Shape Asymmetry and Developmental Stability

Jean-Christophe Auffray[1], Vincent Debat[1] and Paul Alibert[2]

[1] *Institut des Sciences de l'Evolution, UMR 5554 CNRS, CC064, Université Montpellier 2, 34095 Montpellier cedex 05, France*
[2] *Biogéosciences-Dijon, UMR 5561 CNRS, Bâtiment Gabriel, Université de Bourgogne, 6 Bd Gabriel, 21 000 Dijon, France*

16.1 Introduction

Waddington (1957) defined developmental homeostasis as the mechanisms responsible for ensuring phenotypic constancy in organisms despite the great variability of genetic, environmental and developmental features. Developmental homeostasis is considered as a combination of two components: developmental canalisation and developmental stability (Zhakarov, 1989). Developmental canalisation corresponds to the processes that buffer development to produce consistent phenotypes under a range of environmental and genetic conditions. Developmental stability, on the other hand, is defined as the ability of organisms to withstand genetic or environmental disturbances during development, so as to produce a predetermined phenotype (Zakharov, 1989). In other words, if canalisation ensures the constancy of phenotypes for different genotypes under varying environmental conditions, developmental stability enhances this constancy for a given genotype and environment.

Inasmuch as developmental stability reflects the capacity of organisms to produce an optimum phenotype despite perturbations encountered during development, its appraisal may thus be used to evaluate these stresses, as well as the ability of genotypes to correct them. The stresses can be of environmental or genetic origin. A large number of environmental stresses – food deprivation, temperature, pollution, and so on – have been shown to significantly impair the developmental stability of organisms (Parsons, 1990; Møller and Swaddle, 1997). Although the genetic basis of developmental stability remains mostly unknown, it is presumed to depend on certain genetic states, including genomic coadaptation and heterozygosity (Graham, 1992; Clarke, 1993). Any cause, which impairs one of these two conditions may thus be considered as a genetic stress such as cases (1) of inbreeding, which is expected to be accompanied both by a

On Growth and Form: Spatio-temporal Pattern Formation in Biology.
Editor Mark A.J. Chaplain, G.D. Singh, J.C. McLachlan

reduction of heterozygosity, and a higher fixation rate of deleterious alleles (Clarke, 1993), (2) of incorporation of new alleles (e.g. an insecticide resistant allele) which may disrupt genomic co-adaptation (McKenzie and Clarke, 1988) as well as (3) most of the cases of hybridisation (Graham, 1992). These properties have led developmental stability to be the focus of an increasing number of studies during the last decade in numerous fields of evolutionary biology and ecology.

Developmental stability is most often appraised by the study of fluctuating asymmetry (FA). Given that the two sides of a bilaterally symmetrical trait are produced by the same gene complexes and develop under similar environmental conditions, any deviation from symmetry will express the inability of the organism to correct developmental errors, i.e. its level of developmental stability (Møller and Swaddle, 1997). Therefore, developmental stability is widely estimated by a measure of the variability, within a population, of the differences between the right and left sides of organisms, the so-called FA. For a single metric trait, FA can be assessed by the variance of the left-minus-right ($L - R$) distribution in a population, as well as by any other expression of this variability (Palmer, 1994).

16.2 Nature of Asymmetries

In bilateral organisms, however, some asymmetries are adaptative and hence thought to be genetically determined, rendering them uninformative for estimating developmental noise. Thus, fluctuating asymmetry needs to be distinguished from two other types of asymmetry: directional asymmetry (DA) and antisymmetry (AA). DA occurs when one side of a bilateral character is systematically larger than the other, so the mean of the ($L - R$) normal distribution of the population will be different from zero. The AA also corresponds to a systematic deviation from symmetry, but in this case the side that is larger varies at random among individuals. A typical antisymmetric trait leads to a bimodal ($L - R$) distribution centered on zero.

Numerous illustrative cases of DA are found in bilateral organisms, e.g. internal organs in mammals. On the other hand, typical antisymmetric traits are more rarely observed; the claws of the male fiddler crabs (genus Uca) are the most frequently cited example – one of them is always much larger than the other. In most cases, however, DA and AA may be subtle. A trait can be characterised by a slight but significant DA, and any normal-like distribution of the ($L - R$) may hide a slight but true AA (Palmer and Strobeck, 1992). A prior assessment of DA and AA is crucial in studies of fluctuating asymmetry, not only for the biological information it provides but also because several indices available for estimating fluctuating asymmetry are potentially biaised by DA, e.g. mean of the $|R - L|$ distributions (Palmer, 1994). If the purely statistical bias is more or less easily resolved (Palmer, 1994; Rowe *et al.*, 1997; Graham *et al.*, 1998), the meaning of DA or AA in terms of developmental stability is far from trivial. Several studies have persuasively suggested that FA, DA and AA are dynamically interrelated (Graham *et al.*, 1993, 1998; Leamy, 1999). In wild house mice, measurements (maximum length and width) of the three lower molars are often used in studies of asymmetry. While five of these six highly related traits do not exhibit any DA, several independent studies have shown that the width of the third lower molar presents significant DA, the right third molar being larger (Alibert *et al.*, 1997; Chatti *et al.*, 1999). However, it should not be concluded that one specific trait

will systematically exhibit a constant DA across populations. Studies on mandible mensurations in a house mouse strain (CV1) have demonstrated that several traits displayed significant DA, but not always towards the same sides (Leamy, 1993, 1999). Additionally, while DA is thought to be heritable and FA to be an expression of developmental noise, heritability estimates of DA have been shown to be low, and only slightly higher than those reported for FA (Leamy *et al.*, 1997; Leamy, 1999). These results have led Leamy (1999) to stress the relevance of investigating DA as a marker of developmental stability. Formerly, the emergence of DA in developmentally unstable organisms during selective experiments (Graham *et al.*, 1993) emphasised the potential of DA in expressing developmental instability. In the same manner, several reports of transition from FA to AA have suggested an equivalent potential of AA (McKenzie and Yen, 1995; Leary and Allendorf, 1989). Although, Rowe *et al.* (1997) have demonstrated that several emblematic studies of FA were actually dealing with AA, several studies have more directly reported the relevance of antisymmetric signals in terms of developmental instability (McKenzie and Clarke, 1988; Rowe *et al.*, 1997).

