Geometric morphometrics out-perform linear-based methods in the taxonomic resolution of a mammalian species complex

Pietro Viacava[1,2*], Simone P. Blomberg[3], Vera Weisbecker[1,2*]

[1]College of Science and Engineering, Flinders University, Adelaide, South Australia, Australia

[2]Australian Research Council Centre of Excellence for Australian Biodiversity and Heritage

[3]School of Biological Sciences, The University of Queensland, St Lucia, Queensland, Australia

Correspondence authors: Pietro Viacava (pietroviama@hotmail.com) and Vera Weisbecker (vera.weisbecker@flinders.edu.au).

Running headline: Morphometric methods for taxonomic resolution

Abstract

1. Morphology-based taxonomic research frequently applies linear morphometrics (LMM) in skulls to quantify species distinctions. The choice of which measurements to collect generally relies on the expertise of the investigators or a set of standard measurements, but this practice may ignore less obvious or common discriminatory characters. In addition, taxonomic analyses often ignore the potential for subgroups of an otherwise cohesive population to differ in shape purely due to size differences (or allometry). Geometric morphometrics (GMM) is more complicated as an acquisition technique, but can offer a more holistic characterization of shape and provides a rigorous toolkit for accounting for allometry.

2. In this study, we used linear discriminant analysis to assess the discriminatory performance of four published LMM protocols and a 3D GMM dataset for three clades of antechinus known to differ subtly in shape. We assessed discrimination of raw data (which are frequently used by taxonomists); data with isometry removed; and data after allometric correction.

3. We found that group discrimination among raw data was high for LMM, possibly inflated relative to GMM when visualised in PCA plots. However, GMM produced better results in group discrimination after the size and allometry treatments. High measurement redundancy in LMM protocols appears to result in relatively high allometry but low discriminatory performance.

4. These findings suggest that taxonomic measurement protocols might benefit from GMM-based pilot studies, because this offers the option of differentiating allometric and non-allometric shape differences between species, which can then inform on the development of the easier-to-apply LMM protocols.

Keywords: allometry – cryptic species – geometric morphometrics – linear discriminant analysis – linear morphometrics – shape variation – taxonomy

Introduction

Morphometric measurements are an important tool in efforts to differentiate mammalian species from each other, and have been used in taxonomic research for centuries. Mammalian skulls in particular are widely used for taxonomic diagnostics and have long provided important data which can be used in the delimitation of species or Evolutionary Significant Units (ESUs). Cranial morphometric measurements are widely used in the separation of closely related mammalian groups around the world, ranging across disparate taxa such as rodents (Alhajeri, 2021), mustelids (Abramov et al., 2018; Gálvez-López et al., 2022) and whales (Rosel et al., 2017).

Morphometrics-based taxonomic differentiation remains mostly the domain of linear morphometrics (LMM) (Jackson & Groves, 2015), mainly because linear measurements are easily taken and comparable to results from past studies. However, this approach has several limitations. For example, taxonomists choose a set of linear distances based on their expertise of the morphology of the taxon in question. This could pose a problem when these linear protocols are not standardized among morphometricians, obtaining potentially different acquisitions of data in studies of similar taxa but based on different protocols. In addition, the linear distances measured in taxonomic diagnoses often include maximum and minimum heights, widths and lengths that are easily identifiable to the eye. However, these measurements may characterize conditions that are not necessarily biologically homologous across taxa. This is because, when shapes differ, maximum or minimum distances may not be comparable among individuals, as the point-to-point distance may relate to different reference points that are not necessarily homologous (Figure 1). Additionally, LMM describes the distance between two points, missing potentially relevant variation between shapes and not retaining information about the original shape. Lastly, size

differences are not often accounted for in taxonomic studies using LMM, but there is a problem of redundancy when linear measurements contain other linear measurements within them (e.g., multiple measurements along the longitudinal axis of the skull). Similarly, the frequent use of proportional ratios (e.g. skull length vs. width) is problematic because many species display intraspecific allometry, such that genetically similar individuals will differ in a ratio simply because they differ in size (Sidlauskas et al., 2011).

A potential refinement of LMM protocols could be offered by the use of Geometric Morphometrics (GMM). GMM uses the coordinates of anatomical reference points as identifiable shape variables in most or all the specimens in a given dataset. Since the 1990s, this technique (Adams et al., 2004) has become the standard method for shape characterization in evolutionary and ecological morphometrics and has a very mature analytical toolkit (Fruciano, 2016; Fruciano et al., 2017). The holistic characterization of the biological specimens under study and the graphical output of the associated shape changes provide a powerful and alternative tool for biological inferences in many contexts (Klingenberg, 2016; Stone, 1997), a definite advantage that linear morphometrics cannot offer.

