

# Disentangling fluctuating asymmetry from otolith shape

M. Palmer\*, M. Linde, B. Morales-Nin

IMEDEA (CSIC-UIB), Instituto Mediterráneo de Estudios Avanzados, Miquel Marquès 21, 07190 Esporles (Mallorca), Spain

**ABSTRACT:** Fluctuating asymmetry (FA), defined as random deviations from ideal, perfect symmetry, has been hypothesized to be negatively associated with fitness and positively associated with stress. Fish otoliths deserve special attention in regard to this topic because their FA is relatively easy to measure, unlike the fitness of wild fish themselves. In addition, otoliths are convenient for estimating FA, because they are systematically extracted for aging purposes, and otolith collections might allow the use of FA as a bioindicator of change at different spatio-temporal scales. Here we demonstrate 2 key methodological tools: the use of shape instead of previously used univariate (metric or meristic) traits, and the use of a statistical protocol capable of disentangling true FA from other types of asymmetry (namely, directional asymmetry and antisymmetry) and from measurement error. In the case of the razor fish *Xyrichthys novacula*, fluctuation asymmetry of the otolith shape of 180 individuals from 5 populations accounted for 23.2% total shape variability (5 fold larger than the measurement error). Moreover, FA clearly differed between populations. In contrast, FA measured using a univariate trait (radius at the sulcus) did not detect between-population differences. Moreover, in a simulation experiment, we demonstrate that FA measured using shape is able to reveal a true pattern, while FA measured using a univariate trait extracted from the same data is unable to do so. Therefore, the analysis of FA using shape paired with an adequate statistical model may increase our knowledge of the potential links between otolith FA and fitness or stress.

**KEY WORDS:** Fluctuating asymmetry · Shape · Geometric morphometrics · Two-factor mixed effects model · Procrustes ANOVA · Directional asymmetry

Resale or republication not permitted without written consent of the publisher

## INTRODUCTION

Fluctuating asymmetry (FA), or small, random deviation in the development of both sides of a bilateral symmetric structure, has been extensively used as an indicator of stress and/or fitness (e.g. Leung & Forbes 1996 and references therein). The underlying hypothesis is that, under ideal conditions, development should produce nearly symmetric final forms (Palmer & Strobeck 1986). Thus, it is assumed that an organism that fares well is less susceptible to small random developmental errors. Developmental instability (DI) is defined as the inability of an organism to adjust its development to an ideal pattern. DI is not directly observable. One of the main reasons for studying FA relies on the hypothesis that FA may reflect DI (Palmer

& Strobeck 1986). Furthermore, it is hypothesized that FA should correlate positively with stress (i.e. disruptions of normal development due to genetic and/or environmental cues are expected to increase the asymmetry level) and negatively with fitness (i.e. well-faring organisms are expected to display both higher fitness and lower DI).

This appealing hypothetical framework has attracted much attention over the past few decades. However, despite much research, the observed patterns do not fully agree with the expected ones. For example, the contention that large values of FA should predict lower fitness has been inconsistently supported (Leung & Forbes 1996, Møller 1997, Clarke 1998). Similarly, the relationship between stress and FA is controversial (e.g. Bjorksten et al. 2000, Møller 2000, van Dongen &

\*Email: palmer@imedea.uib-csic.es

Lens 2000). Publication bias has been proposed as an additional source of uncertainty (Palmer 2000). Possibly because of this uncertainty, van Dongen (2006) reported a decreased interest in this topic in the last few years. Following van Dongen (2000), there is a growing awareness that FA might not be as straightforward a measure of stress as initially hypothesized. Nevertheless, van Dongen (2000) suggests that it is too early to abandon asymmetry as a potentially accurate estimator of stress. He calls for a better understanding of the underlying mechanisms that drive DI. This deficit in current understanding is part of a larger shortcoming with respect to our knowledge of the origin and control of morphological variability and the processes promoting and buffering such variability (modularity, integration, canalization, phenotypic plasticity and also developmental stability; Debat et al. 2008).

More closely related to the aim of this study, the necessity of correct statistical protocol is emphasized by van Dongen (2008). Some of the ambiguous results obtained previously may be due to improper statistical analyses. For example, most of the earlier papers reporting FA patterns did not include estimates of their measurement error (ME), so some variation attributed to FA in these studies may actually have been due to ME (Palmer 2000). This is of great importance, since it is expected that ME can be large compared with FA (Palmer & Strobeck 1986). This is the case if FA is estimated by the difference between a single measure of the left and the right otolith of the same individual. Alternatively, FA can be better estimated by the deviation of the measured width (as an example of any univariate trait) of the individual  $i$  (or even better, from the average of a number of replicated measurements) from the ideal width (Fig. 1). The ideal width of the right side of the individual  $i$  is estimated by combining the mean width of the right side across all individuals and the mean width of the 2 otoliths from the individual  $i$  (Fig. 1). This method of measuring FA is implicitly assumed by one of the statistical models used for analyzing FA (Leamy 1984, Palmer & Strobeck 1986, Klingenberg & McIntyre 1998, Klingenberg et al. 2002, van Dongen 2006, Debat et al. 2008). This model is able to disentangle FA from directional asymmetry (DA), which reflects a consistent bias of a character within a population toward asymmetric development (flatfishes are an extreme example). DA implies the existence of an ideal left side that is different from the ideal right side for all individuals in a population. Therefore, DA is not expected to reflect DI. Antisymmetry, a third type of asymmetry, is similar to DA, but the ideal left and right sides are reversed in some individuals (for example, the claws of some crabs). Therefore, antisymmetry is not expected to reflect DI either.

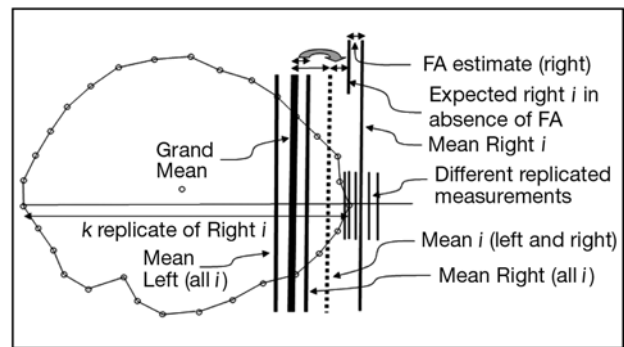


Fig. 1. Scheme of the statistical model used in a univariate (width) analysis of replicated measurements of the otoliths. The outline of a replicated image ( $k$ ) of the right otolith is shown. Each width measurement of the otolith is indicated by one of the thinner vertical lines. Grand mean is the mean of all widths. Mean right indicates the mean width of all right otoliths. Mean  $i$  indicates the mean of all replicates (left and right) from the individual  $i$ . Mean right  $i$  is the mean of all replicates of the right otolith of the individual  $i$ . The distance between Grand mean and Mean right is the side effect. The distance between Grand mean and Mean  $i$  is the individual effect. The Expected right  $i$  in absence of fluctuating asymmetry (FA) is denoted by adding the latter 2 effects (Ind and Side). FA estimate is based on the deviation between the Expected Right  $i$  and the (observed) Mean Right  $i$  (in fact, the FA estimate combines this distance with the same one corresponding to the left side). This FA estimate takes into account the measurement error and the possible occurrence of differences between the mean right and left sides

Another concern regarding the statistical characteristics of FA analysis of univariate traits (e.g. size, weight, etc.) is that individual FA is estimated by only 2 (left and right) replicates (van Dongen 2006). Therefore, some blurring of the patterns depicted by FA can be expected, and a very large sample size is needed in some cases (van Dongen 2007). The analysis of structures with multiple subunits (e.g. the petals of radially symmetric flowers) has been suggested as an alternative (van Dongen 2007). Combining a number of univariate traits with bilateral symmetry has been proposed also (Leung et al. 2000). However, a more general alternative could be the analysis of FA using shape. The shape of an otolith can be defined, for example, by the coordinates of a large number of points (landmarks) located along its perimeter. FA of each one of these points is then defined as random, left-right departures from their ideal position. As depicted in Fig. 1 for size, the ideal position of a point  $i$  in the (e.g. right) otolith is estimated from the average position of the same point  $i$  across all right otoliths in the population. However, FA at the individual level may accumulate FA's of all the points defining the shape of the structure under study. Therefore, FA estimated by shape differences could be more capable of

detecting any pattern in FA, provided that it exists. The extension of the statistical model depicted in Fig. 1 for shape (i.e. a multivariate analysis) has been proposed by Klingenberg & McIntyre (1998) and is also known as Procrustes analysis of variance (Procrustes ANOVA).

