Lost in the Other Half: Improving Accuracy in Geometric Morphometric Analyses of One Side of Bilaterally Symmetric Structures

Andrea Cardini[1,2,*]

[1]*Dipartimento di Scienze Chimiche e Geologiche, Universita' di Modena e Reggio Emilia, Via Campi 103, 41125 Modena, Italy;*
[2]*Centre for Forensic Science, The University of Western Australia, 35 Stirling Highway, Crawley, WA 6009, Australia*
[*]*Correspondence to be sent to: Dipartimento di Scienze Chimiche e Geologiche, Universita' di Modena e Reggio Emilia,*
Via Campi 103, 41125 Modena, Italy; Email: alcardini@gmail.com; cardini@unimo.it

Abstract.—Systematists and evolutionary biologists have widely adopted Procrustes-based geometric morphometrics for measuring size and shape in biology. Many structures, and in fact most animals, are bilaterally symmetric with an internal plane of symmetry (also called object symmetry). Often, when quantifying asymmetric variation is not an aim, only one or the other side is measured and analyzed. This approach has been used in hundreds of studies. Its implicit assumption is that the information on the other side is redundant and a single side will, therefore, produce results mirroring those one would have obtained from the analysis of the entire structure with all its left and right landmarks. However, the extent to which this assumption is met has, to my knowledge, never been explored. Using two example data sets, I will show that congruence may be high in analyses at a macroevolutionary level but much lower at a microevolutionary one, and inaccuracies might especially affect shape. I will discuss some of the other factors that may influence results and will suggest a simple expedient that can improve both the visualization and accuracy of shape analyses in one-side-only studies. [Allometry; landmarks; macroevolution; microevolution; mirror reflection; object symmetry; Procrustes; shape visualization.]

Over the last two decades, geometric morphometrics (GMM) has become the leading set of techniques for the quantitative analysis of biological forms (Rohlf and Marcus 1993; Adams et al. 2004, 2013). In this field, GMM methods based on Cartesian coordinates of anatomical landmarks and generalized Procrustes analysis (Rohlf and Slice 1990) have a special prominence. A Procrustes analysis (or Procrustes superimposition) is, in fact, a fairly simple mathematical "expedient" to extract size and shape data from the raw coordinates of anatomical landmarks. However, because of its desirable statistical properties (Adams et al. 2004, and references therein), this has led to the extensive adoption of the "Procrustes paradigm: a methodological approach to shape analysis arising from the intersection of the statistical shape theory and analytical procedures for obtaining shape variables from landmark data" (Adams et al. 2013, p. 7).

Procrustes-based GMM has produced innovative approaches in many disciplines including evolutionary developmental biology, comparative studies, palaeontology, biomechanics, and locomotion analysis (Cardini and Loy 2013). It allows also powerful analyses of structures with one or more planes of symmetry (Klingenberg et al. 2002; Savriama and Klingenberg 2011). Animals offer a particularly good and simple example of this type of morphological variation, because the vast majority of them are bilaterally symmetric. This means that their anatomical parts tend either to be in two separate copies, one on each side of the body, such as left and right hands (matching symmetry), or to have an internal midplane, which separates two mirror halves, as in the human face (object symmetry). However, when the researcher is not interested in quantifying asymmetry, GMM analyses are often performed on only one or the other side of the body. This is a shortcut to avoid redundancy and speed up data collection. A special case is that of fragmentary fossils, where one-side-only analyses can be the only option whenever one side is incomplete or missing. In fact, even with present-day material, when midsagittal sutures are loosely fixed (as in many reptile crania but also often in rodent mandibles), left and right sides might be disarticulated and that could push a researcher to study sides separately.

Using only one side inevitably leads to discarding some information from a complete symmetric structure. For instance, in assessing morphological differences of marmots in relation to their phylogenetic divergence, I compared size and shape using a configuration of nine landmarks on left hemimandibles, but made no use of right ones (Cardini 2003). The implicit assumption was that asymmetry was small and results would be similar. With the same reasoning, Cardini and Polly (2013) quantified evolutionary allometry in placentals using GMM on a sample of almost 900 specimens, in which only the left half of the cranium had been measured. Cardini and Polly (2013), and other examples of one-side-only analyses of structures with object symmetry, provide a particularly interesting case. This is because they not only discard the information on the other

side, but also employ a subset of the total landmark configuration. This makes differences potentially even more profound, as data, based on one or both sides, inevitably have different covariance structures and, in fact, even belong to different Procrustes shape spaces. This raises an important question: will the choice of using one or both sides produce differences in results that crucially depend on the covariation and set of interindividual distances in a sample?

EXAMPLES OF "ONE-SIDE-ONLY" ANALYSES

The selection of examples that follows illustrates how commonly one-side-only studies have been performed on different structures with object symmetry in vertebrates and invertebrates, and both at micro- and macroevolutionary levels. For instance, in some of the earliest applications of Procrustes-based GMM in systematics, Loy et al. (1993) and Rohlf et al. (1996) used landmarks on the right side of the dorsal cranium to compare species of Old World moles. Since then, the 1996 paper has become almost a template for studies of differences within and between populations (Rohlf 1998; Viscosi and Cardini 2011). One of the standard GMM references on multivariate regression using Procrustes data (Monteiro 1999) also made use of a one-side-only approach on dorsal views of lizard crania. Microevolutionary studies of morphological variation across the range of a species or a species complex require large samples and data collection can be made faster by focusing on one side only. Thus, Monteiro et al. (2003) studied adult punarè rats across their Brasilian range using right half of dorsal, lateral, and ventral views of the cranium; Cardini et al. (2007) analyzed vervet monkeys clines with 3D data on left cranial sides; and Tabatabaei Yazdi and Adriaens (2013) investigated differences in the left side of the cranium of jirds from a large area of the Middle-East.