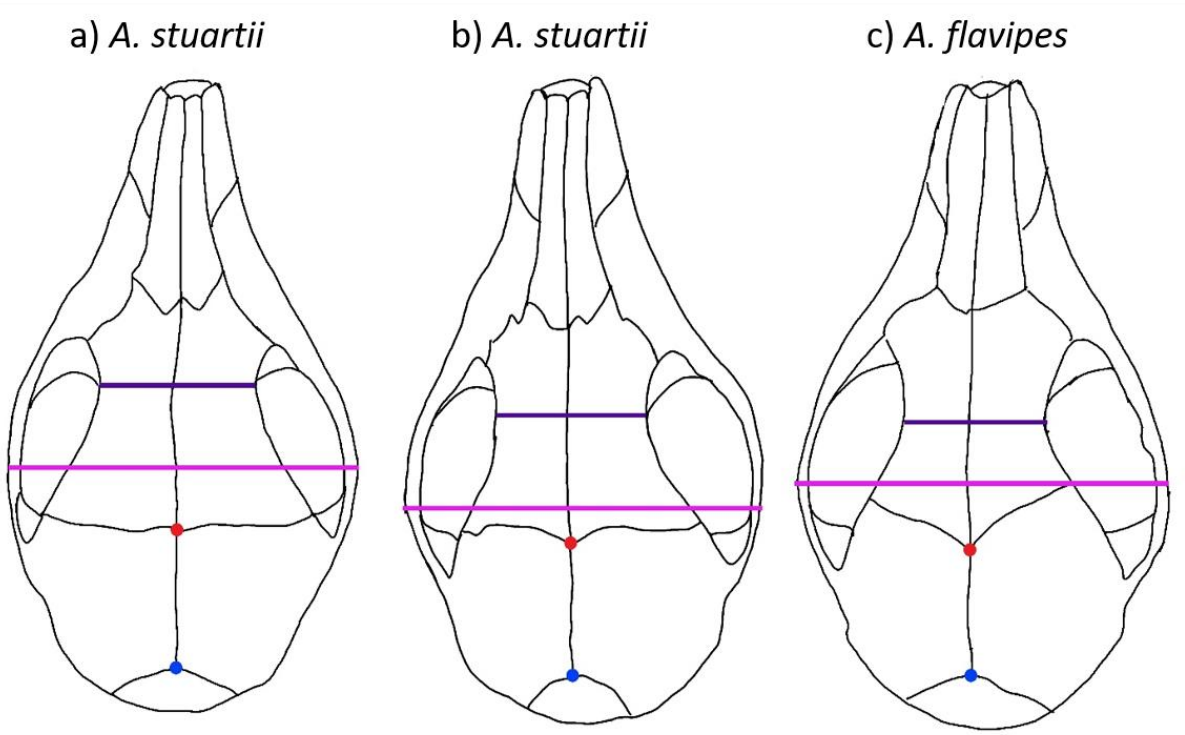


Figure 1: Three specimens (CM12785, CM6540 and CM10548) outlining two commonly used linear distances: the width of greatest constriction of orbitotemporal fossa (dark purple) and the maximum width of cranium measured across zygomatic arches (pink). In addition, two type I homologous landmarks (by suture intersection) are depicted: the fronto-parietal suture in midline (red dot) an the parietal-interparietal suture in midline (blue dot). The two examples of maximum and minimum distances are measured at different anatomical positions relative to the homologous landmarks and other sutures in the skull, indicating a possibly serious lack of homology.

GMM is rarely used in taxonomic studies, probably because linear measurements are easier to acquire and analyse, and because there is a large body of literature on taxonomic measurements that has been widely used for centuries (Sidlauskas et al., 2011). In addition, geometric morphometric data acquisition can be more complex, requiring digitisation of either photographs or 3D specimen representations, which generally involve specialized equipment. The statistical analyses required, while very well developed and versatile, are also specialised and involve high-dimensional data (Adams & Otárola-Castillo, 2013; Klingenberg, 2011; Zelditch et al., 2012). Therefore, they may not be perceived to be straightforward as the statistical toolkits used in LMM analyses.

We argue that GMM analyses are a useful addition to taxonomic studies because the technique addresses several of the issues of LMM raised above. First, the analyses of the main variation of a dataset based on a GMM protocol that optimizes anatomical coverage could inform a LMM protocol to include linear distances with potential taxonomic differentiators. Second, GMM can directly address primary homology by using fixed homologous landmarks (e.g., suture intersections) and curve and surface semilandmarks that correspond anatomically to each other (Gunz & Mitteroecker, 2013; Palci & Lee, 2019; Zelditch et al., 2012). This contrasts to LMM, where often

measurements are chosen by the maximum and minimum distance from one point to another regardless of whether they correspond anatomically. This LMM approach results in distances potentially measured at different places relative to homologous reference points such as tissue juxtapositions (Figure 1). Third, the distance between two points usually measured in LMM protocols does not provide information on the geometry (i. e., curvature) of the line drawn between them. GMM can use semilandmarks along curves or surfaces that describe the shape between points in three dimensions, augmenting other potential maximum or minimum distances to be taken into account in LMM protocols.

An additional advantage for GMM is the explicit treatment of size. The Procrustes superimposition procedure inherent to GMM (Zelditch et al., 2012) allows the removal of the size component from the dataset by scaling all specimens to the same size. This procedure results in two components: a proxy for size called centroid size and a multivariate shape component (Kendall, 1989). These can then be used for analyses of allometry (shape changes disproportionate to size) in the form of a shape vs size regression (Klingenberg, 2016, 2022). This substantially improves on the issue of accounting for isometric and allometric variation, which can have serious implications for taxon delimitation. Allometric

effects in particular can give an impression of species differentiation, when cranial allometry is generally present within most mammalian species (Cardini et al., 2015; Marcy et al., 2020; Viacava et al., 2020, 2021) and may not be related to morphological divergence due to a speciation event (Sidlauskas et al., 2011). Such allometric variation has been regarded as irrelevant to taxonomy. This is because, if shape differences were strictly due to size differences, they are likely to be the differences between small and large animals within a taxonomic group (Pilbeam & Gould, 1974; Seifert, 2008; Wood & Stack, 1980). In contrast, non-allometric shape changes are thought to be caused by independent adaptive processes such as are involved in species divergences (Huxley, 1931; Gould, 1975). Thus, it is recommended to include allometric analyses in taxonomic studies in order to interpret the shape variation and to properly delimit species (Cardini & Polly, 2013; Kaliontzopoulou et al., 2008; Outomuro & Johansson, 2017; Seifert, 2008; Sidlauskas et al., 2011; Yazdi, 2014). However, even in a case where taxonomic differentiation is driven purely by selection for size and coincides only with allometry effects, this represents important information on the differentiation process and should be considered.

Here, we compare the taxonomic differentiation performance of conventional, linear-based morphometrics and 3D geometric morphometrics in a species complex that includes three genetically differentiated taxonomic groups containing subtle morphological differences (Viacava et al., 2021), the *A. stuartii* / *A. subtropicus* species complex. This represents a scenario where finer-grained differences (e.g., allometric effects or differences in non-homologous measurements) need to be identified and interpreted with great care to understand their pertinence to group differentiation. The taxonomic situation of this species complex is also useful because three linear morphometric protocols have been used pertaining to the genus, allowing an assessment of how important protocol choice can be to the delimitation of taxonomic units. We add to this also a more generic protocol developed

for bandicoots (Travouillon, 2016), with a particularly high number of linear measurements. We use linear discriminant analysis to ask how well the four protocols and our GMM protocol perform in a context typical to taxonomy (without consideration of size, or with consideration only of isometry) and compare it to the analysis pipeline typically taken in GMM studies (with allometric variation accounted for as well).

Materials and methods

All analyses are based on a 3D landmark coordinate dataset from Viacava et al. (2021), which includes high-coverage 3D landmarked crania with 412 landmarks (82 fixed landmarks, 185 curved semilandmarks and 145 surface semilandmarks) of 136 adult individuals reconstructed in virtual 3D images. All analyses were performed in R version 4.0.4 (R Core Team, 2021). The code and raw 3D data are available on Github (https://github.com/pietroviama/Viacavaetal_LMMvsGMM).