Obviously, proper comparison of point positions across all individuals requires some preprocessing. Fortunately, the core methodology for coordinate-based shape analysis has already been well established (Rohlf & Marcus 1993, Adams et al. 2004, Slice 2007). The basic steps are shown using fish otoliths (Fig. 2) and include (1) reflection of the left otoliths, for example, (2) scaling of all otoliths to the same size, (3) superimposing the centroid of all otoliths and (4) rotating the otoliths until optimal fit is achieved (Klingenberg & McIntyre 1998).

Fish otoliths are excellent structures for FA studies. Surprisingly, there are very few studies on the relationship between otolith FA and stress/fitness; one such study was based on shape, and another used the statistical model shown in Fig. 1. None of the studies done to date has combined the separate results obtained from univariate traits. Moreover, results show no clear trend (Table 1). Our aim is to demonstrate the capabilities of Procrustes ANOVA, using the shape of the saggita otoliths of a littoral fish (the razor fish *Xyrichtys novacula*, Labridae). Specifically, we sought to: (1) test for the existence of different types of asymmetry with respect to shape (multivariate analysis), (2) isolate the FA-related component of this asymmetry, (3) test the relationship between this FA-related component and

some putative explanatory variables, (4) compare the results obtained in (1) to (3) with those obtained using size (univariate analysis) and (5) demonstrate, using a simulation experiment, that FA measured using shape

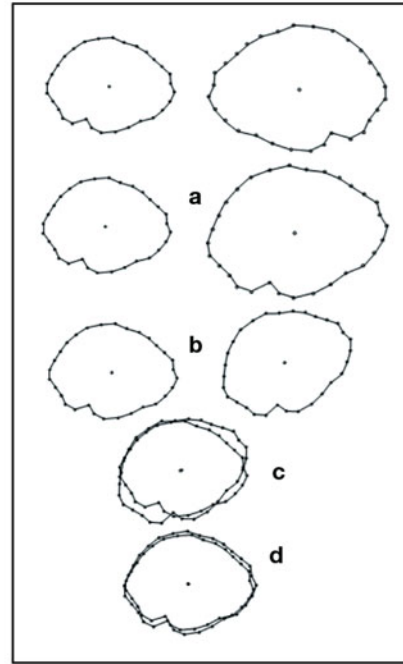


Fig. 2. Preprocessing of the  $x$  and  $y$  coordinates of the left and right otolith outlines prior to the statistical analyses. (a) First step: reflection of one of the 2 otoliths, (b) scaling of all the otoliths to the same size, (c) superimposing the centroid of both otoliths; and (d) rotating the centroids until optimal fit is reached

Table 1. Summary of other studies dealing with otolith asymmetry. The results support (Yes) or do not support (No) the existence of significant patterns between putative effects and fluctuating asymmetry (FA). Cases of partial or weak support were indicated (Yes/No). The number (in brackets) and nature (univariate vs. shape) of the traits analysed are detailed. Type of asymmetry and method are indicated (no ME: no measurement error estimation was completed). The method for testing absence of directional asymmetry (DA) is indicated

| Goal   | Results                              | Trait                    | Type of asymmetry  | Study                        |
|--|--------------------------------------|--------------------------|--|------------------------------|
| Nutritional condition (DNA/RNA ratio)                        | Yes                                  | Univariate (3)           | FA. Two factor mixed model   | Grønkjær & Sand (2003)       |
| Correlation with body condition                              | Yes/No                               | Univariate (4)           | FA; no ME; absence of DA by Left-Right distribution                          | Somarakis et al. (1997)      |
| Index of nutritional condition                               | <i>A priori</i> assumed (not tested) | Shape (geometric method) | Total asymmetry; no ME   | Chen & Chiu (2003)           |
| Correlation with an enzymatic biomarker of organic pollution | Yes/No                               | Univariate (3)           | FA; no ME; absence of DA by Left-Right distribution                          | Franco et al. (2002)         |
| FA along an osmotic gradient                                 | No                                   | Univariate (5)           | FA; no ME; absence of DA by Left-Right distribution                          | Panfili et al. (2006)        |
| Growth   | Yes/No                               | Univariate (3)           | FA; no ME; absence of DA by Left-Right distribution                          | Fey & Hare (2008)            |
| Larval abundance, growth, body condition                     | Yes/No                               | Univariate (22)          | FA; ME estimated in other data set; absence of DA by Left-Right distribution | Lemberget & McCormick (2009) |

has more statistical power than univariate analysis. Thousands of otoliths are systematically extracted from fish every year; thus the protocol proposed here may improve future exploration of the potential link between FA and fitness and provide a better indicator of performance in fish populations.

## MATERIALS AND METHODS

**Study species and populations.** The razor fish *Xyrichthys novacula* is a protogynous hermaphrodite fish; that is, members of the species first mature as females and later change into males (Linde & Palmer 2008). In the Western Mediterranean, this species is found on sandy bottoms near *Posidonia oceanica* beds (Riera & Linde 2001). A total of 180 individuals of *X. novacula* were fished by hook and hand line during June and July 2003 (reproductive season). Fishing was carried out at 5 sites, at Tarragona ( $n = 54$ , TL [mean total length, mm] = 146.5, SD = 2.16); Minorca Island (North Minorca marine reserve,  $n = 34$ , TL = 190.0, SD = 2.03); and 3 sites on Majorca Island, namely Cap Salines ( $n = 37$ , TL = 147.7; SD = 2.11), Andratx ( $n = 24$ , TL = 144.4, SD = 2.24) and Cabrera Island National Park ( $n = 31$ , TL = 161.0, SD = 2.15). Fish total length (mm) and sex were recorded, and saggita otoliths were dissected, cleaned and stored dry.

**Obtaining and managing images of the left and right otoliths.** Otolith images (proximal side) were captured using a Leica DC300 digital camera coupled to a Leica MZ16 binocular microscope and stored in JPG format (2088 × 1552 Pixels, or 3.24 MPixels). Concave otoliths may imply additional problems, but the saggital otolith of *Xyrichthys novacula* is nearly flat. Two independent (non-consecutive) photographs were taken for each of the 2 otoliths (4 images per specimen). We analyzed 180 individuals of *X. novacula* from 5 populations and 720 (180 × 4) images. Image files were named with an ID code for the individual and added suffixes (*left* or *right*, magnification and replicate number). Thousands of images can be managed in this way, which allows for automatic selection of subsets of images, with the individual ID code linking the images with additional information (e.g. population and otolith size are used in further analyses; see below).

