.

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

HUPPERT, S. S., T. L. JACOBSEN, AND A. T. MUSKAVITCH. 1997. Feedback regulation is central to Delta-Notch signalling required for *Drosophila* wing vein morphogenesis. Development 124: 3283–3291.

JOBSON, J. D. 1992. Applied multivariate data analysis. Vol. II. Categorical and multivariate methods. Springer-Verlag, New York.

KIMBLE, J., AND P. SIMPSON. 1997. The LIN-12/Notch signaling pathway and its regulation. Annu. Rev. Cell Dev. Biol. 13:333–361.

KLINGENBERG, C. P. 1996. Multivariate allometry. Pp. 23–49 *in* L. F. Marcus, M. Corti, A. Loy, G. J. P. Naylor, and D. E. Slice, eds. Advances in morphometrics. Plenum Press, New York.

KLINGENBERG, C. P., AND H. F. NIJHOUT. 1998. Competition among growing organs and developmental control of morphological asymmetry. Proc. R. Soc. Lond. B Biol. Sci. 265:1135–1139.

KLINGENBERG, C. P., AND M. ZIMMERMANN. 1992. Static, ontogenetic, and evolutionary allometry: a multivariate comparison in nine species of water striders. Am. Nat. 140:601–620.

KLINGENBERG, C. P., G. S. MCINTYRE, AND S. D. ZAKLAN. 1998. Left-right asymmetry of fly wings and the evolution of body axes. Proc. R. Soc. Lond. B Biol. Sci. 265: *In press.*

LEAMY, L. 1984. Morphometric studies in inbred and hybrid house mice. V. Directional and fluctuating asymmetry. Am. Nat. 123: 579–593.

———. 1993. Morphological integration of fluctuating asymmetry in the mouse mandible. Genetica 89:139–153.

LEAMY, L. J., E. J. ROUTMAN, AND J. M. CHEVERUD. 1997. A search for quantitative trait loci affecting asymmetry of mandibular characters in mice. Evolution 51:957–969.

———. 1998. Quantitative trait loci for fluctuating asymmetry of discrete skeletal characters in mice. Heredity 80:509–518.

MANLY, B. F. J. 1991. Randomization and Monte Carlo methods in biology. Chapman and Hall, London.

MANTEL, N. 1967. The detection of disease clustering and a generalized regression approach. Cancer Res. 27:209–220.

MARKOW, T. A. 1995. Evolutionary ecology and developmental instability. Annu. Rev. Entomol. 40:105–120.

MCINTYRE, G. S., AND R. H. GOODING. 1998. Maternal effects on offspring quality in tsetse flies (*Glossina* spp.). J. Med. Entomol. *In press.*

MØLLER, A. P. 1996. Development of fluctuating asymmetry in tail feathers of the barn swallow *Hirundo rustica*. J. Evol. Biol. 9: 677–694.

MØLLER, A. P., AND J. P. SWADDLE. 1997. Asymmetry, developmental stability, and evolution. Oxford Univ. Press, Oxford, U.K.

PALMER, A. R. 1994. Fluctuating asymmetry analyses: a primer. Pp. 335–364 *in* T. A, Markow, ed. Developmental instability: its origins and evolutionary implications. Kluwer, Dordrecht, The Netherlands.

———. 1996. Waltzing with asymmetry: is fluctuating asymmetry a powerful new tool for biologists or just an alluring new dance step? BioScience 46:518–532.

PALMER, A. R., AND C. STROBECK. 1986. Fluctuating asymmetry: measurement, analysis, patterns. Annu. Rev. Ecol. Syst. 17:391–421.

PAULSEN, S. M. 1994. Quantitative genetics of butterfly wing color patterns. Dev. Genet. 15:79–91.

PAULSEN, S. M., AND H. F. NIJHOUT. 1993. Phenotypic correlation structure among elements of the color pattern in *Precis coenia* (Lepidoptera: Nymphalidae). Evolution 47:593–618.

RISKA, B. 1986. Some models for development, growth, and morphometric correlation. Evolution 40:1303–1311.

ROHLF, F. J. 1996. Morphometric spaces, shape components and the effects of linear transformations. Pp. 117–129 *in* L. F. Marcus, M. Corti, A. Loy, G. J. P. Naylor, and D. E. Slice, eds. Advances in morphometrics. Plenum Press, New York.

———. 1997. tpsRelw: relative warps analysis. Dept. of Ecology and Evolution, State Univ. of New York at Stony Brook, Stony Brook, NY.

ROHLF, F. J., AND D. SLICE. 1990. Extensions of the Procrustes method for the optimal superimposition of landmarks. Syst. Zool. 39:40–59.

SAKAI, K.-I., AND Y. SHIMAMOTO. 1965. A developmental-genetic study on panicle characters in rice, *Oryza sativa* L. Genet. Res. Cambr. 6:93–103.