The study of ontogeny may be less demanding in terms of sample size, because differences among age classes are typically large. However, especially if several species are measured and compared, more data are needed, and that might be part of the reasons why also ontogenetic analyses have often been done using only one side. For instance, Cardini and Thorington (2006) and Piras et al. (2010), both focused on cranial ontogenetic trajectories and assessed their differences across species (respectively, marmots and crocodylians). In the same study, Cardini and Thorington (2006) also compared results from one-side-only 2D and 3D landmarks, finding a generally good congruence. One-side-only analyses have been performed in many other contexts and taxonomic groups. Ivanović et al. (2009) studied the phylogenetic signal in dorsal and ventral views of one side of the cranium of alpine newts from the Balkans. Drotz (2003) and Drotz et al. (2012) investigated patterns of speciation and differences in ecotypes of a large species complex of diving beetles, and compared results from the right dorsal side of the body to molecular

findings. Macagno et al. (2011) estimated the divergence within and among species of dung beetles in Europe using one side of the head and epipharynx. Rufino et al. (2006) landmarked the left half of the carapace to quantify microevolutionary variation in crabs, whereas Hopkins and Lidgard (2012) explored frequencies of evolutionary modes in more than 150 species of fossil trilobites using a set of right-side landmarks on cranidia.

This brief overview represents a subset from more than 150 one-side-only studies I came across without doing an extensive bibliographic search. The approach does not seem to be losing popularity as shown in 2015, among others, by Meloro et al.'s (2015) and Yazdi et al.'s (2015) ecomorphological analyses of, respectively, the Canini (left half of the palate) and *Meriones* (left bulla and cranium); Álvarez et al.'s (2015) 3D study of integration in one side of the skull of caviomorph rodents; Pečnerová et al.'s (2015) analysis of the tree squirrel radiation using genetics and the left half of the cranium; Ruane's (2015) integrative taxonomy of milk snakes (left side of the head in dorsal view), and Head and Polly's (2015) macroevolutionary study of the covariation between hox genes and vertebral shape (one side only for each vertebra) in the evolution of snake body form.

Despite the evident, persistent and common use of one-side-only approaches, however, little or no attention seems to have been paid to understanding its consequences.

ARE WE LOSING SOMETHING BY ANALYZING JUST ONE SIDE OF STRUCTURES WITH OBJECT SYMMETRY?

That the quantification of asymmetries is not an aim of one-side-only studies is, of course, obvious. However, these studies do not generally either concern hypotheses explicitly requiring a test of variation exclusively on that side. As anticipated, the usually implicit assumption is that, in the absence of strong asymmetry, the loss of information is negligible and the pattern of differences in size and shape is faithfully captured by landmarking just one side. But is this really true?

This Point of View does not aim at providing a definitive answer to this question but wants to show that it is a fundamental issue and one clearly overlooked in hundreds of studies. Thus, I will be using two data sets as examples of object symmetry to quantify the accuracy of one-side-only analyses. Also, I will suggest a simple preliminary operation that might increase accuracy. In this specific context, the term accuracy is used in a relative sense to express how close one-side-only results are to results obtained from the total configuration (TC) with all bilateral landmarks.

CASE STUDIES: MICROEVOLUTION IN VERTEBRATES AND MACROEVOLUTION IN INVERTEBRATES
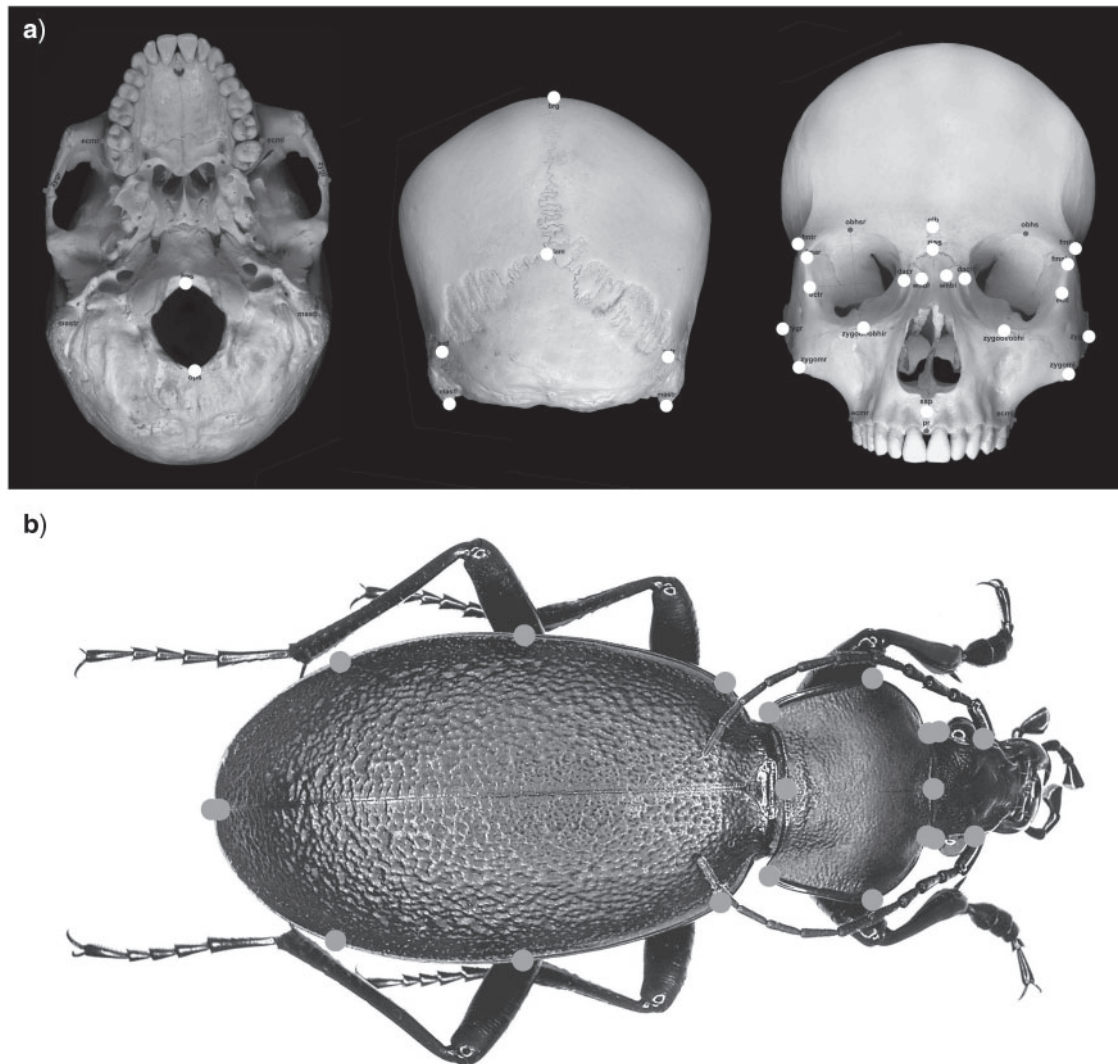
The data sets (Online Material available on Dryad at http://dx.doi.org/10.5061/dryad.mr2mh)