In summary, despite the fact that the nature of DA and AA, as well as their meaning in terms of developmental stability, are increasingly debated, these types of asymmetry are still believed to present a certain genetic component. Therefore, their statistical significance has still to be carefully assessed to avoid confounding effects. The recent studies of Graham *et al.* (1998) and Rowe *et al.* (1997) respectively reconsider the statistical distinction between FA and DA and between FA and AA.

16.3 Nature of Characters

The influence of the number and the nature of characters used to appraise FA has to be considered. As a matter of fact, there is no general rule in the choice of these characters, unless it is guided by the aim of the study, such as secondary sexual characters for sexual selection studies. The influence of the choice of a peculiar set of traits in FA studies strongly depends on the congruence of these characters in revealing developmental perturbations. This question may be adressed at an individual or a population level.

Clarke (1998a) pointed out that when several characters are considered among individuals, a non-random developmental noise among these traits would lead to a significant concordance of individual asymmetry. This would then allow the estimation of an individual asymmetry parameter (IAP) over all characters. However, it is noteworthy that such a parameter is most often undetected (Clarke, 1998a; see table W2.1 at `http://www1.oup.co.uk/MS-asymmetry/`; Møller and Swaddle, 1997). Several reasons could account for this. A first reason for not detecting an IAP, which is currently assessed by correlative procedures between individual asymmetries exhibited by several characters, would be related to statistical bias due to sample sizes and to low repeatability of measurements (Whitlock, 1996). Hence, it has been reported that an individual asymmetry parameter would be conveniently detected if high sample sizes were used (Leamy, 1993).

Another reason would lie with the independence of developmental processes and control of developmental stability of the characters considered. As already suggested in the literature, differences in developmental time, morphogenetic or developmental independence between characters as well as differences in the developmental control

systems may also explain the absence of concordance between individual asymmetries (Møller and Swaddle, 1997; Clarke, 1998a). Such hypotheses could be more conveniently tested if one used the concordance of characters among samples to express the levels of developmental instability. In other words, if a population exhibits the highest level of FA for one character, this population is expected to present the highest levels of FA for the other characters. If this concordance of characters is verified, it would allow the establishment of a population asymmetry parameter (PAP, Soulé, 1967), i.e. a synthetic estimation of FA among populations across characters.

Møller and Swaddle (1997) reported that PAP estimates were often significant (for 21 out of 30 studies reported in table 2.1 at the URL cited above). In contrast, Clarke (1998a) has recently shown that, in most of the cases, the characters were not concordant in providing the correct ranking of FA across samples within several invertebrate species. In his review the studies involved three to six characters which were not closely related in terms of development. The biological interpretation of the presence or absence of such a PAP is undoubtely the same as for the IAP. Yet it has been shown that the PAPs were more significant when their estimates, instead of being based on the whole set of characters of a morphological structure, were based on several sets of morphogenetic units (Leamy, 1993). For example, in rodent teeth, i.e length and width of lower and upper molars, this concordance is verified most of the time (Alibert *et al.*, 1994, 1997; Auffray *et al.*, 1999; Chatti *et al.*, 1999). It is likely that teeth share a common and synchronous developmental process and potentially a common mechanism for controlling stability. In contrast, among more independent morphological structures, these processes could differ (Møller and Swaddle, 1997), leading to insignificant PAPs. Moreover, each case in which the concordance of characters was not verified, corresponded to sets of populations which did not exhibit significant differences in FA levels for any of the traits considered (Auffray *et al.*, 1996b; Fontanillas and Auffray, unpublished). Clarke (1998a) reported a similar pattern, i.e. significant PAPs arising only when FA levels were significantly different, trait by trait.

However, the concordance of characters in ranking different populations according to their level of developmental stability and leading to a significant PAP is most often reported in the literature. This suggests a genome-wide process of control of asymmetry, but the respective amounts of FA displayed by the different characters might differ and thus their sensitivity to developmental stability.

16.4 Developmental Stability and Canalisation

Palmer (1994) stressed the fact that inferring differences in developmental stability among samples required the use of developmentally independent characters. However, even developmentally related characters, which are expected to be congruent in providing the rank of FA levels among samples, may considerably differ in their sensitivity to developmental stability. Any researcher in the field of developmental stability has experienced the heterogeneity in the amplitude of response of characters to developmental perturbations. A clear example was provided by the mole rat (*Spalax ehrenbergi*) populations of the Near East (Auffray *et al.*, 1999). Among eight tooth traits (six for molars, two for incisors) which were concordant in providing ranks, the transversal diameter of the incisor clearly exhibited one of the lowest levels of FA

whereas the other incisor parameter, the anteroposterior diameter, presented some the highest. In these populations the level of FA was related to the mean size of characters. Whitlock (1996) stressed the necessity to use the coefficient of variation of FA to compare levels of FA among traits. In the case of the mole rat, while the FA estimates strongly differed among traits, their respective coefficients of variation were more homogeneous (between 1.28% and 1.71%) suggesting that the amount of FA partly depended on character size.