**Extracting points from the outline of each image.** The images were submitted to a MATLAB program (available on request), which extracted the contour (using a threshold-based method) and adjusted it to an elliptic Fourier transformation as formulated by Rohlf (1993). The starting point used in all cases was the intercept of the sulcus, a valley corresponding to the macula acustica, on the otolith contour. This starting

point, along with the nucleus of the otolith, was clicked by hand on each image. In addition, 29 equidistant points along the contour, defined by the Fourier function, were automatically generated. Note that elliptic Fourier transformation and back transformation to the coordinate space was used only to generate equidistant coordinates. Coordinates of one of the 2 otoliths were mirrored (by changing the sign of the x-coordinates of all points) in order to facilitate the next step (Fig. 2a).

**Processing point coordinates for comparison.** In this step, the shape of an otolith was characterized by a set of 30 points along its perimeter plus the position of the nucleus. However, the x and y coordinates of these points were meaningless, since (for example) they depend on the specific position of each otolith within the image field, and did not allow for proper comparison within and between individuals.

Landmarks are points whose coordinates are assumed to be comparable (i.e. geometrically homologous) across all specimens. The set of landmarks defining the shape of a structure constitutes its landmark configuration. Geometric morphometry (Rohlf & Marcus 1993) provides the framework for making landmark configurations comparable between individuals. Landmark coordinates must represent shape only. Thus, all landmark configurations were first translated (Fig. 2b; the centroid [averaged x and y coordinates] of each landmark configuration is subtracted from all 31 x and y coordinates), then scaled (Fig. 2c; coordinates are divided by the centroid size [square root of the sum of the squared distance between the landmarks and the centroid]), and finally rotated around the centroid (Fig. 2d). After one landmark configuration was chosen to be the reference, a second configuration was rotated around the centroid until the summed squared distance between homologous landmarks, along all 31 landmarks, was minimized. This was achieved by means of an iterative procedure that uses the mean configuration (across all configurations) as a reference configuration (Generalized Procrustes Superimposition, GPS; Rohlf 1990). Note that the scaling step eliminates any isometric effect of size on shape.

This is the conventional procedure for processing true landmarks. However, only 2 points are true landmarks; the nucleus and the intercept of the sulcus at the otolith contour are fully determined by the surrounding anatomy to render their coordinates well defined in all directions (Slice 2007). Therefore, any change in any direction implies shape change. Conversely, the biological meaning of the coordinates of the remaining 29 evenly distributed points along the outline is different, because a change of position along the curve (i.e. parallel) does not imply any shape change, because the biological relevance lies in the shape of the curve. Such

points are called semilandmarks, because only variation orthogonal to the curve may imply biologically relevant information (Slice 2007). An additional processing step is applied to the semi-landmarks. They are allowed to slide on the curve until they match, as well as possible, the positions of the corresponding point along the curve in the reference configuration (Adams et al. 2004). Two criteria have been proposed for sliding landmarks: Procrustes distance and minimum bending energy. Since results obtained using these 2 criteria can differ (Perez et al. 2006), we used both methods and compared the results.

In summary, the full alignment process (i.e. landmark superimposition) involves sliding semi-landmarks, translation, rotation and scaling. Landmark superimposition was completed using *tpsRelw* (at <http://life.bio.sunysb.edu/morph/>; Rohlf 2007). After alignment, the shape of each otolith is fully defined by a set of 31 vectors connecting the 31 landmarks of the reference configuration with the 31 homologous landmarks. These 62 ( $31 \times 2$  x and y coordinates) shape descriptors (called Procrustes residuals) can be used directly as input for multivariate shape analysis (Rohlf & Marcus 1993).

The dimensionality of the shape matrix ultimately obtained is 62 columns by 720 (180 individuals  $\times$  2 sides  $\times$  2 replicates) rows. These 720 rows can be viewed as points in a shape space of 62 dimensions. The Euclidean distance between 2 points is an estimate of shape difference. Note, however, that the true dimensionality of the shape matrix is  $720 \times 58$ , because 4 parameters are lost (i.e. they became linearly dependent on the remaining parameters) during the alignment process; Rohlf 2007).

**Procrustes ANOVA.** The next step consisted of disentangling FA from DA. To do this, we used Procrustes ANOVA, the multivariate extension of the univariate mixed model proposed by Leamy (1984) and Palmer & Strobeck (1986). This model can be written as:

$$Y_{ijk} = \text{Grand mean} + \text{Ind}_i + \text{Side}_j + (\text{Ind} \times \text{Side})_{ij} + \text{ME}_{ijk} \quad (\text{Model 1})$$

where *Ind* is considered a random factor and *Side* is a fixed factor ( $i = 1$  to  $n$ ,  $n$  being the number of individuals, and  $j = \text{left or right}$ ). The number of replicated measurements of each side of the same individual ( $k$ ) was 2 in the case of *Xyrichthys novacula*. As illustrated in Fig. 1, the main advantage of this model is that the *Ind*  $\times$  *Side* interaction is an estimate of FA, even in the case that significant DA exists. DA is tested using the *Side* factor (fixed) against the interaction (*Ind*  $\times$  *Side*) because of the random nature of the factor *Ind*. FA is tested using the *Ind*  $\times$  *Side* interaction against ME (i.e. the residual variation is an estimate of the measurement error).

One of the implicit assumptions of this model is that the mean left side (like the right side) across all individuals reflects the optimal state of every individual  $i$ , and thus the departure from this optimal state is a true measurement of  $\text{FA}_i$ . Note, however, that any difference in DA between some subgroups of individuals would invalidate this assumption (van Dongen 2006). The existence of any pattern in DA can be tested using the model:

$$Y_{ijk} = \text{Grand mean} + \text{Ind}_i + \text{Side}_j + (\text{Ind} \times \text{Side})_{ij} + Z_i + (\text{Side} \times Z)_{ji} + \text{ME}_{ijk} \quad (\text{Model 2})$$

where  $Z$  is any variable being tested for an association with DA. Note that the effect of interest is not  $Z$  but *Side*  $\times$   $Z$ . Here we exemplify it by testing the between-population differences in DA (as in the case described by van Dongen 2007). The effects of otolith size have also been tested to show an application with a quantitative variable.

The response variable ( $Y_{ijk}$ ) of these models can be both a matrix in the case of multivariate analyses of shape (720 rows [replicates] and 62 columns [shape descriptors]) and a vector, in the case of univariate analyses. Here we show the univariate analyses using the length of the radius at the sulcus (i.e. the distance between the 2 true landmarks). The results of univariate and multivariate analyses were compared in terms of the significance of different sources of variation (specifically, FA and DA), and with respect to percentage of variance explained (i.e. sum-of-squares of a specific source of variation/total sum-of-squares).

In practice, Procrustes ANOVAs were performed with the VEGAN library (<http://cc.oulu.fi/~jarioksa/>; Oksanen et al. 2008) of the R package ([www.r-project.org/](http://www.r-project.org/)). This choice is justified for 4 reasons. First, the null-hypothesis ( $H_0$ : a specific effect is zero;  $H_1$ , the alternative hypothesis, is that a specific effect is different from zero) is tested using a permutation procedure (Legendre & Legendre 1998). Thus, multinormality, an assumption that can rarely be met in multivariate analyses, is relaxed. Second, VEGAN was developed for ecological community analyses. Therefore, much emphasis is placed not only on testing but also on visualizing the patterns between rows (here, landmark configurations) and columns (here, shape variables) related to specific effects (Legendre & Legendre 1998; ter Braak & Smilauer 2002). The usefulness of this for studies of asymmetry is expanded on the next section. Third, exactly the same syntax is used for univariate and multivariate models (therefore, we reduce the risk of running slightly different analyses). Finally, here we have performed 3 sets of simulations (see the next sections) that are easily implemented within a programming environment like R. However, for doing routine Procrustes ANOVA, we strongly recommend the use of



MorphoJ ([http://www.flywings.org.uk/MorphoJ\\_page.htm](http://www.flywings.org.uk/MorphoJ_page.htm); Klingenberg 2008). We have compared the results obtained using VEGAN and R against those obtained with MorphoJ. All R-scripts (including those used for simulations) are available on request.