FIGURE 1.   Landmark configurations: a) humans (modified from the pdf help manual of 3d-id: http://www.3d-id.org/) and b) carabids (pictures modified from an original by Udo Schmidt CC BY-SA: https://commons.wikimedia.org/wiki/File:Carabus_coriaceus_%28Linn%C3%A9,_1758%29_%283093276668%29.jpg).

represent distant taxonomic groups (vertebrates and invertebrates), different types of data (3D and 2D) and anatomical structures, and both micro- and macroevolutionary levels of analysis. The first example (Ross 2014, http://www.3d-id.org/forensic-3d-coordinates) originates from a larger sample and a bigger configuration of 3D human cranial landmarks used in forensics. This intraspecific data set consists of 138 American men of European ancestry with seven midplane landmarks and 10 paired landmarks (Fig. 1a). Landmarks were selected from the original set of anatomical points so that the sample had no missing data. The second example is 2D data from 32 different genera of Carabidae and a total of 40 different species from a larger sample used in a macroevolutionary ecomorphological analysis (Panisi 2015). Each species is represented by a single adult male whose body form was measured with four midplane landmarks and 10 bilateral ones (Fig. 1b), of which three are sliding

semilandmarks (slid using a minimum Procrustes distance criterion, Rohlf 2015).

For each sample, five sets of landmarks were analyzed:

(1) Total configuration (TC) with all landmarks (midplane together with left and right bilateral points).

(2–3) Midplane landmarks [same ones as in (1)] and either left (LS) or right (RS) side paired landmarks.

(4–5) The same left- and right-side data as in (2–3) but now with the missing side estimated by mirror reflection of the paired landmarks. Mirror reflection is a common method to estimate missing landmarks in fossils with object symmetry (Gunz et al. 2009; Gunz and Mitteroecker 2013) but has been used also in modern species (e.g., Cardini et al. 2010). After

mirror reflection, the data (whose names will be abbreviated as "mirrored LS" and "mirrored RS") have the same number of landmarks as TC. However, they are in fact based on information from only one or the other side plus the common midplane landmarks.

EXPLORING THE CONGRUENCE OF RESULTS FROM ONE-SIDE-ONLY DATA AND DATA USING BOTH SIDES

The first step of the analysis, after creating the different one-side-only sets of landmarks, was superimposing each of them using separate Procrustes analyses (Rohlf and Slice 1990). Using separate superimpositions is a requirement to be in the same analytical setting as in the real case of an operator having measured either both sides or just one of the two. The five data sets could, in fact, be superimposed all together, which would be straightforward for TC and the mirrored data but could be achieved also for LS and RS, by coding the data on the other side as missing landmarks and using a software which can handle this type of data (e.g., Morpheus et al., Slice 1999). A common superimposition would have the advantage of bringing all data sets in the same shape space, making them apparently easier to compare. However, to also have the same dimensionality in all sets of data, one would need to estimate the missing side and that, for the specific purpose of this study, should be done in a way such that LS and RS distances are the same as when they are superimposed on their own one at a time, as in a typical one-side-only analysis. One simple estimate that achieves this is to replace the missing side with the average of TC. Unfortunately, this whole operation, of common superimposition and mean substitution, introduces a bias (i.e., a consistent difference between data sets) and, therefore, does not really help to make them directly comparable. Alternatively, one could, at least for 2D data, employ a different superimposition method, such as Bookstein baseline registration (with or without sliding, Zelditch et al. 2004) but that too would not solve the issue of different dimensionality (due to the missing landmarks in LS/RS) and also would not replicate the conditions in which the vast majority of one-side-only studies are performed, which is by using a Procrustes superimposition.

After separate Procrustes superimpositions, the strength of the congruence in centroid size (henceforth, called simply "size" for brevity) and shape extracted from one-side-only data in relation to TC was assessed using correlations. Similarity relationships were quantified as distances among specimens in a sample. If one-side-only data are accurate, these distances should be either almost perfectly proportional or virtually identical to those from TC. It might help to stress at this point that one-side-only data are not simply halving the TC landmarks, as only paired landmarks are in fact halved, and that for results to be accurate (i.e., lead to the same conclusion as using TC) what matters is the

relative similarity of size and shape differences (which is what is implied by talking about proportionality and estimates of congruence based on correlations). For univariate size, the strength of the similarity can be assessed by a simple Pearson correlation between TC size and one-side-only size. For multivariate shape, the corresponding estimate is the correlation between matrices of pairwise Procrustes shape distances. In both cases, accurate data should have a correlation as close as possible to one.

Besides looking at size and shape separately, one can also analyse them together using a multivariate regression of shape coordinates on size to estimate allometry (Mitteroecker et al. 2013). The percentage of shape variance explained by centroid size (multivariate R square—R^2) can be used as an estimate of the magnitude of allometry. Accurate one-side-only data should produce R^2 identical or at least very close to the R^2 of TC.

Procrustes superimpositions were done in the TPS Series (Rohlf 2015). Correlational analyses and multivariate regressions were performed in NTSYSpc 2.3 (Rohlf 2013) and R (R Core Team 2014). As an alternative to these programs, these simple analyses can be easily replicated in MorphoJ (superimposition and regressions, Klingenberg 2011) and PAST (matrix correlation, Hammer et al. 2001). The reconstruction of the missing side in LS/RS can be done following the guidelines of Cardini et al. (2010) in Morpheus et al. (Slice 1999) in combination with a spreadsheet. This operation, however, can be done even more easily by taking advantage of programs like MatchPoint.exe (O'Higgins, available on request). This is a software originally designed to match two regions of the same structure using a subset of common landmarks. For the reconstruction by mirror reflection, one creates a replica of one side (e.g., LS), where the sign of one coordinate has been swapped, and then registers the original side and its mirror version in MatchPoint.exe using the midplane landmarks. One way or the other, it is advisable to first order the landmarks in TC so that first are the midplane points, then all those of one side and finally all those of the other side. Reordering of 2D or 3D landmarks, if needed, is easily accomplished in TPSUtil (Rohlf 2015).