Most of the attention should, however, be given to the relationship between FA levels and variances among traits. The phenotypic variance of a character in a population arises from several components: (1) the genetic variance, which is related to differences in genetic constitution, (2) the environmental variance resulting from exposure to different environmental conditions, (3) an interaction term between these two factors and (4) a stochastic part due to the developmental noise. It is well known that all characters are not similarly sensitive to genetic or environmental conditions. The less variable they are, the more they are canalised. The genetic control of canalisation is poorly known, but it has clearly been established that a high level of canalisation for a character can result from selective processes (Scharloo, 1991; Møller and Swaddle, 1997).

Waddington (1957) has suggested that characters for which phenotypic constancy is important in terms of individual fitness will be developmentally less variable (better buffered) than characters for which constancy is less important. He has thus proposed a form of stabilising selection, called canalising selection. Canalising selection acts to eliminate from the population those genotypes that render developmental pathways for a given character sensitive to genetic or environmental variation. This suggests that the degree of developmental homeostasis is character-dependent and is hypothesised to reflect the functional contribution of the character to the fitness of the organism. In the same manner it also implies that the genetic mechanism of developmental homeostasis is character-dependent. Such assumptions, however, still remain to be tested rigorously (Clarke, 1998b).

Waddington (1957) also presumed that the two components of developmental homeostasis, canalisation and developmental stability, have phenotypical and genetic independent properties. Although the genetic control of these two mechanisms is mostly unknown, this idea has been widely accepted. A simple way, however, to assess the level of dependence between canalisation and developmental stability is to appraise the relationship between morphological variability and FA. For example, Clarke (1998b) has recently shown that, in some insect and crustacean species, FA levels exhibited by characters were usually correlated with their variability estimated by the coefficient of variation.

However, in order to study the relative levels of phenotypic variance of characters, and their covariance with FA levels, and to potentially relate these patterns with the functional importance of characters, a high number of characters of presumably different functional importance is required. Given that, in studies of asymmetry, both sides of a single character have to be measured with several replicates on sufficiently high sample sizes, the number of characters studied is rarely very large. The necessity to increase the number of characters would thus undoubtely discourage this kind of approach. What was required was a method which made it possible to appreciate the whole morphology as well as to estimate FA, while preserving reliable information on

the covariation among traits.

Such an opportunity was provided by geometric morphometrics based on landmark data, which are suitable to precisely describe and quantify the covariation among components of a morphological structure (Bookstein, 1991; Rohlf and Marcus, 1993). In addition, geometric morphometrics can properly estimate fluctuating asymmetry using morphological structures as a whole.

16.5 Shape Asymmetry

The use of geometric morphometrics to appraise fluctuating asymmetry was initially proposed by Bookstein (1991). Geometric morphometric methods are based on landmark data for the study of shape variation (Rohlf and Marcus, 1993; Dryden and Mardia, 1998). In this case, instead of being represented by several pairs of measurements, the right and the left sides of an individual are represented by their configuration of landmarks in two or three dimensions.

The estimation of individual asymmetry is based on the calculation of Procrustes distances between the configuration of one side and the mirror image of the other. Procrustes methods, first introduced in 1970 by Schanemann and Gower (Bookstein, 1991), are based on the least-squares technique in order to compute the best-fitting superimposition of landmark configurations (Rohlf, 1990). The mathematical and computational procedures of Procrustes least-squares superimposition are detailed in Rohlf (1990) and in Rohlf and Slice (1990). Briefly, least-squares superimposition of two configurations (Figure 16.1), e.g. one side and the mirror image of the other, requires both to be conventionally scaled to unit centroid size, i.e. the root of the sum of the squared distances of all landmarks to the centroid of their own configuration will be equal to unity. Once the configurations are scaled, the best-fit superimposition of the two configurations involves the translation of the configurations in order to superimpose their centroids, and the rotation of one of them into a position of best-fit to the other one. The optimal superimposition is obtained by minimising the sum of squared distances (A^2) between the corresponding landmarks of the two configurations.

Bookstein (1991) as well as Smith *et al.* (1997) depicted the geometric meaning of this estimate of asymmetry: it is proportional to the sum of the areas of circles centred on each landmark, of which the radii are the vectors connecting corresponding landmarks. The square root of A^2, A, is thus a measure of net (or raw) asymmetry between the two sides of an organism. An analogy can be made between this measure of asymmetry and the $|R - L|$ parameter used in the traditional calculation of the FA1 index of Palmer (1994). Both are unsigned and when these individual asymmetry values are averaged over a sample to provide a mean estimate of asymmetry, both can be very biased by DA or AA. Thus, the mean Procrustes distance between the two sides in a sample, as well as $\overline{|R - L|}$ in traditional approches of FA, can be considered as relevant measures of fluctuating asymmetry only if DA or AA is absent (Palmer, 1994).