**Extracting FA.** Model 2 can be used for testing the differences in FA between groups of individuals. The variable  $Z$  may be, for example, control versus stressful experimental conditions. However, it is simpler, and therefore preferable, to derive a scalar estimate of  $FA_i$  (FA of the individual  $i$ ) and use it in further analysis (Klingenberg & McIntyre 1998, Klingenberg & Monteiro 2005). In the control/stress example, this analysis will be a conventional one-way ANOVA, with the usual degrees of freedom (i.e. a single measure of FA for each individual). A simple  $FA_i$  estimate for a univariate trait is shown in Fig. 1 (i.e. the distance between the 2 sides after removing the effect of Side and Ind). This is equivalent (in the absence of DA) to the unsigned differences between the left and right sides (Palmer & Strobeck 1986). The extension of this measure of  $FA_i$  to a multivariate space was proposed in Klingenberg & McIntyre (1998). Remember that the shape of a specific landmark configuration is defined by a vector of 62 (but 58 independent) shape descriptors, but the procedure is the same that is shown in Fig. 1. Thus, the side effect is removed by subtracting this vector from the vector describing the mean shape of the corresponding side. Similarly, the individual effect is removed by subtracting the vector describing the mean shape of the individual  $i$ . The result is a new vector of 58 elements. These 58 shape descriptors define a new shape space.  $FA_i$  is the Euclidean distance between the points in this space corresponding to the left and right otoliths of the individual  $i$  (after averaging the  $k$  replicates per otolith).

Principal components analysis (PCA) is a conventional tool for visualizing the general patterns depicted by multivariate data. PCA of the multivariate shape descriptors after removing the effects of Side and Ind has been used to visualize the main patterns given by the FA-related fraction of shape (Klingenberg & Monteiro 2005). The Euclidean distance between left and right sides of the individual  $i$  (i.e.  $FA_i$ ) in the original space and in the PCA space is the same. However, in this new space, dimensions are uncorrelated and ordered according to decreasing importance of explained variance (thus, the few first dimensions give us an approximate view of the general patterns; see next section). We computed  $FA_i$  as the Euclidean distances (in the PCA space) between the right and left sides.  $FA_i$  estimates were calculated using VEGAN, and the results were compared with those obtained with MorphoJ.

In addition, Klingenberg & Monteiro (2005) suggest normalizing the columns of PCA variables so that they

have the same variance (before calculating  $FA_i$ ). This normalization is assumed to remove the possible effects of an anisotropic variability. This variability refers to the scatter that forms the same landmark of all individuals after the alignment process. Here we have compared the results obtained with and without applying this normalization.

**Testing for antisymmetry.** Unlike DA (see simulation experiments below), the presence of antisymmetry will be indistinguishable from true FA when the general methodology proposed here is used. This is related to the implicit use of unsigned differences between the right and left sides (i.e. Euclidean distances in the multivariate case). Therefore, there is a need for testing for antisymmetry in the data. For a univariate trait, the usual procedure is to inspect the signed differences between the right and the left sides (Palmer & Strobeck 1986). Extreme antisymmetry is expected to yield a bimodal distribution, but any departure from normality is considered to be related to antisymmetry. Here, we propose to test for normality of the signed left-right differences in each one of the PCA axes representing FA (i.e. the PCA of the multivariate shape descriptors after subtracting the means of Side and Ind). The usefulness of this procedure was confirmed using a simulation experiment, the full details of which are provided in the last section of 'Materials and methods'.

**Testing  $FA_i$  patterns.** The final step in our analysis was to analyze  $FA_i$  patterns. Here, as a very simple application, we explore (1) the existence of between-population differences in mean FA using a conventional one-way ANOVA and (2) the relationships between FA and otolith size. In addition, and as one of the main points of this methodological note, the full procedure was repeated using the length of the otolith radius along the sulcus in order to compare the performance of univariate and multivariate approach.

**Simulation experiments.** Three simulation experiments were done. We simulated a data set consisting of 100 individuals (31 landmarks each) with known (simulated) DI. The simulation procedure was based on the method of van Dongen (2006).  $DI_i$  (DI of the individual  $i$ ) was randomly drawn from a uniform distribution between 5 and 7. All simulated otoliths were computer-generated from the same otolith that was randomly extracted from the actual 180 otoliths of *Xyrichthys novacula*.

In order to simulate FA, the  $x$  and  $y$  coordinates of the landmark  $j$  of the simulated individual  $i$  ( $L_{ij}$ ) were randomly moved from their original position by a distance drawn from the normal distribution  $N(0, DI_i)$ . DA was simulated by adding a fixed value ( $X$ ) to  $x$  and  $y$  coordinates of half the landmarks of all the right otoliths (thus producing a fixed change of shape unrelated to DI). Antisymmetry was simulated in the same

way as DA, except that the side with a different shape was randomly selected from a binomial distribution (i.e. allowing for changing of the proportion of right or left sides affected by the shape change). ME was simulated by moving each  $L_{ij}$  a distance drawn from the normal distribution  $N(0, 2)$ . Four replicates (2 left and 2 right otoliths) were simulated in the same way for each individual  $i$ .

The 400 simulated landmark configurations (100 individuals  $\times$  2 sides  $\times$  2 replicates) were aligned using GPS. Procrustes residuals were extracted from the aligned coordinates and directly used as input for a (multivariate) Procrustes ANOVA. The FA descriptors corresponding to the side  $\times$  individual effects after removing the effects of side and individual were retained (i.e. following the same procedure used for the case of *Xyrichthys novacula*), and the Euclidean distance between them for the left and right side of the individual  $i$  is assumed to be an estimate of  $FA_i$ .

The first simulation experiment addressed the power of multivariate versus univariate approaches in detecting the true relationship between DI and FA. In this case, DA and antisymmetry were set to zero. The same procedure was repeated with one univariate trait (radius at the sulcus). This procedure was repeated 1000 times.

In the second simulation experiment,  $X$  was set to 10 units (approx. 1% of the otolith width) in order to simulate a significant DA. The capacity of the Procrustes ANOVA to recover the true relationship between FA and DI was used to determine if the multivariate approach correctly extracted individual FA values, even in case significant DA was present.

The purpose of the final simulation was to detect antisymmetry. The final simulation demonstrated the usefulness of testing the normality of the signed differences between the left and right sides. These differences were determined in the FA shape space and after completing a PCA (see above). In this simulation, antisymmetry was generated by setting  $X$  to 10, but the

otolith experiencing the shape change (left or right) was randomly selected from a binomial distribution with  $p = q = 0.5$ .

All simulations were completed with R package at [www.r-project.org/](http://www.r-project.org/) using the libraries VEGAN (multivariate analysis; by J. Oksanen et al., <http://cc.oulu.fi/~jarioksa/softhelp/vegan.html>) and SHAPES (Procrustes alignment and shape analysis; by I. Dryden). The R scripts are available on request.