NUMERICAL RESULTS

Centroid size estimated from one-side-only data was always very highly correlated to that of TC. In the human data set, correlations ranged from 0.958–0.965 (RS and LS, respectively) to 0.984–0.985 (mirrored RS and mirrored LS, respectively). In the carabid sample, correlations were 0.9996 in LS–RS and 1.000 in the same data sets after mirror reflection of the missing side. Using Procrustes shape distances, one-side-only correlations to TC were, ranked from the lowest to the highest: in humans, 0.894 (LS), 0.901 (mirrored LS), 0.908 (RS), 0.919 (mirrored RS); in carabids, 0.930–0.931 (RS and LS, respectively), 0.980 (mirrored RS), and 0.983
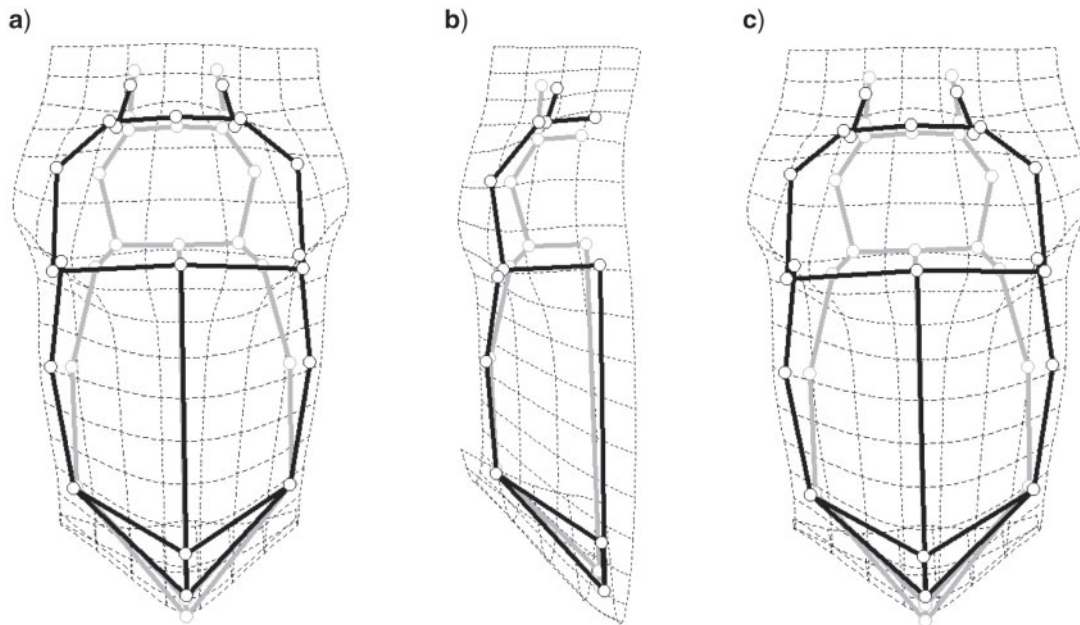
FIGURE 2. Carabid mean shape (gray) warped into the shape (black) corresponding to the positive end of PC1 (magnified 1.5 times): a) TC; b) LS; c) mirrored LS.

(mirrored LS). Neither humans nor carabids showed strong allometries. In humans, allometry accounted for just 1.3% ($P=0.0089$) of total shape variance in TC; in one-side-only data sets from the same sample allometric variance, ranked from the least to the most accurate in relation to TC, was 1.6% (RS), 1.6% (LS), 1.5% (mirrored RS), and 1.3% (mirrored LS). In carabids, the corresponding estimates of allometries were: TC, 4.2% ($P=0.1153$); LS, 5.3%; RS, 4.7%; mirrored RS, 3.8%; mirrored LS, 4.5%.

Figure 2 shows an example of visualization of shape variation using the carabid sample and both deformation grids and wireframes (Klingenberg 2013). The target shape (black) is the positive extreme of the first principal component (PC) of the shape coordinates and it is shown in relation to the sample mean shape (gray). The two shapes are shown one superimposed onto the other only to emphasize the differences between the three sets of data. The midplane of the start and target shapes overlaps almost perfectly in TC and mirrored LS. However, it looks transposed to the right in LS with the pronotum of the target apparently bending slightly laterally relative to its medially displaced abdomen. This is purely an artifact of the superimposition in the absence of bilateral landmarks. The visualization using RS and mirrored RS (not shown) shows the same exact pattern as in LS.

IS SIZE ACCURATE?

In terms of centroid size, one-side-only data sets were all quite accurate ($r > 0.95$). This was especially true for the carabids ($r \approx 1$), as expected in an interspecific sample where the largest species is more than 20 times bigger

than the smallest. With such a large range of variation, the relative variation in size is almost unaffected by the choice of using one or both sides. It is, in contrast, much more interesting that, even in the human data set, a much more homogeneous sample with the largest cranium just 15% bigger than the smallest, correlations were high, although smaller and more variable than in carabids. In both samples, reconstructing the missing side by mirror reflection consistently increased accuracy in size, but this was clearly more evident in the intraspecific human sample.

Overall, however, estimates of centroid size variation seem fairly robust to the lack of complete information. The issue is different but this finding is not dissimilar to that of a recent study (Cardini 2014) on the approximation of 3D structures using flat 2D pictures. That is a case in which inaccuracies can be huge, as not only some information (i.e., the third dimension) is not used, but also measurements are inevitably distorted by their projection onto a plane. Despite these issues, Cardini (2014) found that centroid size from pictures of marmot mandibles and crania accurately measured relative differences among specimens with a correlation of 0.987 or more to the corresponding estimates from 3D landmarks.

IS SHAPE ACCURATE?

The accuracy of one-side-only shape distances was lower than with size. Consistently with size, however, mirrored LS/RS did relatively better than LS/RS and carabids were more accurate ($r \approx 0.93$–0.98) than humans ($r \approx 0.90$–0.92). As it happened with size, the lower inaccuracy in carabids is again because relative

differences between one-side-only data and TC become smaller in the beetle sample, made of very different genera and species from a large and morphologically disparate family of insects. In contrast, as humans were all of the same sex, age class, ancestry, and geographical region, their shape distances were more strongly influenced by the choice of measuring one side instead of both.