SLICE, D. E. 1994. GRF-ND: generalized rotational fitting of N-

dimensional data. Dept. of Ecology and Evolution, State Univ. of New York at Stony Brook, Stony Brook, NY.
SLICE, D. E., F. L. BOOKSTEIN, L. F. MARCUS, AND F. J. ROHLF. 1996. A glossary for geometric morphometrics. Pp. 531–551 *in* L. F. Marcus, M. Corti, A. Loy, G. J. P. Naylor, and D. E. Slice, eds. Advances in morphometrics. Plenum Press, New York.
SMALL, C. G. 1996. The statistical theory of shape. Springer-Verlag, New York.
SMITH, D. R., B. J. CRESPI, AND F. L. BOOKSTEIN. 1997. Fluctuating asymmetry in the honey bee, *Apis mellifera*: effects of ploidy and hybridization. J. Evol. Biol. 10:551–574.
SMITH, L. D., AND A. R. PALMER. 1994. Effects of manipulated diet on size and performance of brachyuran crab claws. Science 264:710–712.
SOKAL, R. R., AND F. J. ROHLF. 1995. Biometry: the principles and practice of statistics in biological research. 3d ed. Freeman, New York.
STURTEVANT, M. A., AND E. BIER. 1995. Analysis of the genetic hierarchy guiding wing vein development in *Drosophila*. Development 121:785–801.
SWADDLE, J. P. 1997. On the heritability of developmental stability. J. Evol. Biol. 10:57–61.
SWADDLE, J. P., AND M. S. WITTER. 1997. On the ontogeny of developmental stability in a stabilized trait. Proc. R. Soc. Lond. B Biol. Sci. 264:329–334.
WADDINGTON, C. H. 1940. The genetic control of wing development in *Drosophila*. J. Genet. 41:75–139.
WALKER, J. A. 1997. Morphometrika: Macintosh software for geometric morphometric analysis of landmark data. Field Museum of Natural History, Chicago.
WHITLOCK, M. 1996. The heritability of fluctuating asymmetry and the genetic control of developmental stability. Proc. R. Soc. Lond. B Biol. Sci. 263:849–854.
ZELDITCH, M. L. 1987. Evaluating models of developmental integration in the laboratory rat using confirmatory factor analysis. Syst. Zool. 36:368–380.

Corresponding Editor: M. Zelditch

APPENDIX

Alternative Procedure for Users of Standard Software

Although we used SAS/IML for our calculations, the method is easy to implement with publicly available software for Procrustes analysis, such as GRF-ND (Slice 1994), tpsRelw (Rohlf 1997), or Morphometrika (Walker 1997), and a standard statistics package or even a spreadsheet program. The following steps constitute the alternative protocol for users of these programs: (1) Arrange all coordinate data (including all replicates) in a single data matrix. Reverse the sign of the x coordinate for all configurations of either the left or the right body side to reflect them to mirror images (step 1 in Fig. 2). (2) Using morphometrics software, perform a Procrustes analysis using the generalized least squares method (GLS) for optimal fitting of specimens. This corresponds to steps 2–4 in Figure 2. The output of the Procrustes procedure contains the centroid sizes of all specimens (for a conventional ANOVA according to Palmer and Strobeck 1986; Palmer 1994) and the coordinates of superimposed landmark configurations for use in subsequent analyses (Procrustes residuals can be used equivalently). (3) For each of the x and y coordinates of the aligned configurations, separately run a two-factor ANOVA following Palmer and Strobeck (1986) or Palmer (1994), with body sides as a fixed effect and individuals as a random effect. (4) Add the sums of squares for each of the effects (sides, individuals, side × individual interaction, and error) across x and y coordinates of all landmarks; the resulting sums are the Procrustes sums of squares used in this study. (5) To calculate degrees of freedom for the Procrustes ANOVA, multiply the degrees of freedom for each of the effects (from the output of the statistics program) by twice the number of landmarks minus four. (6) Compute mean squares for each effect as the Procrustes sum of squares (from step 4) divided by the degrees of freedom. These mean squares, with the appropriate degrees of freedom (step 5), can be used for parametric testing using F ratios.

This analysis quantifies the shape variation and asymmetry for a sample of specimens and provides statistical tests for fluctuating and directional asymmetry.

A Measure of Overall Shape Asymmetry

In a variety of other contexts, investigators require a single estimate of overall asymmetry for every individual, for example, to relate it to heterozygosity or measures of stress resistance. A convenient measure of the amount of overall shape asymmetry is the Procrustes distance between left and right sides (Bookstein 1991; Smith et al. 1997). To compute this distance, subtract the Procrustes-aligned x and y coordinates of the landmark configuration for the right side from the corresponding coordinates for the left side, sum the squared differences, and calculate the square root of the resulting sum. (Actually, this is an approximation of Procrustes distance, but it will only deviate noticeably if shape asymmetry is extremely large.) This distance measure is similar to the absolute difference between left and right sides for a linear measurement, $|R - L|$ (see Palmer and Strobeck 1986; Palmer 1994), but it accounts for overall size because of the initial standardization to unit centroid size.

Alternatively, a measure for nondirectional asymmetry can be derived by subtracting the component for directional asymmetry (Smith et al. 1997): Subtract the sample mean of the left-right coordinate differences from the left-right coordinate differences of the specimen, then square the resulting differences, sum across x and y coordinates of all landmarks, and calculate the square root. This measure of asymmetry does not have an equivalent in the reviews by Palmer and Strobeck (1986) or Palmer (1994), but derives directly from their two-factor ANOVA. Whether this is a proper measure of fluctuating asymmetry is contentious (CPK and GSM have opposite opinions on this point); it depends on whether fluctuating asymmetry is defined as deviations from "ideal" or "perfect" symmetry (e.g., Palmer 1996) or from the "norm" of average (directional) asymmetry as implied by the factorial design of the ANOVA.

This measure of overall asymmetry of individuals can be used in further analyses, for example, to explore its correlation with measures of heterozygosity, stress, or performance. For comparisons of fluctuating asymmetry among groups, the measure can be entered as the dependent variable into a one-way ANOVA (with the usual degrees of freedom, because it is truly a univariate analysis). This is analogous to a comparison of asymmetry levels by Levene's test, which Palmer (1994) recommended for univariate studies. To avoid assumptions about the distribution of Procrustes differences, we recommend the use of permutation methods for significance tests.