We identified four linear measurement protocols that represent traditional morphometric methods commonly used in Australian mammal taxonomy, but are also specific to *Antechinus*. These include a protocol used for a species contained in the species complex studied here, *A. subtropicus* (Van Dyck & Crowther, 2000), a sister species of the species complex studied here, *A. agilis* (Dickman et al., 1998), a species within the genus *Antechinus*, *A. flavipes* (Baker & Van Dyck, 2013), and a comprehensive protocol that was developed for Peramelemorphians (bandicoots) (Travouillon, 2016). The last protocol is not necessarily expected to apply well to the genus *Antechinus* because it was designed for a different order of marsupials. However, it was chosen as a useful comparison of performance with the other three sets of linear measurements, representing one of the most

comprehensive protocols in the morphometric study of Australian mammals. All of these protocols differ from each other, but overlap in some measurements (Table 1).

Table 1. Degree of overlap of linear measurements between protocols. The LMM protocols in the rows cover a fraction of the LMM protocols in the columns.

	Van Dyck & Crowther, 2000	Dickman et al., 1998	Baker & Van Dyck, 2013	Travouillon, 2016
Van Dyck & Crowther, 2000		61.9 %	66.67 %	27.27 %
Dickman et al., 1998	100 %		72.22 %	40.91 %
Baker & Van Dyck, 2013	92.31 %	61.9 %		31.82 %
Travouillon, 2016	46.15 %	42.86 %	33.33 %	

To obtain linear measurement data, we extracted the linear distances of each protocol that could be estimated most appropriately from the coordinates of the landmarks used for the geometric morphometric approach (Supplementary Table S4.1). These measurements were not exactly the same as calliper measurements; however, we assume that slight inconsistencies between linear-based and 3D landmark-based distances are acceptable because they were taken in a consistent fashion and the representation of shape taken with the linear distances is not lost. We averaged right and left measurements whenever possible.

Isometry and allometry

In geometric morphometrics, the isometric component of shape (i.e. the shape that varies in a 1:1 proportion with size) is generally removed from the dataset through the scaling

192 procedure of the Procrustes superimposition. This step brings all specimens to the same
193 size, producing “isometry-free” shape coordinates and a centroid size (Dryden & Mardia,
194 2016; Klingenberg, 2016) for each specimen. Centroid size can be used subsequently as a
195 proxy for specimen size. To approximate this effect in the LMM context, we used an
196 approach that is analogous to centroid size extraction by deriving the geometric mean of all
197 variables as the centroid size, and using log-shape ratios [$\log_{10}$(measurement/geometric
198 mean)] as isometry-free shape variables. This ensures that each dataset can be analysed
199 in an approximately equivalent way (Claude, 2013; Mosimann, 1970).

200 In order to assess the effect of allometry, we regressed the Procrustes shape
201 variation vs log(centroid size) with the ‘geomorph’ (Adams & Otárola- Castillo, 2013)
202 function ‘procD.lm’. For LMM, we regressed the linear data vs log(geometric mean) with the
203 *lm.rrpp* function of the ‘RRPP’ package (Collyer & Adams, 2018). We considered both
204 centroid size and geometric mean as proxies for size in the context of geometric and linear
205 morphometrics. We also computed “allometry-free” datasets for the classification analyses
206 below, by using the residuals from the allometric regressions. In summary, three types of
207 morphological data were obtained and analysed for the LMM protocols and the GMM
208 dataset: a) raw 3D coordinates obtained from a partial Procrustes superimposition (GMM)
209 and raw linear measurements (LMM), b) shape after Procrustes superimposition (GMM) and
210 log-shape ratios as explained above (LMM), and c) allometry-corrected shape for both. In
211 the case of raw shape, this type of data is typically called “form” in geometric morphometrics
212 (shape plus size). However, for practical purposes, we will further call the types of
213 morphological data explained above as “raw”, “isometry-free” and “allometry-free” shape,
214 respectively. Allometric regressions were performed with 1000 permutations and p-values
215 were calculated using Goodall’s F-test (Goodall, 1991).

Ordination

To assess if the main variation of shape related to differentiation between species, we computed Principal Component Analysis (PCA) for each treatment (raw, isometry-free and allometry-free measurements) and each linear measurement protocol and geometric morphometrics. However, note that lack of differentiation of groups in PC1/PC2 space does not mean that the groups are not differentiated; PCA is agnostic to groupings, such that variation that differentiates a particular group can also be "smeared" across many Principal components (Bookstein, 2015, 2017a, 2017b; Klingenberg et al., 1996; Weisbecker et al., 2019).

Classification rule

To assess how well specimens are predicted to belong to each group based on the different analyses, we used 95% of the PC variance of each dataset to perform a Linear Discriminant Analysis. We used the clade identity as a group factor and provided an equal prior on class membership to the three groups. We plotted the two linear discriminants for each treatment (raw, isometry-free and allometry-free measurements), and for each linear measurement protocol and the geometric morphometrics protocol. Next, we used a machine learning model known as leave-one-out cross validation procedure to calculate the posterior probability values (Venables & Ripley, 2002). These values allowed us to use the 'klaR' package for R (Weihs et al., 2005) to calculate a number of metrics termed Garczarek's classification performance measures (Garczarek & Weihs, 2003), which include Correctness Rate (CR), Accuracy (AC), Ability to Separate (AS), Confidence (CF), and Confidence for each class. The CR and AC values estimate the degree of validity ("quality") of the linear discriminant analysis from the predicted values based on the true values. AS

corresponds to the distance between the posterior values and the assigned groups and CF measures the degree of confidence to which the groups have been assigned – both AS and CF estimate the "certainty" of the result of the linear discriminant analysis (Dr. Karsten Luebke pers. comm.). Finally, we predicted the identity of unidentified specimens (n = 32). For this, we predicted the PC scores of the unidentified specimens and then used the LDA model of our "isometry-free" datasets to predict the class provenance for each specimen (Supp. Table 1).

Results

Allometry

All LMM and GMM protocols were significantly allometric (Table 2). The amount of shape variation attributable to allometry differed substantially, from 7.9% using Van Dyck & Crowther's linear measurement protocol, to over 25% using Travouillon's linear measurement protocol. Dickman et al., Baker & Van Dyck and GMM identified a similar allometric effect of between 11 and 14% of shape variation explained by size (Table 2).

Ordination

The first Principal Component (PC1) of the three LMM protocols developed for antechinuses accounted for more than 70% of morphological variation in raw, isometry-free and allometry-free contexts (Figure 2). Travouillon et al.'s linear measurement protocol was substantially affected in the amount of morphological variation accounted for by PC1 after the removal of isometry, dropping from 73.36% to 38.33%, and to 24.65% after allometric correction. We also observed a reduction in morphological variation accounted by PC1 after removal of isometry in the GMM protocol, from 78.77% to 19.43%, and a slight decrease

after allometric correction to 14.78%. Grouping of the three clades was affected in all three stages of data treatment; in all cases, isometry removal contributed to the assemblage of the groups along PC1 and allometric correction scattered the three groups showing unclear grouping (Figure 2).