As for size, I first contextualize shape results by comparing them with other sources of inaccuracy. The already mentioned (Cardini 2014) analysis of 2D approximations of 3D landmarks showed that 2D Procrustes shape distances had correlations with the more accurate 3D ones in the range of ca. 0.5–0.8. Using both analyses of variances and cluster analyses, that study suggested that at least the 2D data sets with the highest correlations (≈ 0.7–0.8) were able to capture similarity relationships in shape almost as faithfully as 3D landmarks. For instance, with the relatively flat mandibles, 85% of the individuals measured in 2D clustered together with their 3D replica. Thus, judged on this "scale," even the smallest matrix correlation for one-side-only data ($r \approx 0.90$) seems to imply an accurate representation of shape variation.

Dimensionality reduction in multivariate data is another example of potential inaccuracy because of a loss of information. This is probably even more common in GMM than using 2D pictures to measure 3D anatomical features and is generally achieved by discarding several of the higher order PCs of the shape coordinates. Matrix correlations between Euclidean distances in reduced PC spaces and Procrustes shape distances provide some information on how well similarity relationships are preserved by different subsets of first PCs. Fadda and Corti (2000) adopted this method to decide how many PCs to retain in a study of cranial variation in the Ethiopian *Myomys–Stenocephalemys* complex. In a sample of 106 specimens, measured using 33 3D landmarks, they argued that shape relationships were accurately represented by the first 15 PCs out of a total of 98. Pairwise Euclidean distances based on these 15 PCs had a matrix correlation to the corresponding Procrustes distances of 0.993 and explained approximately 80% of total variance. Following Fadda and Corti (2000), Cardini et al. (2010) selected the first 20 PCs to summarize most shape variation in a sample of 122 African monkey skulls with 86 3D landmarks. Those 20 PCs had a correlation of 0.986 with Procrustes shape distances and accounted for 76% of total shape variance. This suggests that matrix correlations used to compare shape spaces can be very high despite having discarded 1/4 to 1/5 of total variance. Indeed, this is true also for TC in the two samples used in this study. If higher order PCs are removed until correlations between TC Procrustes shape distances and distances in the reduced TC shape space are of the same magnitude as those of one-side-only data, about 20% to almost 50% of total variance is lost. More precisely, in humans, correlations of 0.90–0.93 would be obtained by including, respectively, only the first 8–10 PCs of TC (out of a total of 74), which explain

about 50–60% of total shape variance. In the carabid data set, correlations of 0.93–0.98 would correspond approximately to those of the shape sub-space of the first 3–4 PCs of TC (out of a total of 27), which explain about 70–80% of variance. In this perspective, as crude and approximate as it might be, the magnitude of the shape information lost by using just one side seems much larger.

A third way of looking at how one-side-only data relate to TC, and obtain partial but less approximate estimates of their differences, is to just focus on mirrored LS/RS. As the mirror reflection reconstructs the missing side, the configuration becomes the same as in TC and the three data sets can be analyzed in a common shape space. In both humans and carabids, this operation leaves shape distances within each data set virtually unchanged, which means that similarity relationships are the same as if they had been superimposed separately. Having data in a common shape spaces makes it possible to partition the total sum of squares between individual and error components, as one would do in an ANOVA to test measurement error (Viscosi and Cardini 2011). The error component represents the differences among the data sets. The *P* value is not considered, because the data are not independent. Thus, only the percentages of variation accounted for by the two components are calculated. Overall, differences among the three sets of data accounted for 10.9% of total sum of squares in humans and 3.6% in the carabids. If data were analyzed pairwise (i.e., TC vs. either mirrored LS or mirrored RS), percentages were slightly lower (ca. 7–9% for humans and 2% for carabids). One can, therefore, conclude that using one-side-only data the loss of information on the complete structure is indeed small in the beetle interspecific data set but can be as high as about 1/10 of total variance in the intraspecific sample of humans. These estimates are much less than suggested by the comparison with the variance discarded in higher order PCs. However, one has to bear in mind that they were possible only for the two most accurate one-side-only data sets (mirrored LS/RS). LS and RS (without mirror reflection of bilateral landmarks) would certainly perform less well and the loss of information in humans would probably be larger than 10%.

Are Allometry and the Visualization Accurate?

When size and shape were analyzed together to estimate allometry, results were largely congruent with those of the correlational analyses of size and shape on their own but somewhat less interesting to look at, because of the small magnitude of allometric variation in both data sets. The mirror reflection of bilateral landmarks improved accuracy making the percentages of allometric variance very close to those of TC but, overall, absolute differences were quite small. Nevertheless, if differences of one-side-only data sets to TC were expressed in relative terms, one-side-only analyses led to under- or overestimates of allometric

variance ranging from 10% to more than 20% of the value observed using both sides.

For the visualization, I chose to use the carabid data set as in 2D it is easier to see the potentially misleading midplane displacement. It is important to notice that this displacement is observable with just a very small magnification of the observed differences (1.5 times the largest PC1 score) and this happens despite using the example data set with the strongest congruence between one-side-only data and data based on both sides. It is also obvious that, after mirror reflection of the bilateral landmarks, the visualization is almost identical to that obtained in TC. Although the conclusions on shape differences between the start and target shape may be about the same, TC and mirrored LC look much more similar, as well as plausible and straightforward to interpret: the pronotum becomes bigger in all directions, whereas the elytra and abdomen become relatively larger and shorter and the eyes are positioned a bit closer to the pronotum. The literature offers many others examples of poor visualization using just half of a structure with object symmetry and that in itself had been previously suggested as a good reason for estimating the missing side by mirror reflection (Cardini et al. 2010).

PREMISES, CONCLUSIONS AND RECOMMENDATIONS

Before trying to draw cautious and clearly preliminary conclusions from these two examples data sets, I want to summarize the premises (1–2) and main question (3) of this study.

(1) The inspiration came from a basic observation: GMM studies using Procrustes methods are often performed using just one side of structures with object symmetry. This has been a common approach since the early days of GMM in biology and has not lost its popularity. However, what the exclusion of one side of a symmetric object implies and whether that may appreciably affect results have not been investigated.

(2) In all the studies I know, when only one side was analyzed, that was not because the hypotheses being tested were specific to that side. It was either because the other side was missing, for instance, in palaeontological studies, or, more simply and more frequently, as a shortcut to speed up data collection and reduce the number of variables. The implicit assumption was that asymmetries are small and the information on the other side is redundant and, therefore, unimportant, unless one is specifically studying asymmetric variation.