## RESULTS AND DISCUSSION

### Variance partitioning

Concerning the empirical case of *Xyrichthys novacula*, multivariate MANOVA of shape descriptors allowed the partitioning of shape variability into 4 sources (Table 2). The effects of the interaction Side  $\times$  Ind (representing FA) accounted for up to 23.2% of the total variation (and 98% of the asymmetry-related variation) and was more than 5-fold the measurement sum-of-squares error (4.2%). It is noteworthy that side effects (DA) were very small (0.5%) but significantly different from zero. However, assuming that there is no bias of DA between individuals (see below), the interaction Side  $\times$  Ind is still a valid estimate of FA (see results below and Söderman et al. 2007). As expected, between-individual differences represent the largest source of variability (72.1%), but this proportion is of no interest in the present context. The results obtained with VEGAN and MorphoJ were identical.

The results obtained using 2 different methods for sliding semi-landmarks (minimum bending energy and Procrustes method) were very close in terms of variance partitioning (Table 3). In addition, measures of  $FA_i$  using the 2 methods are clearly correlated ( $r$ [Pearson correlation] = 0.87, 95% confidence interval: 0.84 to 0.91;  $n = 180$ ). Perez et al. (2006) reported some differences in the results obtained using these 2

Table 2. MANOVA table indicating the amount of variance explained in terms of sum-of-squares (SS) by each of the factors considered in Model 1 (Var SS), degrees of freedom (for each one of the 58 shape descriptors), the percentage of the variance explained in relation to total variance (%Total) and to asymmetry-related variance (%Asymm.), the  $F$  values and the probability ( $p$ ). Probability estimates were based on 999 permutations. Otolith shape is defined using the Minimum Bending Energy method for sliding semi-landmarks

|                                    | Var SS    | df  | %Total | %Asymm. | $F$                                | $p$        | Permutations   |
|------------------------------------|-----------|-----|--------|---------|------------------------------------|------------|--|
| Individual                         | 0.0036492 | 179 | 71.9   |         | 3.0766<br>(Ind/Ind $\times$ Side)  | Not tested |  |
| Side (DA)                          | 0.0000239 | 1   | 0.5    | 2.0     | 3.6070<br>(Side/Ind $\times$ Side) | <0.001     | Randomly exchanging sides within individuals but not between individuals |
| Interaction Ind $\times$ Side (FA) | 0.0011861 | 179 | 23.4   | 98.0    | 11.0830<br>(Ind $\times$ Side/ME)  | <0.001     | Unconstrained permutation  |
| Residual (ME)                      | 0.0002152 | 360 | 4.2    |         |                                    |            |  |

Table 3. Partitioning shape variability: comparison between an univariate method (Size) and 2 multivariate methods (Shape-Minimum Bending Energy [MBE] and Shape-Procrustes [Proc.]). Shape methods differ in how semi-landmarks are aligned. Variance explained in terms of sum-of-squares by each of the factors considered in Model 1 (Var SS) are indicated in the left 3 columns. The corresponding  $F$ -values are indicated in the right 3 columns. Probability values based on 999 permutations are also added in brackets (nt, not tested)

|                                    | Var SS |             |               | $F$             |                 |                 |
|------------------------------------|--------|-------------|---------------|-----------------|-----------------|-----------------|
|                                    | Size   | Shape (MBE) | Shape (Proc.) | Size            | Shape (MBE)     | Shape (Proc.)   |
| Individual                         | 152285 | 0.0036492   | 0.0018470     | 211.411 (nt)    | 3.0766 (nt)     | 3.3125 (nt)     |
| Side (DA)                          | 42     | 0.0000239   | 0.0000111     | 10.519 (0.002)  | 3.6070 (<0.001) | 3.5551 (<0.001) |
| Interaction Ind $\times$ Side (FA) | 720    | 0.0011861   | 0.0005576     | 24.399 (<0.001) | 11.083 (<0.001) | 9.748 (<0.001)  |
| Residual (ME)                      | 59     | 0.0002152   | 0.0001150     |                 |                 |                 |

methods on human facial skeletons. However, these researchers also reported a case in which between-method differences were smaller (human molars). Human molars are very similar to fish otoliths in that both comprise equidistant points along a closed outline. By contrast, human facial skeletons are more variable and are described by a combination of true landmarks with open and closed outlines. In addition, points are not equidistant on the curves, and point density is smaller (Perez et al. 2006). Obviously, the effects of the method used for sliding landmarks needs further attention (Slice 2007), but the results reported here suggest that differences in some cases may be less important than previously suggested (Perez et al. 2006). Concerning the case of *Xyrichthys novacula*, both methods rendered the same qualitative results for all the analyses detailed below. Therefore, only the results concerning shape measured using minimum bending energy as a method for sliding semi-landmarks are reported hereafter.

Analysis of the univariate trait showed similar trends; the contribution of both DA and FA to size asymmetry was significantly different from zero (Table 3).

#### Testing between-group differences in DA

As previously mentioned, one of the key assumptions implicit in Model 1 is that the mean left side (the same for the right side) across all individuals reflects the optimal state of every individual. This reference point is needed in order to consider any departure from this optimal state as a valid measurement of  $FA_i$ . Note, however, that any difference in DA between some subgroups of individuals would invalidate this assumption (van Dongen 2006). The existence of these between-group differences was tested using Model 2 and focusing on the group  $\times$  side interaction. In the case of *Xyrichthys novacula*, we tested the null hypothesis that the interaction population  $\times$  side (i.e. population DA) is zero; the  $F$  value was determined using Ind  $\times$  Side as the denominator. The results obtained suggest that DA

is the same for all populations ( $F = 1.0605$ ;  $df = 4, 175$ ; Probability ( $p$ ) based on 1000 within-population permutations = 0.299). Similarly, the interaction otolith size  $\times$  side was not significant ( $F = 0.9982$ ;  $df = 1, 177$ ;  $p$  based on 1000 within-population permutations = 0.405). Obviously, there may be other unknown variables affecting DA that were not included in the model. However, usually it is possible to test almost all variables that have a biological basis.

#### Patterns in FA

Individual  $FA_i$  estimates were obtained for each of the 180 fish analyzed. One of the most important results reported here is that  $FA_i$  estimates obtained from shape are uncorrelated with those obtained from radius at the sulcus ( $r = 0.10$ ;  $p = 0.150$ ; Fig. 3).

The distribution of  $FA_i$  measured using shape was clearly skewed (Fig. 4). Concerning between-population

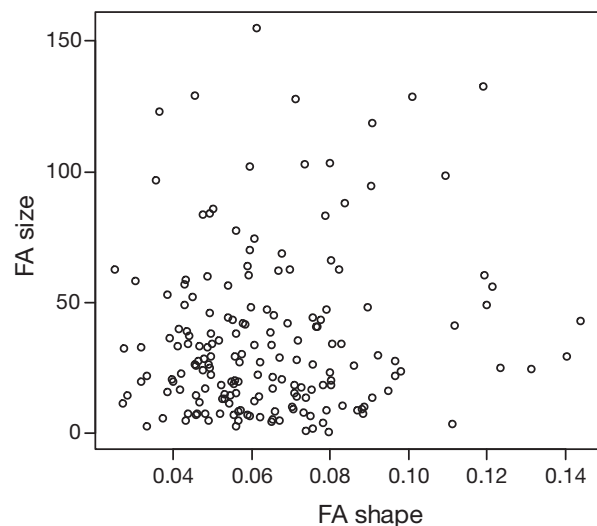


Fig. 3. Relationship between estimates of fluctuating asymmetry using shape (FA shape) and using the radius at the sulcus (FA size). Each point corresponds to an individual. FA shape has no units of measure because it is standardized by size (see Fig 2). FA size is in pixels (after scaling by magnification)