Classification rule

The Linear Discriminant Analysis plots display similar groupings of clades for raw and isometry-free measurements. Interestingly, the removal of isometry increased group differentiation in the GMM protocol, whereas the LMM protocols showed a considerable decrease in group differentiation after removal of allometry (Figure 3 and Table 2). For GMM, the removal of allometry did not affect group differentiation as much as for LMM (Figure 3 and Table 2).

For size-unadjusted raw data, the classification performance measures were reasonably high in all four LMM protocols (Table 2). After isometry removal and allometric correction, these measures decreased to varying degrees for all LMM protocols. GMM performed better than LMM at group discrimination at the raw data stage. After the removal of isometry, GMM performed similarly to LMM protocols in CR and AC (“quality” measures) and better in AS and CF (“certainty” measures). After allometric correction, a large decrease in CR and AC was observed in GMM data despite similar performance in AS and CF.

Class predictions for unidentified specimens are shown in Supp. Table 1.

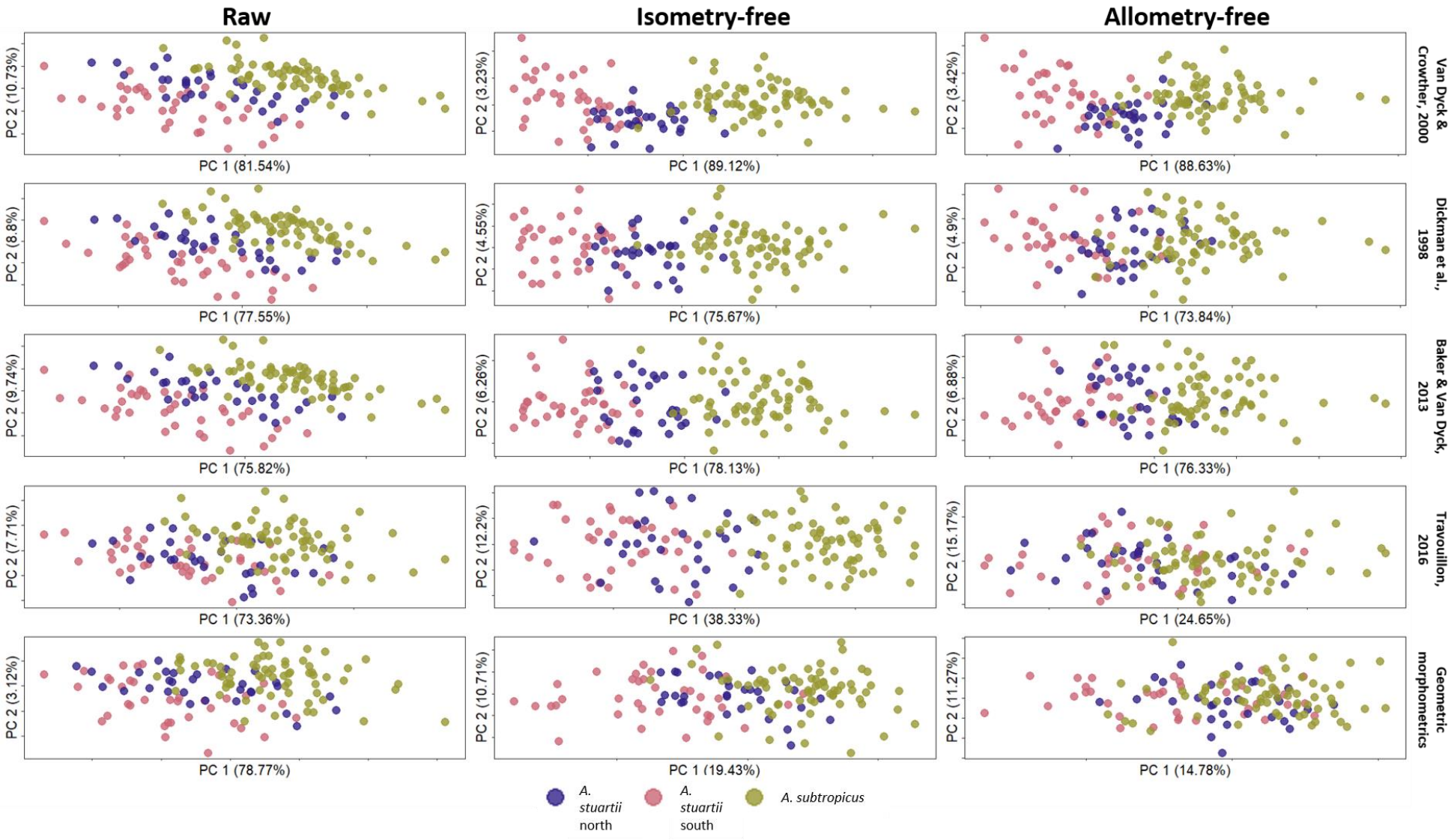
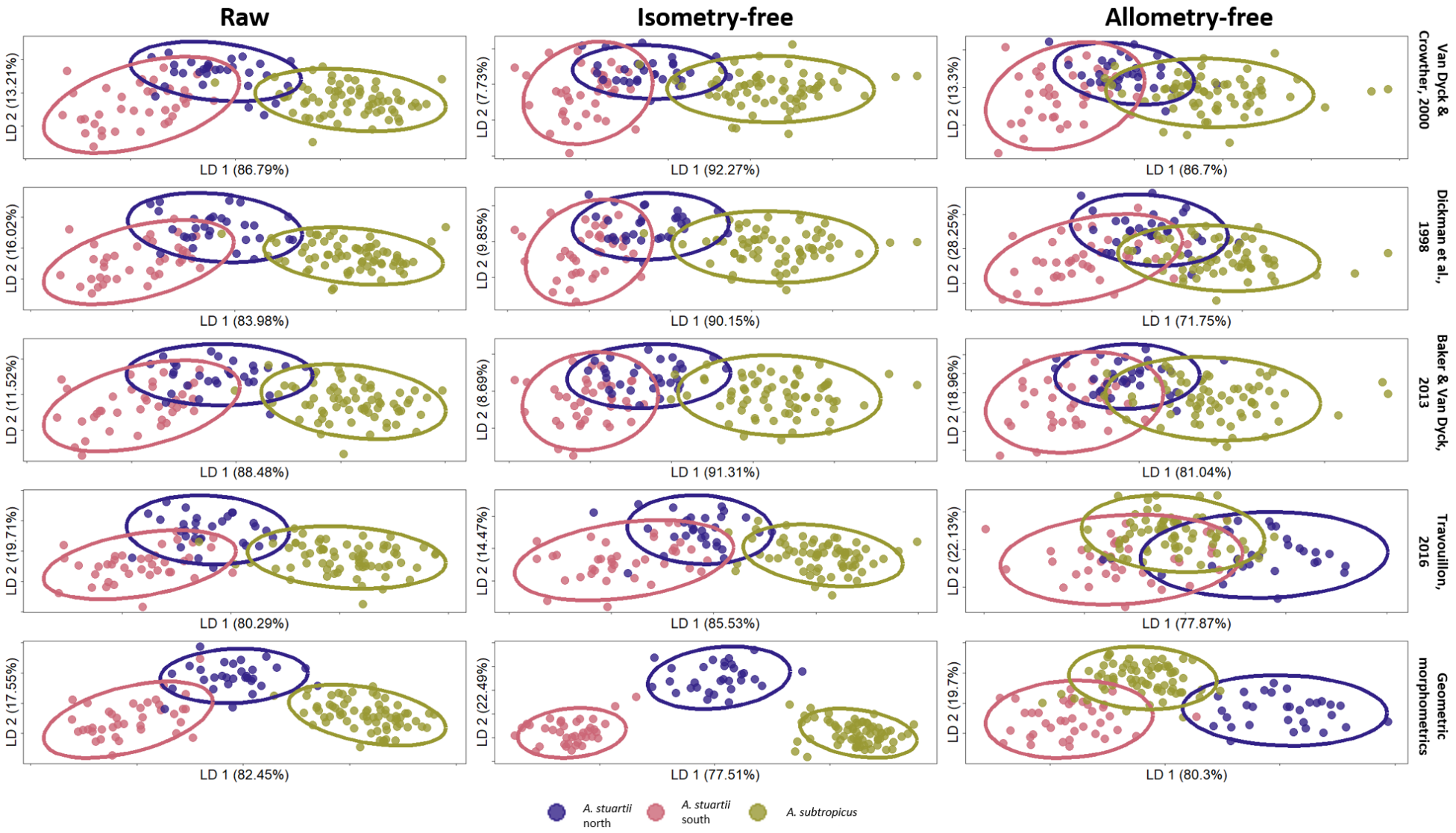