Here the directional asymmetry is estimated in a sample by the Procrustes distance between the mean right and the mean left configurations after their superimposion. These two mean configurations are computed by a generalised least squares (GLS) superimposition which is a generalisation over a sample of the two-configuration least-

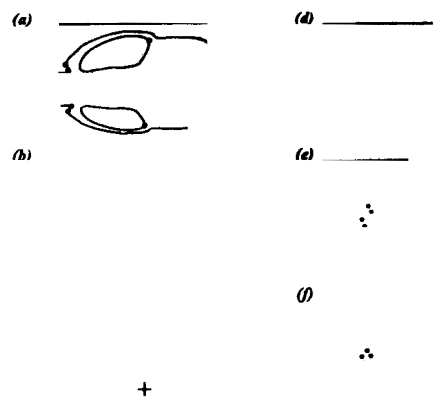

Figure 16.1 Two-dimensional representation of the Procrustes least-squares superimposition of the two sides of a mouse skull: (a) location of digitised landmarks on the two sides of the skull (landmarks on each side have a different color); (b) representation of the raw configurations of landmarks; (c) reflection of one of the sides and location of the centroid for each side (+); (d) scaling configurations to equal centroid size (in practice both configurations are scaled to unit centroid size); (e) translation of configurations in order to superimpose their centroids; (f) rotation of one configuration using the criterion of minimum squared distances between corresponding landmarks on both sides; the rotation angle is shown by the + signs (Bookstein, 1991; Auffray et al., 1996b).

square superimposition. Computational details can be found in Rohlf (1990) or Rohlf and Slice (1990). This procedure minimizes the sum of Procrustes distances from all individuals to the Procrustes mean configuration. GLS superimposition is an iterative procedure which begins by fitting all configurations to one of them used as a reference. Once all configurations are superimposed according to the criterion of minimizing their Procrustes distances to this reference, a mean configuration is calculated onto which all configurations are re-superimposed. A new mean configuration is then calculated and so on. The procedure converges rapidly to a stable mean configuration (Rohlf, 1990; Slice, 1996).

In Bookstein's (1991) studies and Auffray *et al.*'s (1996b) studies of Procrustean asymmetry, GLS superimposition was only used to estimate the mean configuration of each side, and individual asymmetries were computed one by one. Smith *et al.* (1997) simplified the procedure by superimposing both sides of all individuals using GLS superimposition. This procedure, which does not modify the Procrustes distance between sides of individuals, allows one to use the new coordinates of superimposed configurations to estimate individual asymmetry as well as directional asymmetry. At this step, it should be noted that in the statistical techniques reported here, the use of coordinates of superimposed configurations is strictly similar to that of residuals around each mean landmark. Once all sides are superimposed by GLS (Figure 16.2a), the whole variability at each landmark is expressed by the scatter of points around the mean landmark. Within the scatter around a given landmark k, each individual is represented by two points I_{kr} and I_{kl}, corresponding to the kth landmark of the

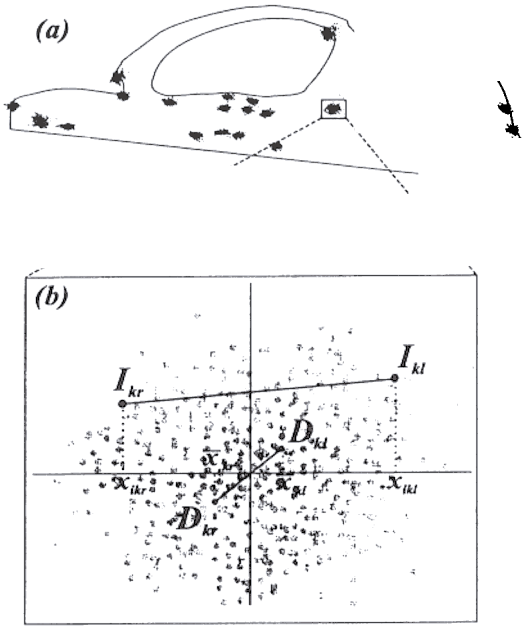

Figure 16.2 Procrustean estimate of shape asymmetry according to the procedure of Smith *et al.* (1997): (a) generalised least squares (GLS) superimposition of configurations for both sides in a sample; (b) basic design of the partition of raw asymmetry into its directional and fluctuating terms shown for one coordinate of a given landmark: I_{kr} and I_{kl} respectively correspond to the kth landmark of the right and left configuration of individual i; D_{kr} and D_{kl} respectively represent the kth mean landmark for the right and left sides in the sample (see text).

right side configuration and of the left one respectively (Figure 16.2b). The mean right point D_{kr} and the mean left one D_{kl} are computed and correspond respectively to the kth landmark of mean right and mean left configurations. If we consider $x_{ikr}, x_{ikl}, \overline{x}_{.kr}$ and $\overline{x}_{.kl}$ as the projections of respectively I_{kr}, I_{kl}, D_{kr} and D_{kl} onto the x-axis, the distances between I_{kr} and I_{kl} projected onto this axis can be expressed as

$$\underset{(1)}{(x_{ikr} - x_{ikl})} = \underset{(2)}{(\overline{x}_{.kr} - \overline{x}_{.kl})} + \underset{(3)}{(x_{ikr} - \overline{x}_{.kr})} + \underset{(4)}{(\overline{x}_{.kl} - x_{ikl})} \qquad (16.1)$$

where (2) corresponds to the difference after projection, between the kth landmarks of the right and left mean configurations, and (3) and (4) correspond to the deviation of the landmarks of each side to their respective mean side landmark over the sample. Similar distances can be computed on the y-axis, using the y-coordinates for the same landmarks.

Computing the distances between I_{kr} and I_{kl} in two dimensions requires that terms be squared and summed. Thus, for landmark k, the square distance between I_{kr} and I_{kl}, $d^2(I_{kr}, I_{kl})$ is

$$d^2(I_{kr}, I_{kl}) = d^2(D_{kr}, D_{kl}) + F_{ik} \tag{16.2}$$

where F_{ik} corresponds to the expression of the distance between the k^{th} landmark of both sides around the norm. The norm is distinguished from perfect symmetry in that it integrates the expected mean difference between sides due to DA.