(3) Starting from these premises, I asked the question: do one-side-only data really provide an accurate representation of the similarity relationships one would infer using total structures complete of all bilateral landmarks? The answer is inevitably specific to the data sets I analyzed and the following "conclusions" (in quotation marks!) should be seen as a way to open the discussion and stimulate further studies:

(a) Size variation can be accurately estimated using just one side. This is suggested by all analyses and data sets, and is in good agreement with previous studies on measurement error (Cardini 2014). However, the relative error will be bigger in microevolutionary studies, and one might speculate whether it could become more important with less landmarks and larger asymmetry. In fact, this same question could be asked also for shape.

(b) Shape is quite accurate in the macro-evolutionary sample. However, in the intraspecific sample, the congruence with shape distances from the TC is smaller and the loss of information might be comparable to the one produced by discarding higher order PCs which account for 20% or more of TC shape variance. Even if that is likely to be an overestimate, differences could still account for about 1/10 or more of total shape variation in humans, as suggested by the common shape analyses of TC and mirrored LS/RS.

(c) The error in estimating allometric variance can be relatively large at least when allometry is modest.

(d) In terms of the visualization, not only the mirror reflection makes it more realistic but also avoids implausible landmark displacements perpendicular to the mid-plane. This type of displacement makes shape diagrams look odd and it is potentially misleading.

(e) In all examples and data sets, one-side-only data whose missing side is reconstructed by mirror reflection were more accurate than the corresponding analyses based on the same side without mirror reflection. As anticipated, because LS and mirrored LS (or RS and mirrored RS) are based on the same original raw data, their difference in accuracy must be mainly related to the superimposition. A likely effect of superimposing LS (or RS) without mirror reflection is that there is more variation perpendicular to the midplane, because there are no points on the other side to "constrain" it. This is the same problem I mentioned talking about the visualization (d). If a structure is symmetric, and asymmetries are small, variation on one side cannot be independent of variation on the other side, because they are built during

development to be mirror images. When only one side is measured, that information is lost. However, the mirror reflection of the paired landmarks, in a sense, reintroduces some of the covariation between the two sides, which might contribute to make shape data more similar to those obtained by landmarking both sides. The variance–covariance structure will not be the same as in the total data sets, as some of that information is simply missing. However, if Procrustes shape distance matrices are more similar to that of the complete landmark data after reconstruction by mirror reflection, that implies a better approximation of the variance–covariance structure as well. In fact, the operation of reconstruction by mirror reflection might also help when semilandmarks are present, as their sliding should take symmetry into account (Bookstein 2014) and that is clearly impossible in LS/RS without mirrored bilateral points.

(f) Based on the previous point, one could make an apparently obvious recommendation: this is that, unless one has a specific hypothesis that needs to be tested on just one side, mirror reflection of the missing side should be a customary operation before analyses of one-side-only data. Indeed, preliminary results from a larger study (Cardini, manuscript in preparation), using 10 different cranial data sets from five mammalian orders (rodents, carnivores, primates, cetarctiodactyls, and diprotodont marsupials) and a variety of taxonomic levels (from intraspecific to infraordinal), support this conclusion and suggest that accuracy, as well as visualization, may indeed be often improved by mirror-reflecting bilateral landmarks to reconstruct the missing side.

(g) There might be exceptions to this last (f) main conclusion, however. Likely they will be those cases in which asymmetry is large, which could happen because asymmetry is really big or it is inflated by measurement error. For instance, in a small ($N = 26$) intraspecific copepod data set (which is a subsample from a larger study published during the review of this article; Karanovic et al. 2016) with 19 paired cephalothoracic landmarks and only two miplane ones, and an unusually high asymmetry (36.3% of total shape variance; 35.1% if only fluctuating asymmetry is considered), one-side-only data were consistently less accurate after, rather than before, reconstructing the missing sides by mirror reflection. In fact, correlations of shape distances with TC dropped from about

0.8 (LS–RS) to 0.3 (mirrored LS or mirrored RS) after mirroring the missing side.

(i) With large directional asymmetry, using only one side may simply not be advisable, as the principal justification for analyzing only one or the other side becomes untenable: with big directional differences, the information provided by the two sides will be simply different. One could still obtain results in good agreement with TC, however, if the directional asymmetry acts like a bias, which is consistent in magnitude and direction across all samples.

(ii) With strong fluctuating asymmetry, the effect of analyzing just one side is less predictable. As with directional asymmetry, fluctuating asymmetry might be large either because it is real or because of measurement error, but the differences, regardless of their origin, are nonsystematic. With this type of random differences between sides, one-side-only data will be a poor proxy for TC. To understand why, one can imagine an extreme case in which differences between sides are close in magnitude to differences among individuals: if that happens, similarity relationships will be remarkably different depending on the side being analyzed. Moreover, with strong fluctuating asymmetry, because sides are so different, mirror-reflected LS or RS could be as inaccurate as or even more inaccurate than LS and RS. This might have been the case of the copepod data set, but did not happen in both samples analyzed in this study. In fact, even in humans, where asymmetry was fairly large (18.3% of total variance) and mostly related to fluctuating asymmetry (14.8%), one-side-only data, especially after mirror reflection of the missing side, were quite accurate. Nevertheless, one cannot make any generalization from this and, besides the amount of asymmetry, the number and dimensionality of the landmark data might have an effect on how well (or poorly) one-side (with or without reconstruction of the missing landmarks by mirroring) approximates results from the landmark configuration complete with both sides. Broader empirical studies as well as simulations (e.g., using different number of landmarks and semilandmarks in 2D/3D, testing the effect of sample size and varying the amount and type of asymmetry) will

be required to better understand when one-side-only data are suitable and how to best treat them to improve accuracy.

Is Then Really Advisable to Perform One-Side-Only Analyses?

Using humans and carabids, I showed that the type of information being extracted (size or shape) and the magnitude of the variation in a sample (microevolutionary vs. macroevolutionary analyses) might affect the outcome of one-side-only analyses. I also discussed another potentially important component, asymmetry, which might need to be considered and may interact with other factors. Indeed, how appropriate one-side-only data are will be largely specific to the study samples and hypotheses, and should thus be explored in preliminary analyses: both sides could be measured in a subsample and the data used for assessing (Klingenberg et al. 2002) the magnitude of asymmetry, its nature (directional/fluctuating), and whether one-side-only data may be improved by mirror-reflecting bilateral landmarks.