Figure 2: Principal Component Analyses plot for all raw, isometry-free and allometry-free datasets. These include the four linear measurement protocols and the geometric morphometrics approach. Only the first two Principal Components are shown.



284

285 **Figure 3: Linear Discriminant Analyses plot for all raw, isometry-free and allometry-free datasets used in this study. These include the**
286 **four linear measurement protocols and the geometric morphometrics approach. Ellipses were computed at 95%condifence intervals.**

Table 2: Classification performance measures (Garczarek, 2002) of the four linear measurement protocols and geometric morphometrics. For each protocol, the classification performance measures were computed with raw datasets, after size treatment, and after allometry correction. Allometric regression results are also indicated in the last row.

	Van Dyck & Crowther, 2000			Dickman et al., 1998			Baker & Van Dyck, 2013			Travouillon, 2016			GMM		
	Raw	Isometry-free	Allometry-free	Raw	Isometry-free	Allometry-free	Raw	Isometry-free	Allometry-free	Raw	Isometry-free	Allometry-free	Raw	Isometry-free	Allometry-free
Correctness Rate	0.904	0.875	0.838	0.904	0.853	0.765	0.875	0.86	0.757	0.882	0.853	0.662	0.926	0.86	0.581
Accuracy	0.759	0.686	0.55	0.779	0.681	0.516	0.741	0.687	0.5	0.764	0.712	0.328	0.843	0.761	0.286
Ability to Separate	0.855	0.763	0.654	0.87	0.784	0.707	0.837	0.793	0.665	0.879	0.858	0.597	0.943	0.983	0.879
Confidence	0.915	0.861	0.79	0.924	0.874	0.817	0.905	0.879	0.794	0.927	0.914	0.749	0.967	0.99	0.928
Confidence for each true class	North: 0.89	North: 0.773	North: 0.72	North: 0.889	North: 0.802	North: 0.794	North: 0.86	North: 0.776	North: 0.726	North: 0.892	North: 0.863	North: 0.806	North: 0.963	North: 0.961	North: 0.947
	South: 0.875	South: 0.807	South: 0.787	South: 0.886	South: 0.817	South: 0.835	South: 0.87	South: 0.819	South: 0.783	South: 0.926	South: 0.897	South: 0.754	South: 0.953	South: 0.997	South: 0.944
	Sub: 0.949	Sub: 0.931	Sub: 0.823	Sub: 0.961	Sub: 0.937	Sub: 0.818	Sub: 0.945	Sub: 0.958	Sub: 0.83	Sub: 0.944	Sub: 0.946	Sub: 0.721	Sub: 0.976	Sub: 0.998	Sub: 0.91
Allometry	R^2 = 0.079, F = 11.475, p = 0.001			R^2 = 0.144, F = 22.48, p = 0.001			R^2 = 0.113, F = 17.064, p = 0.001			R^2 = 0.251, F = 44.99, p = 0.001			R^2 = 0.132, F = 20.403, p = 0.001		

Discussion

Our results showed that linear morphometrics performed well in distinguishing the three closely related species of antechinus in our dataset. However, the confidence of differentiations was better for the geometric morphometric protocol, particularly after size correction. There is also a clear indication that measurement choice has a substantial influence on the discriminatory performance of a linear measurement protocol.

We found that GMM performed relatively better at discriminating groups based on raw and isometry-free data. However, the traditional morphometric protocols were highly dependent on the choice of the measurements. The fewer variables relative to GMM may therefore improve the discrimination of LMM protocols, but only if the selected linear distances are the “real” best discriminatory ones. In the case of 3D GMM, this dependence on measurement choice is expected to be less pronounced if the creation of the landmarking template relies on the agnostic and comprehensive placement of homologous reference points present in all specimens in a given dataset. The selection of the landmarks should involve the construction of a template that attempts the optimal anatomical coverage with diverse homologous points. This process does not necessarily focus on the most variable regions because it can rely on finding those differentiable shape patterns at the analytical step (Webster & Sheets, 2010). Thus, proportionally, some linear measurement protocols might be best at discriminating but this will only be the case if the linear distances selected are best at discriminating in “reality”. This is probably why the protocol developed for bandicoots (Travouillon, 2016) had the lowest classification performance metrics among all

314 protocols (possibly exacerbated by the fact this was the only LMM protocol not
315 optimised for *Antechinus*).