Summing the terms across all landmarks for an individual i corresponds exactly to the calculation of the square Procrustes distance, A_i^2, between the configurations of its two sides. It is expressed by

$$A_i^2 = Ad^2 + F_i,$$

where Ad^2 corresponds to the Procrustes distances between the mean right and the mean left configurations, in other words the squared DA, and F_i is an expression of the distance between the two sides around the norm.

Summing A_i^2 across the n individuals in a sample and averaging them provides the mean raw asymmetry $\overline{A^2}$ of the sample as

$$\overline{A^2} = Ad^2 + \overline{F}$$

where $\overline{F}$ corresponds to the fluctuating component of the mean square raw asymmetry averaged over the sample.

Testing differences of mean raw asymmetry among samples can be done by a one-way ANOVA on the root of the square Procrustes distances, i.e. A_i (Smith *et al.*, 1997; Klingenberg and McIntyre, 1998). As such an index of asymmetry can be very biased by DA, it is necessary to assess the absence of DA in the samples considered. On a single sample, the significance of DA can be appraised by testing the equality of the mean right and mean left shapes. For reasons related to the geometric procedure of Procrustes superimposition, and to the loss of degrees of freedom (Bookstein, 1996, Dryden and Mardia, 1998), the use of general linear models such as MANOVAs on Procrustes residuals is allowed only when variation is small and the number of individuals is high with regard to the number of landmarks. If these conditions are fulfilled, a one-way MANOVA on all coordinates with the factor side can be used to appraise the equality of mean side shapes. If not, non-parametric tests such as permutation tests should be used (Dryden and Mardia, 1998).

Since 1991, when this method was first proposed by Bookstein, it has gathered few followers. So far the studies of Auffray *et al.* (1996b) and Smith *et al.* (1997) are the only published works using this method, and Smith *et al.* (1997) is a re-examination of the study reported by Bookstein in 1991. These studies have yielded relevant biological signals, but as is often the case for newly arising methods, they were unsatisfactory from several points of view. None of them took into account the size asymmetry. However, although each configuration had been scaled to a unit centroid size (see above), the size component of FA could have been assessed from the raw centroid sizes as for any traditional trait (Palmer, 1994). Another problem lay with appraising the measurement error, which was absolutely required since FA results from subtle differences between sides. Auffray *et al.* (1996b) introduced a procedure to control the level of the error of measurement by estimating Procrustes distances between the two replicates of left configurations in a sample. However, this procedure did not yield an

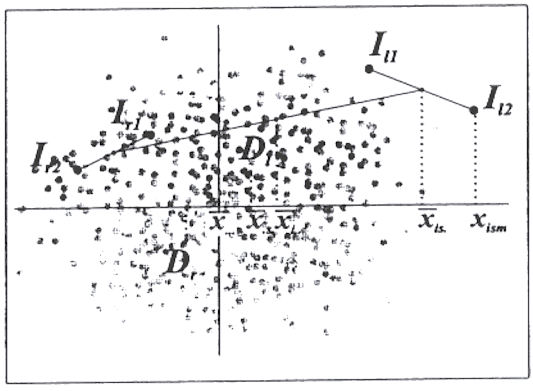

Figure 16.3 Basic design of the partition of whole shape variability according to several effects following the procedure of Klingenberg and McIntyre (1998), shown for one coordinate at a given landmark. $I_{r1}, I_{r2}, I_{l1}, I_{l2}$ respectively correspond to two replicates of the right and left sides of an individual i for the same landmark, D_r and D_l respectively depict the mean landmark for the right and left sides over the sample (see text).

estimate of the fluctuating asymmetry index from which the error effect was removed, just as for the FA10 index of traditional approaches (Palmer, 1994).

Such an estimate of fluctuating asymmetry using geometric morphometrics has been recently proposed by Klingenberg and McIntyre (1998). Coordinate by coordinate, the basic design of this procedure differs from that of Smith *et al.* (1997) in that, instead of partitioning the total distance between the two sides into directional and fluctuating components, it partitions the deviation of each replicate to the grand mean according to several potential sources of variation, i.e. side, genotype (or individual) as well as their interaction and error. By analogy with the two-way mixed-model ANOVA proposed by Palmer and Strobeck (1986) and Palmer (1994) to test the significance of FA relative to the measurement error, this approach can thus involve several replicates of both sides. The generalisation of the basic partition design, by considering the sum of squares at a landmark or at the whole configuration, provides a means to appraise the significance of these effects as sources of variation.

In practice each side of an individual is represented by the number of its digitised replicates. By analogy with Figure 16.2, the basic partition design of the Procrustes residual used in the approach of Klingenberg and McIntyre (1998) is represented in Figure 16.3.