At the very end, I would like to go back to one important point, which had been briefly mentioned before but becomes even more relevant after showing some of the potential issues of one-side-only approaches. This is why, in the first place, one might want to use only one side when using both, if possible, produces data which are more accurate, have a better visualization and potentially allow to quantify asymmetries. In the published GMM literature, sometimes no reasons are given at all. When there is a justification, as anticipated, that is either that by measuring only one side one avoids redundant data or that it speeds up data collection.

The first explanation does not really matter. As I have explained, there might be a degree of redundancy because of symmetry but, even when asymmetry is small and not the subject of the study, bilateral landmarks also capture information on the covariance of the two sides. That information is lost using just one side but potentially partly recovered by reconstructing the other side by mirror reflection. Also, using PCs instead of shape coordinates in parametric tests will solve potential issues with the number of variables and the accuracy of the degrees of freedom (Viscosi and Cardini 2011). In fact, if the problem was simply to avoid redundancy, one should follow Klingenberg et al. (2002), measuring both sides but only using the symmetric component or its PCs. That would be equivalent to averaging structures with matching symmetry (e.g., the left and right hands) to increase accuracy but have the advantage of preserving the relationships between paired landmarks in the analysis and visualization of object symmetry.

The second point on the efficiency of data collection, however, is more interesting and bears another question: is speeding up data collection any useful or just indicates lack of care and effort on the operator side? I would argue that, when sample size is crucial (Cardini and Elton 2007; Cardini et al. 2015), and funds and time are limited, one-side landmarking may have a strong justification. To make this point less abstract, I take as an example my own 2004–2005 data collection on Old World monkeys. The main aim of that data collection was to build a big reference database for modeling evolutionary divergence within and between species with no specific interest in asymmetries. This is a type of application which may take many different forms but is the kind of work that lies at the heart of many studies in systematics: from biogeographical analyses (Cardini et al. 2010, and references therein) to ecomorphological investigations and studies of the tempo and mode of adaptive radiations (Cardini and Elton 2008a, 2008b). For that project, in 123 working days, I measured 3547 skulls using 86 landmarks on the left side only. This means a total of 305,042 landmarks and ca. 29 specimens/day or 2480 landmarks/day. Even without taking into account the time for repositioning each specimen (which may be partially counterbalanced by the lack of need of reselecting a specimen and taking it out of its drawer), if landmarks on both sides had been measured, each specimen would have had 155 landmarks; at a speed of little less than 2500 landmarks/day, that translates into an average of 16 specimens/day for a total of slightly less than 2000 skulls in 123 days. Thus, even if landmarking both sides would have been desirable, that would have reduced the total sample size by 45%. As in this case, it is very likely that many other researchers might have to choose between a modest increase in accuracy and a significantly larger sample size, and, if sample size is crucial (Cardini and Elton 2007; Cardini et al. 2015), and especially if differences are fairly large and asymmetry small, one-side-only landmarking, with the reconstructing of the missing side by mirror reflection, could be the best choice.

Supplementary Material

Acknowledgments

References

Adams D.C., Rohlf F.J., Slice D.E. 2004. Geometric morphometrics: ten years of progress following the "revolution". Ital. J. Zool. 71:5–16.

Adams D.C., Rohlf F.J., Slice D.E. 2013. A field comes of age: geometric morphometrics in the 21st century. Hystrix Ital. J. Mammal. 24:7–14.

Álvarez A., Perez S.I., Verzi D.H. 2015. The role of evolutionary integration in the morphological evolution of the skull of caviomorph rodents (Rodentia: Hystricomorpha). Evol. Biol. 42:312–327.

Bookstein F.L. 2014. Measuring and reasoning: numerical inference in the sciences. New York: Cambridge University Press.

Cardini A. 2003. The geometry of the marmot (Rodentia: Sciuridae) mandible: phylogeny and patterns of morphological evolution. Syst. Biol. 52:186–205.

Cardini A. 2014. Missing the third dimension in geometric morphometrics: how to assess if 2D images really are a good proxy for 3D structures? Hystrix Ital. J. Mammal. 25:73–81.

Cardini A., Elton S. 2007. Sample size and sampling error in geometric morphometric studies of size and shape. Zoomorphology 126:121–134.

Cardini A., Elton S. 2008a. Does the skull carry a phylogenetic signal? Evolution and modularity in the guenons. Biol. J. Linn. Soc. 93:813–834.

Cardini A., Elton S. 2008b. Variation in guenon skulls (II): sexual dimorphism. J. Hum. Evol. 54:638–647.

Cardini A., Filho J.A.F.D., Polly P.D., Elton S. 2010. Biogeographic analysis using geometric morphometrics: clines in skull size and shape in a widespread African arboreal monkey. In: Elewa A.M.T., editor. Morphometrics for nonmorphometricians. Berlin, Heidelberg: Springer. p. 191–217.

Cardini A., Jansson A., Elton S. 2007. A geometric morphometric approach to the study of ecogeographical and clinal variation in vervet monkeys. J. Biogeogr. 34:1663–1678.

Cardini A., Loy A. 2013. On growth and form in the "computer era": from geometric to biological morphometrics. Hystrix Ital. J. Mammal. 24:1–6.

Cardini A., Polly P.D. 2013. Larger mammals have longer faces because of size-related constraints on skull form. Nat. Commun. 4:2458.

Cardini A., Seetah K., Barker G. 2015. How many specimens do I need? Sampling error in geometric morphometrics: testing the sensitivity of means and variances in simple randomized selection experiments. Zoomorphology 134:149–163.

Cardini A., Thorington R.W. 2006. Postnatal ontogeny of marmot (rodentia, sciuridae) crania: allometric trajectories and species divergence. J. Mammal. 87:201–215.

Drotz M.K. 2003. Speciation and mitochondrial DNA diversification of the diving beetles Agabus bipustulatus and A. wollastoni (Coleoptera, Dytiscidae) within Macaronesia. Biol. J. of Linn. Soc. 79:653–666.

Drotz M.K., Brodin T., Saura A., Giles B.E. 2012. Ecotype Differentiation in the Face of Gene Flow within the Diving Beetle Agabus bipustulatus (Linnaeus, 1767) in Northern Scandinavia. PLoS ONE. 7:e31381.

Fadda C., Corti M. 2000. Three dimensional geometric morphometric study of the Ethiopian *Myomys-Stenocephalemys* complex (Murinae, Rodentia). Hystrix Ital. J. Mammal. 11:131–143.