316 Visual display of the main variation (PC1 vs PC2 plots; “PCA plots” from hereon
317 in) highlight the important issue that an interpretation solely based on the first principal
318 components can be misleading (Schreiber, 2021; Weisbecker et al., 2019) and in our
319 case can lead to a misunderstanding on the performance of GMM data. For GMM
320 data, the PCA plots revealed unclear grouping of the clades (see Figure 2), compared
321 to the much clearer differentiation of clades for the LMM protocols. However, the
322 classification performance measures that used 95% of PC variance of all protocols
323 reflect the ability of GMM to differentiate among clades exceedingly well (see Table
324 2).

325 This superficially better differentiation of the LMM PCA plots relative to the
326 GMM PCA plot is chiefly due to the lower dimensionality of the LMM dataset and the
327 fact that, in these particular cases, the linear distances chosen largely reflected the
328 differences among clades. However, this consideration can be deceptive given the low
329 variance explained by the first principal component in GMM, where in truth the
330 morphological variance tends to be distributed along a larger number of principal
331 components. This simply reflects the fact that the GMM dataset contains far more
332 variation overall, much of which does not differentiate clades. As PCA is agnostic to
333 group membership, this meant that the principal components containing variation that
334 discriminates groups were “hidden” in the lower ranks (Bookstein, 2017b; Klingenberg
335 et al., 1996; Weisbecker et al., 2019). The relevance of ignored morphological
336 variance in a PCA biplot in GMM is emphasized in our LDA results where 95% of the

PC variance was taken into account. This showed a more similar performance to group discrimination in GMM and LMM relative to what we observed in the PCA plots.

The GMM protocol had an interesting property of numerically (and visually; Figure 3) increasing the “certainty” measures of classification after the isometry removal step (between raw and isometry-free datasets). The contribution of geometric morphometrics towards isometry-free group separation may be a substantial improvement in the way we regard size and shape as independent variables for subsequent allometric analyses. In the case of geometric morphometrics, the large number of landmarks may contribute to a holistic characterization of size – in the form of centroid size (Mitteroecker & Gunz, 2009). In the LMM context, the linear distances may contain fewer aspects of the size of the skull (Farkas et al., 2002; Slice, 2006). For example, if we measured only the length or the width of a skull, other linear distances associated with size-related shape could be ignored, such as the width of the snout. This can be a problem because it disregards measures that are characterizing the size of a three-dimensional object (Adams et al., 2004). Furthermore, if size is not characterized well, further consequences on the independence of a size and shape variable can undermine allometric analyses in the form of a size vs shape regression (Klingenberg, 2016).

The removal of shape variation due to allometry (the step from “isometry-free” to “allometry-free”) had a generally larger decrease in all classification performance measures compared to the previous step of removal of isometry (from “raw” to “isometry-free”). In the GMM dataset, this step of removal of allometry had a greater decline in Correctness Rate and Accuracy but lesser decline in Ability to Separate and Confidence (see Table 2) compared to the LMM protocols. This could be either

because of redundancy in the information of nearby landmarks and semilandmarks resulting in “low quality” classification (low CR and AC), or because GMM deals more effectively with allometric variation resulting in “highly certain” classification (highest AS and CF among the datasets). We note that this result may be an indicator of the former where the linear discriminant analysis may wrongly assign classes with false “certainty” due to the poor ratio between variables (PC scores) and observations (number of individuals) typically encountered in GMM. However, we suspect that the latter is the case because the step of removal of allometry has a similar large decrease on Accuracy in GMM and Travouillon’s protocol (2016), despite the particularly stronger allometric relationship captured by Travouillon’s linear measurements. This “large amount of allometry” captured by Travouillon may be caused by the redundancy of some linear measurements that exacerbate some shape patterns driven by size. However, if this was the case in GMM (i.e., high amount of allometry captured due to redundant measurements), the Ability to Separate, reduced drastically in Travouillon, should also be drastically reduced in GMM after removal of allometry. In contrast, what we observe is high performance in Ability to Separate in GMM after the removal of allometry. These contrasting results suggest that geometric morphometric techniques provide a more thorough way of dealing with allometry-driven shape patterns compared to linear measurements.

Our study suggests that GMM and its statistical toolkit provides improved insights into taxon discrimination and particularly the influence of allometric patterns that might not be taxonomically relevant. While GMM-based taxonomic studies are not practical and are very time consuming, they are an excellent first “pilot” step to identify linear measurements that are most likely to discriminate best within a group of interest.

Subsequently, linear morphometric studies can rely on the measurements that account for the relevant shape patterns identified in GMM and be applied to larger sample sizes.

Our results also highlight the dangers of introducing artificial group separation due to “splitting” differently sized, but allometrically uniform taxonomic units. This is not to say that size cannot be a distinguishing feature of two real taxonomic groups; rather, understanding when variation relates to size, as opposed to other factors, can help clarify the distinction between taxa as part of a wider discrimination toolkit (such as genetics, pelage or behavioural traits).

Acknowledgments

We thank all museum collection curators who granted us access to scan the specimens in their care: Heather Janetzki (Queensland Museum), Sandy Ingleby (Australian Museum) and Christopher Wilson and Leo Joseph (Australian National Wildlife Collection from the Commonwealth Scientific and Industrial Research Organisation). We also thank P. V.’s Ph. D. thesis examiners Bastien Mennecart and Laura Wilson for providing generous feedback, and finally Meg Martin for help in the making of Figure 1.

P. V. was supported by a University of Queensland Research Training Tuition Scholarship and a University of Queensland Research Higher Degree Living Stipend Scholarship. This study was also financed by the Australian Research Council Future

406 Fellowship (FT180100634), Discovery Project (DP170103227) and Centre of
407 Excellence (CE170100015) awarded to V. W..

408

409 **Conflict of interest statement:** The authors declare no conflict of interest.

410 **Author's contribution statement:** P. V., V. W. and S. P. B. conceived the ideas and
411 designed methodology; P. V. collected the data; P. V. analysed the data; P. V. and V.
412 W. led the writing of the manuscript. All authors contributed critically to the drafts and
413 gave final approval for publication.

414 **Data accessibility statement:** All data and coded analyses are available in Github
415 (https://github.com/pietroviama/Viacavaetal_LMMvsGMM).

416

417 ## References

418 Abramov, A. V., Puzachenko, A. Y., & Masuda, R. (2018). Cranial Variation in the
419 Siberian Weasel *Mustela sibirica* (Carnivora, Mustelidae) and its Possible
420 Taxonomic Implications. *Zoological Studies*, *57*, e14.
421 https://doi.org/10.6620/ZS.2018.57-14

422 Adams, D. C., & Otárola-Castillo, E. (2013). geomorph: An R package for the collection
423 and analysis of geometric morphometric shape data. *Methods in Ecology and*
424 *Evolution*, *4*(4), 393–399.