As for any two-way ANOVA, it can be stated that, in terms of the projection on the x-axis of the scatter of points around a given landmark, the deviation, in equation (16.5), of one of the replicates m of side s of individual i, x_{ism}, from the grand mean $\bar{x}_{...}$, can be partitioned as

$$\underset{(1)}{(x_{ism} - \bar{x}_{...})} = \underset{(2)}{(\bar{x}_{.s.} - \bar{x}_{...})} + \underset{(3)}{(\bar{x}_{i..} - \bar{x}_{...})} + \underset{(4)}{(x_{ism} - \bar{x}_{is.})} + \underset{(5)}{F_{ism}} \qquad (16.5)$$

where (2) is the distance between the mean side over replicates and individuals, $\bar{x}_{.s.}$ and

the grand mean, (3) the distance between the mean individual over sides and replicates, $\bar{x}_{i..}$ and the grand mean, (4) the residual error between the replicate considered and the mean side s of the individual i over replicates, $\bar{x}_{is.}$, and (5) an interaction term F_{ism} between the two factors such as

$$F_{ism} = (\bar{x}_{is.} - \bar{x}_{i..} - \bar{x}_{.s.} + \bar{x}_{...} \tag{16.6}$$

For a given coordinate, this expression is squared and summed over all replicates, sides and individuals, leading to a classical ANOVA sum of squares. As stated already, the main interest of this procedure is to partition morphological variability according to several sources of variation: (1) the variation among individuals, (2) the variation related to side, (3) the added random variance component due to non-directional asymmetry, which reflects the variation in the between-sides difference among individuals such as FA, AS or error, and (4) the variance due to the error of measurement. This partition of the whole variability at one landmark coordinate is strictly similar to the one performed in the two-way mixed-model ANOVA used for the estimation of FA10 (Palmer, 1994).

Summing mean squares for each effect accross the x- and y-coordinates for a given landmark could provide a relevant picture of the sensitivity of each landmark to the sources of variation considered. However, applying F-ratios of ANOVAs at this step would yield only scarce information on the significance of the different sources of variation. Landmark by landmark, these residuals are not independent (Bookstein, 1991; Auffray *et al.*, 1996a). If one landmark is more variable than the other ones, i.e. if it represents a very localised variability, the GLS superimposition will dilute this localised variability over all landmarks and lead to its underestimation, as noted by Klingenberg and McIntyre (1998).

To take into account the variability of the whole morphological structure over a sample, sums of squares are added for each of the effects across the x- and y-coordinates of all landmarks. These Procrustes sums of squares are then used to compute mean squares for each effect by dividing them by the relevant degrees of freedom, which are the conventional degrees of freedom (Palmer, 1994) multiplied by the number of coordinates minus 4. Four degrees of freedom are lost during the GLS superimposition, two for translation, one for scaling and one for rotation (Bookstein, 1991). F-tests can then be applied on these mean squares. However, as for any inferences following Procrustes superimposition (see above), non-parametric tests such as permutation tests are recommended for all statistics dealing with Procrustes coordinates or residuals.

In traditional FA as well as in the Procrustes approach, the significance of the third term indicates that the interaction variance is greater than the error of measurement. In other words, it indicates that the differences between sides among individuals vary more than would be expected given the size of the measurement error (Palmer, 1994). If this term is significant, it is possible to extract the measurement error variance from the between-side variance in order to get a fluctuating asymmetry index from which the effect of measurement error has been removed. The shape fluctuating asymmetry index can then be computed in a similar way to the traditional FA10 of Palmer (1994). A corresponding size FA10 index for size can also be computed from the centroid size of each replicate. In this case, the significance of the ANOVA effects can be classically

assessed by F-ratios, i.e. using parametric statistics.

More interestingly, Klingenberg and Mcintyre (1998) developed a method to assess and depict the morphological variability related to each effect as well as their relationships. Their procedure involved performing principal components analyses (PCAs) on different covariance matrices computed from the whole Procrustes coordinates or the residual matrix, but corresponding to each potential source of variation. The covariation of landmarks in relation to the inter-individual source of variation was then appraised by a PCA based on a covariance matrix computed from the individual means, FA from individual left–right differences, and the error of measurement from the residual of the replicate measurements over the mean of each side.

Furthermore, the computation of eigenvectors defining the principal components emerging from these analyses allows one to depict, at each landmark of the original mean configuration, the vectors representing the variability expressed at a given principal component for any of the effects considered. At this step, permutation tests are used to establish the independence between eigenvectors, or to appraise the agreement between the morphological expression of developmental instability (FA), the interindividual variability and the error of measurement.

In their study, Klingenberg and McIntyre (1998) provided a detailed example, in which they focused on the similarity of covariation of landmarks for different sources of variation (interindividual variability, FA and error of measurement) rather than on the comparison of FA levels among groups. The significant correlation between these three covariance matrices, as well as the non-independence between emerging PCs and their highly similar pictorial patterns, led the authors to consider that there was no reason to suspect that specific developmental processes were affecting asymmetry or buffering against it, other than those influencing the mean shape across body sides. They also admitted that, since this study was the first one, their conclusions could not be considered as a general rule.

16.6 Towards the Form Asymmetry

Further uses and developments of geometric morphometric methods in the field of developmental stability would undoubtely be extremely informative about the processes governing developmental stability as well as those patterning the emergence of FA. This would be the case in particular for the relationship between size, shape, and their respective levels of asymmetry. But even though examination of the relationships between the centroid size and the shape or size asymmetry has the same interest as in traditional approaches to developmental stability (Palmer and Strobeck, 1986; Palmer, 1994; Rowe *et al.*, 1997; Graham *et al.*, 1998), assumptions about the patterns of the relationship between size and shape asymmetries are far from being stated. One can, however, hypothesise that an independence between these asymmetries as well as a negative relationship would provide interesting clues to numerous unanswered questions such as the relative functional importance of size and shape, or the potential complementing effects between size and shape components in the maintenance of symmetry. Our experience on several independent samples of the house mouse skull has indicated that shape and size asymmetries are positively correlated, both among individuals and among samples (Debat and Auffray, unpublished). This leads us to