Gunz P., Mitteroecker P. 2013. Semilandmarks: a method for quantifying curves and surfaces. Hystrix Ital. J. Mammal. 24:103–109.

Gunz P., Mitteroecker P., Neubauer S., Weber G.W., Bookstein F.L. 2009. Principles for the virtual reconstruction of hominin crania. J. Hum. Evol. 57:48–62.

Hammer O., Harper D., Ryan P. 2001. PAST: paleontological statistics software package for education and data analysis. Paleontol. Electron. 4(1):1–9.

Head J.J., Polly P.D. 2015. Evolution of the snake body form reveals homoplasy in amniote Hox gene function. Nature 520:86–89.

Hopkins M.J., Lidgard S. 2012. Evolutionary mode routinely varies among morphological traits within fossil species lineages. Proc. Natl Acad. Sci. 109:20520–20525.

Ivanović A., Sotiropoulos K., Džukić G., Kalezić M.L. 2009. Skull size and shape variation versus molecular phylogeny: a case study of alpine newts (*Mesotriton alpestris*, Salamandridae) from the Balkan Peninsula. Zoomorphology 128:157–167.

Karanovic T., Djurakic M., Eberhard S.M. 2016. Cryptic species or inadequate taxonomy? Implementation of 2D geometric morphometrics based on integumental organs as landmarks for delimitation and description of copepod taxa. Syst. Biol. 65:304–327.

Klingenberg C.P. 2011. MorphoJ: an integrated software package for geometric morphometrics. Mol. Ecol. Res. 11:353–357.

Klingenberg C.P. 2013. Visualizations in geometric morphometrics: how to read and how to make graphs showing shape changes. Hystrix Ital. J. Mammal. 24:15–24.

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

Loy A., Corti M., Marcus L.F. 1993. Landmark data: size and shape analysis in systematics. A case study on Old World Talpidae (Mammalia, Insectivora). In: Marcus L.F., Bello E., Valdecasas, A.G., editors. Contributions to morphometrics. Madrid: Consejo Superior de Investigaciones Cientificas (Monografias Museo de Ciencias Naturales; vol. 8). p. 215–240.

Macagno A.L.M., Pizzo A., Rolando A., Palestrini C. 2011. Size and shape interspecific divergence patterns partly reflect phylogeny in an Onthophagus species-complex (Coleoptera: Scarabaeidae). Zool. J. Linn. Soc. 162:482–498.

Meloro C., Hudson A., Rook L. 2015. Feeding habits of extant and fossil canids as determined by their skull geometry. J. Zool. 295:178–188.

Mitteroecker P., Gunz P., Windhager S., Schaefer K. 2013. A brief review of shape, form, and allometry in geometric morphometrics, with applications to human facial morphology. Hystrix Ital. J. Mammal. 24:59–66.

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

Monteiro L.R., Duarte L.C., dos Reis S.F. 2003. Environmental correlates of geographical variation in skull and mandible shape of the punaré rat Thrichomys apereoides (Rodentia: Echimyidae). J. Zool. 261: 47–57.

Panisi, M. 2015. *Morphoscape*: 'paesaggi' morfologici per lo studio della biodiversità. Tesi di Laurea, Università di Modena e Reggio Emilia.

Pečnerová P., Moravec J.C., Martínková N. 2015. A skull might lie: modeling ancestral ranges and diet from genes and shape of tree squirrels. Syst. Biol. 64:1074–1088.

Piras P., Colangelo P., Adams D.C., Buscalioni A., Cubo J., Kotsakis T., Meloro C., Raia P. 2010. The *Gavialis–Tomistoma* debate: the contribution of skull ontogenetic allometry and growth trajectories to the study of crocodylian relationships. Evol. Dev. 12:568–579.

R Core Team. 2014. R: a language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing.

Rohlf F.J. 2013. NTSYSpc: numerical taxonomy system. Version 2.3. Setauket, New York: Exeter Publishing, Ltd.

Rohlf F.J. 2015. The tps series of software. Hystrix Ital. J. Mammal. 26:9–12.

Rohlf F.J., Loy A., Corti M. 1996. Morphometric analysis of Old World Talpidae (Mammalia, Insectivora) using partial-warp scores. Syst. Biol. 45:344–362.

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

Rohlf F.J., Marcus L.F. 1993. A revolution morphometrics. Trends Ecol. Evol. 8:129–132.

Rohlf F.J. 1998. On Applications of Geometric Morphometrics to Studies of Ontogeny and Phylogeny. Syst. Biol. 47:147–158.

Ross, A.H. 2014. Forensic 3D database. http://www.3d-id.org/forensic-3d-coordinates.

Ruane S. 2015. Using geometric morphometrics for integrative taxonomy: an examination of head shapes of milksnakes (genus Lampropeltis). Zool. J. Linn. Soc. 174:394–413.

Rufino M.M., Abelló P., Yule A.B. 2006. Geographic and gender shape differences in the carapace of *Liocarcinus depurator* (Brachyura: Portunidae) using geometric morphometrics and the influence of a digitizing method. J. Zool. 269:458–465.

Savriama Y., Klingenberg C. 2011. Beyond bilateral symmetry: geometric morphometric methods for any type of symmetry. BMC Evol. Biol. 11:280.

Slice D. 1999. Morpheus et al. (beta version). Department of Ecology and Evolution, State Univ. New York: Stony Brook.

Tabatabaei Yazdi F., Adriaens D. 2013. Cranial variation in Meriones tristrami (Rodentia: Muridae: Gerbillinae) and its morphological comparison with Meriones persicus, Meriones vinogradovi and Meriones libycus: a geometric morphometric study. J. Zoolog. Syst. Evol. Res. 51:239–251.

Viscosi V., Cardini A. 2011. Leaf morphology, taxonomy and geometric morphometrics: a simplified protocol for beginners. PLoS One 6:e25630.

Yazdi F.T., Colangelo P., Adriaens D. 2015. Testing a long-standing hypothesis on the relation between the auditory bulla size and environmental conditions: a case study in two jird species (Muridae: Meriones libycus and M. crassus). Mammalia 79:185–200.

Zelditch M., Swiderski D., Sheets D., Fink W. 2004. Geometric morphometrics for biologists: a primer. Waltham (MA): Elsevier Academic Press.