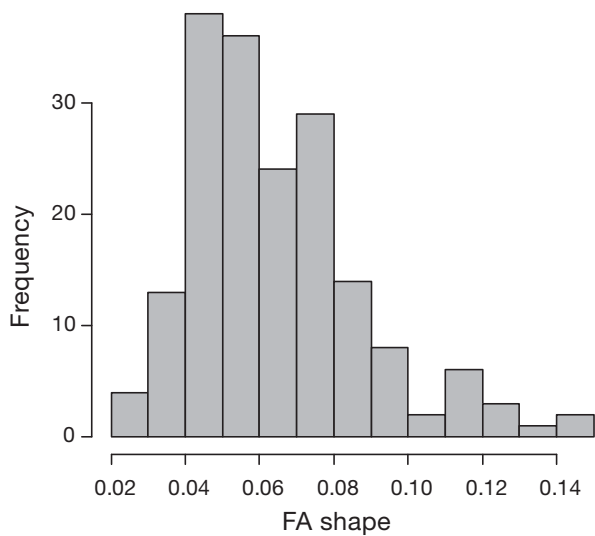


Fig. 4. Distribution of the fluctuating asymmetry values ( $FA_i$ ) using shape

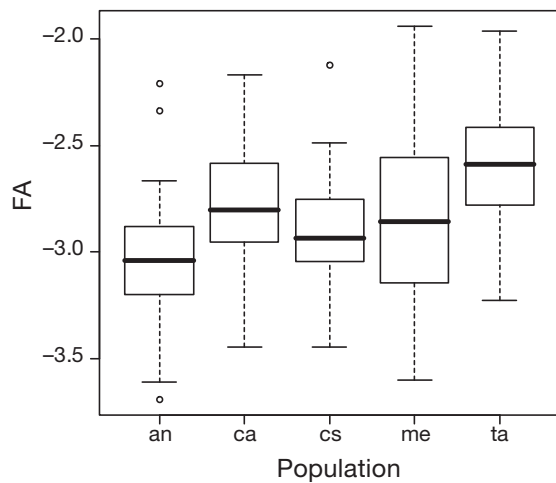


Fig. 5. Box-plot of the log-transformed fluctuating asymmetry values ( $FA_i$ ) values grouped by population. The central line is the median, the boxes indicate the first and third quartiles, and the whiskers indicate 2 standard deviations. Outliers are those outside median  $\pm 2$  SD. Population codes are: an (Andratx, Mallorca), ca (Cabrera Island National Park [Marine protected area]), cs (Cap Salines, Mallorca), me (North Menorca Marine Reserve), and ta (Tarragona)

differences, one-way ANOVA demonstrated that mean FA of the 5 populations was significantly different (Fig. 5; SS[sum-of-squares] = 4.31;  $df = 4$ , residual SS = 16.67; residual  $df = 175$ ;  $F = 11.4$ ;  $p < 10^{-8}$ ). The distribution of the residuals was also skewed, with a larger number of individuals showing lower asymmetry than the median. However, a simple log-transformation rendered normally distributed residuals and comparable within-group variances. Therefore, the probability values reported above have been ob-

tained using a conventional parametric test of the log-transformed FA.

Between-population differences of the mean FA did not differ significantly from zero when FA estimates were based on the radius at the sulcus ( $F = 1.10$ ,  $df = 4$ ,  $p = 0.35$ ). In this case, the probability value was obtained using randomization methods, because residuals were clearly skewed, even after applying conventional transformations.

The effects of otolith size on FA were non-significant for shape-based estimates (SS = 0.07;  $df = 1$ , residual SS = 16.60; residual  $df = 174$ ;  $F = 0.80$ ;  $p = 0.37$ ). In this case, population was included in the model as a covariate because one of the populations (Menorca) shows larger otoliths. Residuals after log-transformation were normally distributed, and the probability values reported have been obtained using a conventional parametric test. Note that shape was corrected for isometric changes during the process of landmark alignment (Fig. 2). Palmer & Strobeck (1986) suggest examining the relationship between size and FA because the existence of such a relationship would complicate the interpretation of the pattern observed (larger size can imply larger variance also). This is not the case with *Xyrichthys novacula*.

The effect on  $FA_i$  of the correction for anisotropy proposed by Klingenberg & Monteiro (2005) seems to be small in the case of *Xyrichthys novacula*. Corrected  $FA_i$  are well correlated with the original (uncorrected)  $FA_i$  ( $r = 0.82$ ;  $n = 180$ ;  $p < 0.001$ ).

### Simulation experiments

**$FA_{\text{shape}}$  versus  $FA_{\text{sulcus}}$ .** In the first simulation experiment, it is clear that at least for some range of variation of DI, the chance of detecting a true relationship between DI and FA is increased when using shape as the FA estimator. Results of a single simulation are shown in Fig. 6. In this case, FA and DI are significantly correlated ( $r^2 = 0.79$ ,  $n = 100$ ). In contrast, the univariate descriptor of FA (radius at the sulcus) failed to show a relationship between DI and FA ( $r^2 = 0.05$ ). After completing 1000 of these simulations, FA in shape correctly identified the existence of a correlation between FA and DI 100% of the time; when radius at the sulcus was used, this proportion dropped to 35.5%.

**DA.** The second simulation experiment focused on the capability of Model 1 to correctly estimate FA in the presence of significant DA. The shape differences in the raw data between the left and the right sides were clearly depicted using a PCA (Fig. 7a). Conversely, when the PCA was completed on the fraction corresponding to the interaction  $\text{Ind} \times \text{Side}$  (after removing the effects of Side and Ind, and the measure-

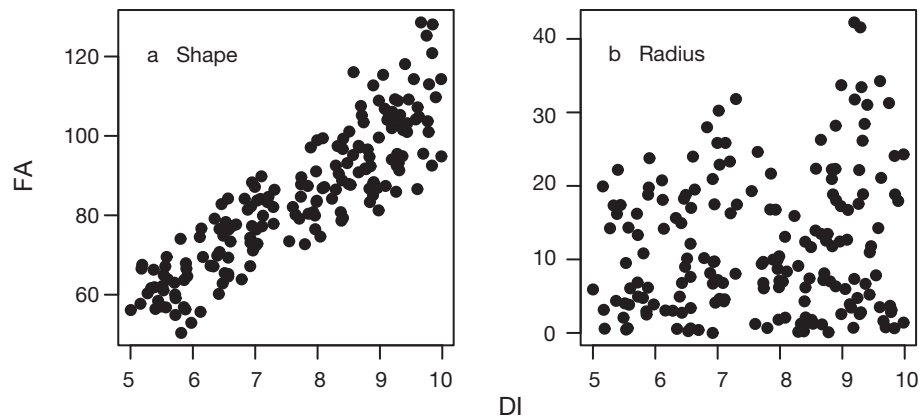


Fig. 6. Correlation between simulated developmental instability (DI) values and fluctuating asymmetry (FA) (see text for details on FA estimation). Panels correspond to the FA values determined using (a) shape and (b) radius at the sulcus