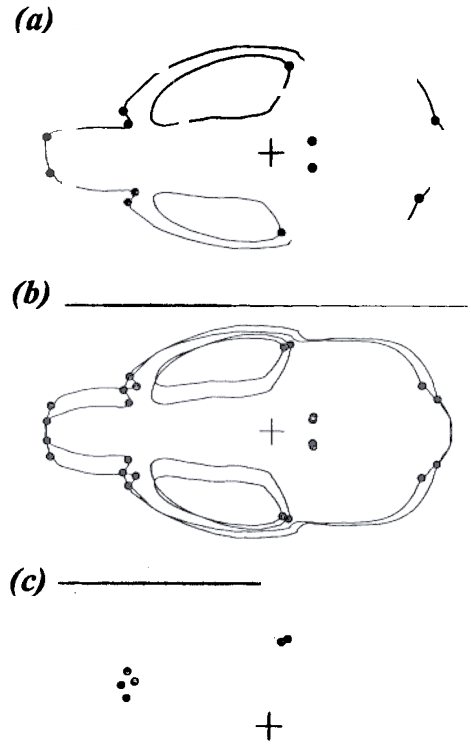

Figure 16.4 Superimposition procedure of a whole symmetric structure in order to estimate form asymmetry: (a) location of landmarks digitised on the two sides (the landmarks on each side have a different color); (b) least-squares superimposition of the whole configuration onto its mirrored image; (c) description on one side of the form distances between sides; since both sides are not scaled to each other, the raw asymmetry shown here corresponds to raw asymmetry of form.

address the question of the potential weakening effect of partitioning of FA signals into size and shape terms on the statistical appraisal of overall FA. As size and shape are classically conceived to be components of form, it would be possible to compute a form asymmetry index combining both size and shape asymmetries. This would, however, require the use of single symmetric morphological structures, i.e. containing a midline (e.g. a skull) instead of two disconnected ones around an axis of symmetry (e.g. wings). Instead of superimposing both sides of an organism, which involves scaling each side to a unit centroid size, the GLS superimposition of whole morphological structures (both connected sides) onto their mirror images would preserve the size of both sides relative to each other (Figure 16.4). This procedure would lead to a perfectly symmetric distribution of residuals (or coordinates) for both sides around the midline, which should not be confused with the axis of symmetry. Consequently, the residuals at each landmark of one side, and eventually those of the expected axis of symmetry, would contain all the information on the covariation of landmarks in terms of form. Except for the fact that in this case size asymmetry does not have to be considered, and that the degrees of freedom should be adequately modified, all the

procedures of Klingenberg and McIntyre (1998) would remain fully applicable.

Although studies in the field of developmental stability as well as geometric morphometrics have been steadily increasing for the last two decades, the use of the geometric morphometrics to infer FA levels has been rare until now. This was probably due to the fact that the use of Procrustes distances as estimates of FA had not been expressed within the general framework of statistical assessement of FA, mostly represented by "the primer" of Palmer (1994). The procedure of Klingenberg and McIntyre (1998) has not only conveniently done this, it has also extended the study of covariation of landmarks to much more integrative aspects of morphological variability, e.g. the relationships between the components of developmental homeostasis.

Acknowledgements

We thank Christian P. Klingenberg and Larry Leamy for helpful discussions during the preparation of the manuscript and for having sent preprints of their work in press; Janice Britton-Davidian for comments on an earlier version of the manuscript; Sandrine Peired and Josette Catalan for their help with the bibliography. This study was supported by the Programme CNRS Paléoenvironnement et Evolution des Hominidés: Stress génétiques et environnementaux subis par les populations actuelles et fossiles sous l'effet de l'anthropisation (coordinator Jean-Christophe Auffray). This is publication ISEM No. 99-134.

REFERENCES

Alibert, P., Renaud, S., Dod, B., Bonhomme, F. and Auffray, J.-C. (1994) Fluctuating asymmetry in the Mus musculus hybrid zone: a heterotic effect in disrupted co-adapted genomes. *Proc. Roy. Soc. Lond.* B258: 53–59.

Alibert, P., Fel-Clair, F., Manolakou, K., Britton-Davidian, J. and Auffray, J.-C. (1997) Developmental stability, fitness, and trait size in laboratory hybrids between European subspecies of the house mouse. *Evolution* 51: 1284–1295.

Auffray, J.-C., Alibert, P., Latieule, C. and Dod, B. (1996a) Relative warp analysis of skull shape across the hybrid zone of the house mouse (*Mus musculus*) in Denmark. *J. Zool.* 240: 441–455.

Auffray, J.-C., Alibert, P., Renaud, S., Orth, A. and Bonhomme, F. (1996b) Fluctuating asymmetry in *Mus musculus* subspecific hybridization: Traditional and Procrustes comparative approach. In: *Advances in Morphometrics*, Marcus L.F., Corti M., Loy A., Slice, D. and Naylor, G. eds. pp. 275–283. Plenum press, New York.

Auffray, J.-C., Renaud, S., Alibert, P. and Nevo, E. (1999) Developmental stability and adaptive radiation in *Spalax ehrenberghi* superspecies in the Near East. *J. Evol. Biol.* 12: 207–221.

Bookstein, F. L. (1991) *Morphometric Tools for Landmark Data: Geometry and Biology.* Cambridge University Press, New York.

Bookstein, F. L. (1996) Combining the tools of geometric morphometrics. In: *Advances in Morphometrics*, Marcus L.F., Corti, M., Loy A., Slice, D. and Naylor, G. eds. pp. 113–152. Plenum press, New York.