425 Adams, D. C., Rohlf, F. J., & Slice, D. E. (2004). Geometric morphometrics: Ten years
426 of progress following the ‘revolution.’ *Italian Journal of Zoology*, *71*(1), 5–16.
427 https://doi.org/10.1080/11250000409356545

428 Alhajeri, B. H. (2021). A morphometric comparison of the cranial shapes of Asian dwarf
429 hamsters (*Phodopus*, Cricetinae, Rodentia). *Zoologischer Anzeiger*, *292*, 184–
430 196. https://doi.org/10.1016/j.jcz.2021.04.001

431 Baker, A. M., & Van Dyck, S. (2013). Taxonomy and redescription of the Yellow-footed
432 Antechinus , *Antechinus flavipes* (Waterhouse) (Marsupialia: Dasyuridae).
433 *Zootaxa*, *3649*(1), 1–62. https://doi.org/10.11646/zootaxa.3649.1.

434 Bookstein, F. L. (2015). The Relation Between Geometric Morphometrics and
435 Functional Morphology, as Explored by Procrustes Interpretation of Individual
436 Shape Measures Pertinent to Function. *The Anatomical Record*, *298*(1), 314–
437 327. https://doi.org/10.1002/ar.23063

438 Bookstein, F. L. (2017a). A method of factor analysis for shape coordinates. *American*
439 *Journal of Physical Anthropology*, *164*(2), 221–245.
440 https://doi.org/10.1002/ajpa.23277

441 Bookstein, F. L. (2017b). A Newly Noticed Formula Enforces Fundamental Limits on
442 Geometric Morphometric Analyses. *Evolutionary Biology*, *44*(4), 522–541.
443 https://doi.org/10.1007/s11692-017-9424-9

444 Cardini, A., & Polly, P. D. (2013). Larger mammals have longer faces because of size-
445 related constraints on skull form. *Nature Communications*, *4*, 2458.

https://doi.org/10.1038/ncomms3458
https://www.nature.com/articles/ncomms3458#supplementary-information

Cardini, A., Polly, P. D., Dawson, R., & Milne, N. (2015). Why the long face? Kangaroos and wallabies follow the same 'rule' of Cranial Evolutionary Allometry (CREA) as Placentals. *Evolutionary Biology*, *42*(2), 169–176. https://doi.org/10.1007/s11692-015-9308-9

Claude, J. (2013). Log-shape ratios, Procrustes superimposition, elliptic Fourier analysis: Three worked examples in R. *Virtual Morphology and Evolutionary Morphometrics in the New Millenium.*, 94.

Collyer, M. L., & Adams, D. C. (2018). RRPP: An r package for fitting linear models to high-dimensional data using residual randomization. *Methods in Ecology and Evolution*, *9*(7), 1772–1779.

Dickman, C. R., Parnaby, H. E., Crowther, M. S., & King, D. H. (1998). *Antechinus agilis* (Marsupialia: Dasyuridae), a new species from the *A. stuartii* complex in south-eastern Australia. *Australian Journal of Zoology*, *46*(1), 1–26. https://doi.org/10.1071/zo97036

Dryden, I. L., & Mardia, K. V. (2016). *Statistical shape analysis: With applications in R* (Vol. 995). John Wiley & Sons.

Farkas, L. G., Tompson, B. D., Katic, M. J., & Forrest, C. R. (2002). Differences Between Direct (Anthropometric) and Indirect (Cephalometric) Measurements of the Skull. *Journal of Craniofacial Surgery*, *13*(1), 105–108.

Fruciano, C. (2016). Measurement error in geometric morphometrics. *Development Genes and Evolution*, *226*(3), 139–158. https://doi.org/10.1007/s00427-016-0537-4

Fruciano, C., Celik, M. A., Butler, K., Dooley, T., Weisbecker, V., & Phillips, M. J. (2017). Sharing is caring? Measurement error and the issues arising from combining 3D morphometric datasets. *Ecology and Evolution*, *7*(17), 7034–7046. https://doi.org/10.1002/ece3.3256

Gálvez-López, E., Kilbourne, B., & Cox, P. G. (2022). Cranial shape variation in mink: Separating two highly similar species. *Journal of Anatomy*, *240*(2), 210–225. https://doi.org/10.1111/joa.13554

Garczarek, U., & Weihs, G. (2003). Standardizing the Comparison of Partitions. *Computational Statistics*, *18*(1), 143–162. https://doi.org/10.1007/s001800300136

Goodall, C. (1991). Procrustes methods in the statistical analysis of shape. *Journal of the Royal Statistical Society: Series B (Methodological)*, *53*(2), 285–321.

Gunz, P., & Mitteroecker, P. (2013). Semilandmarks: A method for quantifying curves and surfaces. *Hystrix, the Italian Journal of Mammalogy*, *24*(1). https://doi.org/10.4404/hystrix-24.1-6292

Jackson, S., & Groves, C. (2015). *Taxonomy of Australian Mammals*. CSIRO Publishing.

Kaliontzopoulou, A., Carretero, M. A., & Llorente, G. A. (2008). Head shape allometry and proximate causes of head sexual dimorphism in *Podarcis* lizards: Joining linear and geometric morphometrics. *Biological Journal of the Linnean Society*, *93*(1), 111–124. https://doi.org/10.1111/j.1095-8312.2007.00921.x

Kendall, D. G. (1989). A survey of the statistical Theory of shape. *Statistical Science*, *4*(2), 87–99. JSTOR.

Klingenberg, C. P. (2011). MorphoJ: an integrated software package for geometric morphometrics. *Molecular Ecology Resources*, *11*(2), 353–357.

Klingenberg, C. P. (2016). Size, shape, and form: Concepts of allometry in geometric morphometrics. *Development Genes and Evolution*, *226*(3), 113–137. https://doi.org/10.1007/s00427-016-0539-2

Klingenberg, C. P. (2022). Methods for studying allometry in geometric morphometrics: A comparison of performance. *Evolutionary Ecology*. https://doi.org/10.1007/s10682-022-10170-z

Klingenberg, C. P., Neuenschwander, B. E., & Flury, B. D. (1996). Ontogeny and Individual Variation: Analysis of Patterned Covariance Matrices with Common Principal Components. *Systematic Biology*, *45*(2), 135–150. https://doi.org/10.1093/sysbio/45.2.135

Marcy, A. E., Guillerme, T., Sherratt, E., Rowe, K. C., Phillips, M. J., & Weisbecker, V. (2020). Australian Rodents Reveal Conserved Cranial Evolutionary Allometry across 10 Million Years of Murid Evolution. *The American Naturalist*, *196*(6), 755–768. https://doi.org/10.1086/711398

509 Mitteroecker, P., & Gunz, P. (2009). Advances in geometric morphometrics.
510 *Evolutionary Biology*, *36*(2), 235–247. https://doi.org/10.1007/s11692-009-
511 9055-x

512 Mosimann, J. E. (1970). Size allometry: Size and shape variables with
513 characterizations of the lognormal and generalized gamma distributions.
514 *Journal of the American Statistical Association*, *65*(330), 930–945.