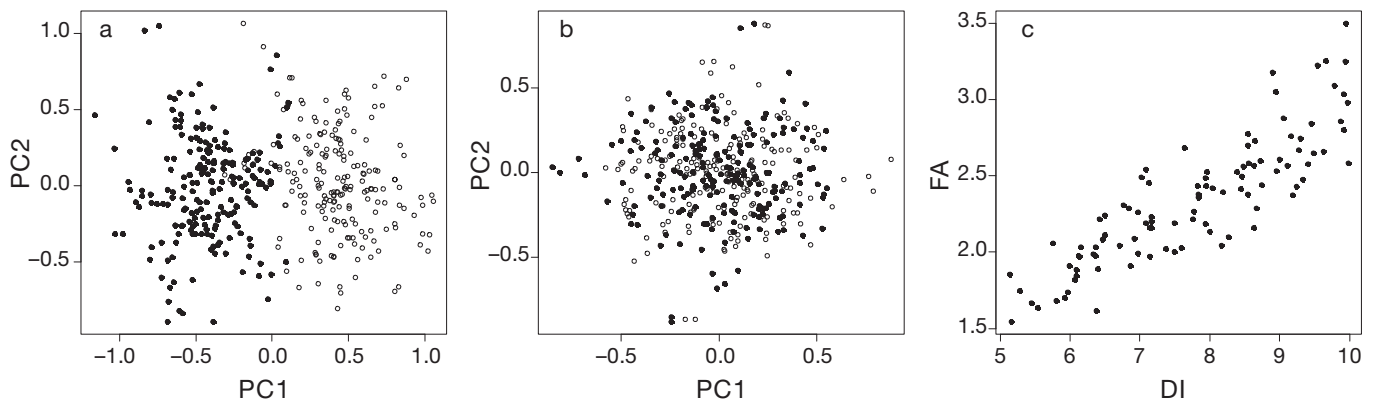


Fig. 7. Simulated data with directional asymmetry. Scatter plot (first 2 dimensions) of the principal components analysis (PCA) scores of all landmark configurations. Configurations corresponding to the left (open points) and right side (closed points) are differentiated. (a) PCA performed on the raw data, and 2 separate clusters corresponding to right and left configurations are evident. (b) PCA performed on the shape descriptors corresponding to Side  $\times$  Ind interaction only (see text). This fraction is now free from directional left-right differences (open and closed points are well mixed). (c) Fluctuating asymmetry of the individual  $i$  ( $FA_i$ ) calculated as left-right Euclidean distances within the space of the central panel; this correctly recovers the true relation between developmental instability (DI) and FA

ment error), the differences between sides disappear (Fig. 7b). FA is estimated as the Euclidean distance between the points corresponding to the right and left sides of the same otolith (along all the dimensions of the space, not only on the first 2 dimensions displayed in Fig. 7b). These  $FA_i$  estimates reflect true FA, because they are highly correlated with the known value of  $DI_i$  (Fig. 7c).

**Antisymmetry.** The third experiment simulated a data set with a significant amount of antisymmetry. This is also shown using a PCA of the raw data. In this case, there are 2 clear clusters of points, but they do not correspond to the left and right otoliths (Fig. 8a). Note that the PCA of the remaining fraction after removing Side and Ind still show 2 clusters (Fig. 8b; compare with Fig. 7b).

This is because there is not an ideal left and right side. In the absence of antisymmetry, it is expected that the signed differences of the scores in every dimension of the PCA will be normally distributed. Therefore, we propose to test these signed differences for normality as a method for detecting antisymmetry. The histograms of the first 4 dimensions of a simulated data set with antisymmetry are shown in Fig. 9 (simulated data with (Fig. 9a) and without (Fig. 9b) antisymmetry) and Fig. 9c (*Xyrichthys novacula*, which seems to be free from antisymmetry).

In conclusion, for physiological reasons, otolith size seems *a priori* better than otolith shape for demonstrating some cause-and-effect link with fitness. However, here the issue is not the size itself but the

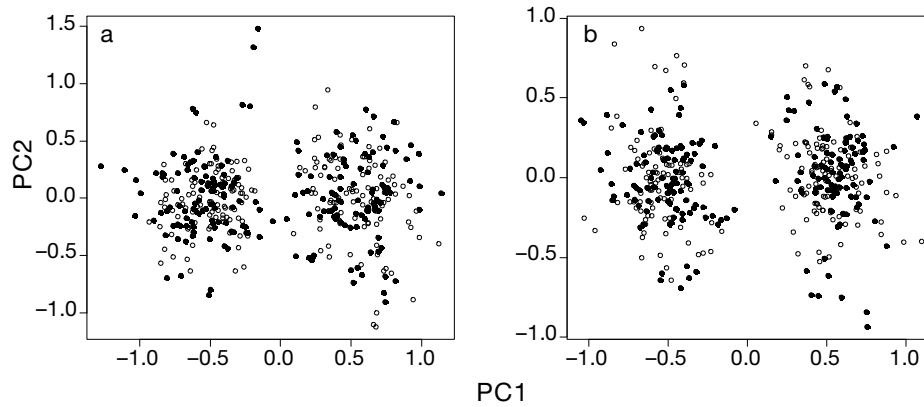


Fig. 8. Simulated data with antisymmetry: Scatter plot (first 2 dimensions) of the principal components analysis (PCA) scores of all landmark configurations. Configurations corresponding to the left (open points) and right side (closed points) are differentiated. (a) PCA performed on the raw data. There are 2 clear clusters of points, but they do not correspond to the left and right sides of the otoliths. (b) PCA performed on the shape descriptors corresponding to Side  $\times$  Ind interaction only (see text). This fraction still shows 2 different clusters; thus left-right Euclidean distances are not reliable estimates of fluctuating asymmetry (FA)

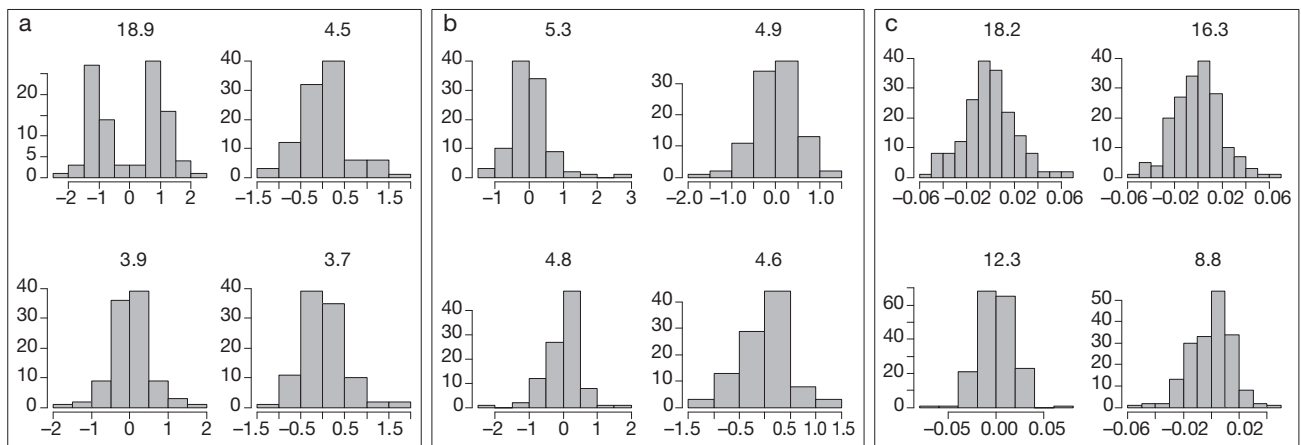


Fig. 9. Detecting antisymmetry: histograms of the distribution of the signed differences between left and right otoliths within the principal components analysis (PCA) space. These differences were calculated for each of the first 4 dimensions of the space corresponding to the interaction Side  $\times$  Ind only (Figs. 7b & 8b). Numbers above the histograms represent the variance explained (%) by each dimension. (a) Simulated data set with antisymmetry clearly reflected in the bimodal distribution of the signed differences at the first dimension. (b) Simulated data set without antisymmetry; all histograms corresponding to normal distributions. (c) Empirical case (*Xyrichthys novacula*) all histograms correspond to normal distributions, thus no antisymmetry is assumed to occur

FA of a biological structure, and the latter is better estimated combining multiple univariate descriptors (Leung et al. 2000) or using shape as proposed here, than using any single univariate trait. The use of shape should be considered a major improvement for FA studies. Multivariate analysis of otolith shape using the proposed 2-factor mixed model (Ind+Side+Ind  $\times$  Side) shows promise as a method to better understand the links between otolith FA and fitness or stress.