Chatti, N., Said, K., Catalan, J., Britton-Davidian, J. and Auffray, J.-C. (1999) Developmental instability in wild chromosomal hybrids of the house mouse. *Evolution* (in press).

Clarke, G.M. (1993) The genetic basis of developmental stability. I Relationships between stability, heterozygosity and genomic coadaptation. *Genetica* 89: 15–23.

Clarke, G.M. (1998a) The genetic basis of developmental stability. IV. Individual and

population asymmetry parameters. *Heredity* 80: 553–561.

Clarke, G.M. (1998b) The genetic basis of developmental stability. V. Inter- and intra-individual character variation. *Heredity* 80: 562–567.

Dryden, I.L. and Mardia, K.V. (1998) *Statistical Shape Analysis.* John Wiley, Chichester.

Graham, J.H. (1992) Genomic coadaptation and developmental stability in hybrid zones. *Acta Zoologica Fennica* 191: 121–131.

Graham, J.H., Freeman, D.C. and Emlen, J.M. (1993) Developmental stability: a sensitive indicator of populations under stress. In: *Environmental toxicology and risk assessment*, Landis, W.G., Hughes, J.S. and Lewis, M.A., eds., pp. 136–158. American Society for Testing and Materials, Philadelphia PA.

Graham, J.H., Emlen, J.M., Freeman, D.C., Leamy, L.J. and Kieser, J.A. (1998) Directional asymmetry and the measurement of developmental instability. *Biol. J. Linn. Soc.* 64: 1–16.

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

Leamy, L. (1993) Morphological integration of fluctuating asymmetry in the mouse mandible. *Genetica* 89: 139–154.

Leamy, L. (1999) Heritability of directional and fluctuating asymmetry for mandibular characters in random-bred mice. J. Evol. Biol. 12: 146–155.

Leamy, L., Routman, E.J., Cherevud, J.M. (1997) A search for quantitative trait loci affecting asymmetry of mandibular characters in mice. *Evolution* 51: 957–969.

Leary, R.F. and Allendorf, F.W. (1989) Fluctuating asymmetry as an indicator of stress: implication for conservation biology. *Trends in Ecology and Evolution* 47: 214–217.

McKenzie, J.A. and Clarke, G.M. (1988) Diazinon resistance, fluctuating asymmetry and fitness in the Australian sheep blowfly, *Lucilia cuprina. Genetics* 120: 213–220.

McKenzie, J.A. and Yen, J.L. (1995) Genotype, environment and the asymmetry phenotype. Dieldrin-resistance in *Lucilia cuprina* (the Australian sheep blowfly). *Heredity* 75: 213–220.

Møller, A.P. and Swaddle, J.P. (1997) *Asymmetry, Developmental Stability, and Evolution.* Oxford University Press, Oxford.

Palmer, A.R. (1994) Fluctuating asymmetry analysis: a primer. In: *Developmental Instability: Its Origin and Evolutionary Implications*, Markow T.A., ed., pp. 335–364. Kluwer Academic, Dordrecht.

Palmer, R.A. and Strobeck, C. (1986) Fluctuating asymmetry: measurement, analysis, patterns. *Annual Review of Ecology and Systematics* 17: 391–421.

Palmer, R.A. and Strobeck, C. (1992) Fluctuating asymmetry as a measure of developmental stability: implications of non-normal distributions and power of statistical tests. *Acta Zoologica Fennica* 191: 57–72.

Parsons, P.A. (1990) Fluctuating asymmetry: an epigenetic measure of stress. *Biol. Rev.* 63: 131–145.

Rohlf, F.J. (1990) Rotational fit (Procrustes) method. In: *Proceedings of the Michigan Morphometrics Workshop*, Rohlf, F.J. and Bookstein F.L., eds., pp. 227–236. University of Michigan Museum of Zoology, Ann Arbor MI.

Rohlf, F.J. and Marcus, L.F. (1993) A revolution in Morphometrics. *Trends in Ecology and Evolution* 8: 129–132.

Rohlf, F.J. and Slice, D. (1990) Extensions of the Procrustes methods for the optimal superposition of landmarks. *Systematic Zoology* 39: 40–59.

Rowe, L., Repasky, R.R. and Palmer, A.R. (1997) Size-dependent asymmetry: fluctuating asymmetry versus antisymmetry and its relevance to condition-dependent signaling. *Evolution* 51: 1401–1408.

Scharloo, W. (1991) Canalization: genetioc and developmental aspects. *Annual Review of Ecology and Systematics* 22: 65–93.

Slice, D. E. (1996) Three-dimensional generalized resistant fitting and the comparison of least-squares and resistant-fir residuals. In: *Advances in Morphometrics*, Marcus, L.F., Corti, M.,

Loy, A., Slice, D. and Naylor, G. eds., pp. 275–283. Plenum press, New York.
Smith, D.R., Crespi, B.J. and Bookstein, F.L. (1997) Fluctuating asymmetry in the honey bee, *Apis mellifera*: effects of ploidy and hybridization. *J. Evol. Biol.* 10: 551–574.
Soulé, M.E. (1967) Phenetics of natural populations. II Asymmetry and evolution in a lizard. *American Naturalist* 101: 141–160.
Waddington, C.H. (1957) *The Strategy of the Genes.* Macmillan, New York.
Whitlock, M. (1996) The heritability of fluctuating asymmetry and the genetic control of developmental stability. *Proc. Roy. Soc. London* B263: 849–854.
Zakharov, V.M. (1989) Future prospects for population phenogenetics. *Sov. Sci. Rev.* 4: 1–79.