**Acknowledgements.** We thank C. P. Klingenberg and J. Oksanen for their clever comments on the statistical analysis. J. Tomas and 2 anonymous referees helped improve the manuscript. We thank F. J. Rohlf, C. P. Klingenberg, J. Oksanen and I. Dryden for making software freely available. We thank the staff of the marine reserves for help in sampling fishes. The results presented in this paper comprise parts of 3 research projects: RAOR (funded by the Conselleria d'Agricultura i Pesca del Govern de les Illes Balears), DONZELLA (funded by the Conselleria de Economia e Innovació del Govern de les Illes Balears) and ROQUER (CTM2005-00283, funded by the Ministerio de Educación y Ciencia).

## LITERATURE CITED

- Adams DC, Rohlf FJ, Slice DE (2004) Geometric morphometrics: ten years of progress following the 'revolution'. *Ital J Zool* 71:5–16
- Bjorksten TA, Fowler K, Pomiankowski A (2000) What does sexual trait FA tell us about stress? *Trends Ecol Evol* 15: 163–166
- Chen CS, Chiu TS (2003) Early life history traits of Japanese anchovy in the northeastern waters of Taiwan, with reference to larval transport. *Zool Stud* 42:248–257
- Clarke GM (1998) Developmental stability and fitness: The evidence is not quite so clear. *Am Nat* 152:762–776
- Debat V, Cornette R, Korol AB, Nevo E, Soulet D, David JR (2008) Multidimensional analysis of *Drosophila* wing variation in Evolution Canyon. *J Genet* 87:407–419
- Fey DP, Hare JA (2008) Fluctuating asymmetry in the otoliths of larval Atlantic menhaden *Brevoortia tyrannus* (Latrobe)—a condition indicator? *J Fish Biol* 72:121–130
- Franco A, Malavasi S, Pranovi F, Nasci C, Torricelli P (2002) Ethoxyresorufin O-deethylase (EROD) activity and fluctuating asymmetry (FA) in *Zosterisessor ophiocephalus* (Teleostei, Gobiidae) as indicators of environmental stress in the Venice lagoon. *J Aquat Ecosyst Stress Recovery* 9: 239–247
- Grønkjær P, Sand MK (2003) Fluctuating asymmetry and nutritional condition of Baltic cod (*Gadus morhua*) larvae. *Mar Biol* 143:191–197
- Klingenberg CP (2008) MorphoJ. Faculty of Life Sciences, University of Manchester, UK. [www.flywings.org.uk/MorphoJ\\_page.htm](http://www.flywings.org.uk/MorphoJ_page.htm)
- Klingenberg CP, McIntyre GS (1998) Geometric morphometrics of developmental instability: analyzing patterns of fluctuating asymmetry with Procrustes methods. *Evolution* 52:1363–1375
- Klingenberg CP, Monteiro LR (2005) Distances and directions in multidimensional shape spaces: implications for morphometric applications. *Syst Biol* 54:678–688
- Klingenberg CP, Barluenga M, Meyer A (2002) Shape analysis of symmetric structures: quantifying variation among individuals and asymmetry. *Evolution* 56:1909–1920
- Leamy L (1984) Morphometric studies in inbred and hybrid house mice. V. Directional and fluctuating asymmetry. *Am Nat* 123:579–593
- Legendre P, Legendre L (1998) Numerical ecology. Elsevier Science, Amsterdam
- Lemberget T, McCormick MI (2009) Replenishment success linked to fluctuating asymmetry in larval fish. *Oecologia* 159:83–93
- Leung B, Forbes MR (1996) Fluctuating asymmetry in relation to stress and fitness: effects of trait type as revealed by meta-analysis. *Ecoscience* 3:400–413
- Leung B, Forbes MR, Houle D (2000) Fluctuating symmetry as a bioindicator of stress: comparing efficacy of analyses involving multiple traits. *Am Nat* 155:101–115
- Linde M, Palmer M (2008) Testing Allsop and West's size at sex change invariant within a fish species: a spurious ratio or a useful group descriptor? *J Evol Biol* 21:914–917
- Møller AP (1997) Developmental stability and fitness: a review. *Am Nat* 149:916–932
- Møller AP (2000) Symmetry, size and stress. *Trends Ecol Evol* 15:330–330
- Oksanen J, Kindt R, Legendre P, O'Hara B and others (2008) Vegan: Community Ecology Package. R package version 1.15-0. <http://cran.r-project.org/>, <http://vegan.r-forge.r-project.org/>
- Palmer AR (2000) Quasireplication and the contract of error: lessons from sex ratios, heritabilities and fluctuating asymmetry. *Annu Rev Ecol Syst* 31:441–480
- Palmer AR, Strobeck C (1986) Fluctuating asymmetry: measurement, analysis, patterns. *Annu Rev Ecol Syst* 17: 391–421
- Panfili J, Durand J, Diop K, Gourene B, Simier M (2005) Fluctuating asymmetry in fish otoliths and heterozygosity in stressful estuarine environments (West Africa). *Mar Freshw Res* 56:505–516
- Perez SI, Bernal V, Gonzalez PN (2006) Differences between sliding semi-landmark methods in geometric morphometrics, with an application to human craniofacial and dental variation. *J Anat* 208:769–784
- Riera F, Linde M (2001) El raor i la cirviola. Conèixer per preservar. Conselleria d'Agricultura i Pesca, Mallorca, Balearic Islands
- Rohlf FJ (1990) Rotational fit (Procrustes) methods. In: Rohlf FJ, Bookstein FL (eds) Proceedings of the Michigan Morphometric Workshop. University of Michigan, Ann Arbor, p 227–236
- Rohlf FJ (1993) NTSYS-pc. Numerical taxonomy and multivariate analysis system, version 1.18. Exeter Software, Setauket, NY
- Rohlf FJ (2007) TPSRelw, relative warps analysis, version 1.45. Department of Ecology and Evolution, State University of New York at Stony Brook
- Rohlf FJ, Marcus LF (1993) A revolution in morphometrics. *Trends Ecol Evol* 8:129–133
- Slice DE (2007) Geometric morphometrics. *Annu Rev Anthropol* 36:261–281
- Söderman F, van Dongen S, Pakkasmaa S, Merilä J (2007) Environmental stress increases skeletal fluctuating asymmetry in the moor frog *Rana arvalis*. *Oecologia* 151: 593–604
- Somarakis S, Kostikas I, Peristeraki N, Tsimenides N (1997) Fluctuating asymmetry in the otoliths of larval anchovy *Engraulis encrasicolus* and the use of developmental instability as an indicator of condition in larval fish. *Mar Ecol Prog Ser* 151:191–203
- ter Braak CJF, Smilauer P (2002) CANOCO reference manual and CanoDraw for Windows user's guide: software for canonical community ordination (version 4.5). Microcomputer Power, Ithaca, NY
- van Dongen S, Lens L (2000) Symmetry, size and stress. *Trends Ecol Evol* 15:330–331
- van Dongen S (2006) Fluctuating asymmetry and developmental instability in evolutionary biology: past, present and future. *J Evol Biol* 19:1727–1743
- van Dongen S (2007) What do we know about the heritability of developmental instability? answers from a bayesian model. *Evolution* 61:1033–1042

Editorial responsibility: Cynthia Jones,  
Norfolk, Virginia, USA

Submitted: March 30, 2009; Accepted: October 2, 2009  
Proofs received from author(s): January 6, 2010