515 Outomuro, D., & Johansson, F. (2017). A potential pitfall in studies of biological shape:
516 Does size matter? *Journal of Animal Ecology*, *86*(6), 1447–1457.
517 https://doi.org/10.1111/1365-2656.12732

518 Palci, A., & Lee, M. S. Y. (2019). Geometric morphometrics, homology and cladistics:
519 Review and recommendations. *Cladistics*, *35*(2), 230–242.
520 https://doi.org/10.1111/cla.12340

521 Pilbeam, D., & Gould, S. J. (1974). Size and Scaling in Human Evolution. *Science*,
522 *186*(4167), 892–901. https://doi.org/10.1126/science.186.4167.892

523 R Core Team. (2021). *R: A language and environment for statistical computing. R*
524 *Foundation for Statistical Computing, Vienna, Austria*.

525 Rosel, P. E., Taylor, B. L., Hancock-Hanser, B. L., Morin, P. A., Archer, F. I., Lang, A.
526 R., Mesnick, S. L., Pease, V. L., Perrin, W. F., Robertson, K. M., Leslie, M. S.,
527 Berta, A., Cipriano, F., Parsons, K. M., Viricel, A., Vollmer, N. L., & Martien, K.
528 K. (2017). A review of molecular genetic markers and analytical approaches
529 that have been used for delimiting marine mammal subspecies and species.
530 *Marine Mammal Science*, *33*(S1), 56–75. https://doi.org/10.1111/mms.12412

Schreiber, J. B. (2021). Issues and recommendations for exploratory factor analysis and principal component analysis. *Research in Social and Administrative Pharmacy*, *17*(5), 1004–1011. https://doi.org/10.1016/j.sapharm.2020.07.027

Seifert, B. (2008). Removal of allometric variance improves species separation in multi-character discriminant functions when species are strongly allometric and exposes diagnostic characters. *Myrmecological News*, *11*, 91–105.

Sidlauskas, B. L., Mol, J. H., & Vari, R. P. (2011). Dealing with allometry in linear and geometric morphometrics: A taxonomic case study in the *Leporinus cylindriformis* group (Characiformes: Anostomidae) with description of a new species from Suriname. *Zoological Journal of the Linnean Society*, *162*(1), 103–130. https://doi.org/10.1111/j.1096-3642.2010.00677.x

Slice, D. E. (2006). *Modern morphometrics in physical anthropology*. Springer Science & Business Media.

Stone, J. R. (1997). The spirit of D'Arcy Thompson dwells in empirical morphospace. *Mathematical Biosciences*, *142*(1), 13–30. https://doi.org/10.1016/S0025-5564(96)00186-1

Travouillon, K. J. (2016). Investigating dental variation in *Perameles nasuta* Geoffroy, 1804, with morphological evidence to raise *P. nasuta pallescens* Thomas, 1923 to species rank. *Zootaxa*, *4114*(4), 351–392.

Van Dyck, S., & Crowther, M. S. (2000). Reassessment of northern representatives of the *Antechinus stuartii* complex (Marsupialia: Dasyuridae): *A subtropicus* sp.

552 Nov. And *A. adustus* new status. *Memoirs-Queensland Museum*, *45*(2), 611–
553 635.

554 Venables, W. N., & Ripley, B. D. (2002). *Modern Applied Statistics with S*. Springer
555 New York. https://doi.org/10.1007/978-0-387-21706-2

556 Viacava, P., Baker, A. M., Blomberg, S. P., Phillips, M. J., & Weisbecker, V. (2021).
557 Using 3D geometric morphometrics to aid taxonomic and ecological
558 understanding of a recent speciation event within a small Australian marsupial
559 (*Antechinus*: Dasyuridae). *Zoological Journal of the Linnean Society*, *zlab048*.
560 https://doi.org/10.1093/zoolinnean/zlab048

561 Viacava, P., Blomberg, S. P., Sansalone, G., Phillips, M. J., Guillerme, T., Cameron,
562 S. F., Wilson, R. S., & Weisbecker, V. (2020). Skull shape of a widely
563 distributed, endangered marsupial reveals little evidence of local adaptation
564 between fragmented populations. *Ecology and Evolution*, *10*(18), 9707–9720.
565 https://doi.org/10.1002/ece3.6593

566 Webster, M., & Sheets, H. D. (2010). A Practical Introduction to Landmark-Based
567 Geometric Morphometrics. *The Paleontological Society Papers*, *16*, 163–188.
568 https://doi.org/10.1017/S1089332600001868

569 Weihs, C., Ligges, U., Luebke, K., & Raabe, N. (2005). KlaR analyzing German
570 business cycles. In *Data analysis and decision support* (pp. 335–343). Springer.

571 Weisbecker, V., Guillerme, T., Speck, C., Sherratt, E., Abraha, H. M., Sharp, A. C.,
572 Terhune, C. E., Collins, S., Johnston, S., & Panagiotopoulou, O. (2019).
573 Individual variation of the masticatory system dominates 3D skull shape in the

herbivory-adapted marsupial wombats. *Frontiers in Zoology*, *16*(1), 41. https://doi.org/10.1186/s12983-019-0338-5

Wood, B. A., & Stack, C. G. (1980). Does allometry explain the differences between "Gracile" and "Robust" australopithecines? *American Journal of Physical Anthropology*, *52*(1), 55–62. https://doi.org/10.1002/ajpa.1330520108

Yazdi, A. B. (2014). Application of geometric morphometrics to analyse allometry in two species of the genus *Myrmica* (Hymenoptera: Formicidae). *Soil Organisms*, *86*(1), 77–84.

Zelditch, M. L., Swiderski, D. L., & Sheets, H. D. (2012). *Geometric Morphometrics for Biologists: A Primer*. Academic Press.

591

592 **Supp. Table 1:** Class predictions for unidentified specimens using isometry-free data.

	Van Dyck & Crowther, 2000		Dickman et al., 1998		Baker & Van Dyck, 2013		Travouillon, 2016		Geometric morphometrics	
	Class	Posterior probability (%)	Class	Posterior probability (%)	Class	Posterior probability (%)	Class	Posterior probability (%)	Class	Posterior probability (%)
CM12785	*A. subtropicus*	99.87	*A. subtropicus*	99.9	*A. subtropicus*	99.66	*A. subtropicus*	99.81	*A. stuartii south*	99.64
CM12786	*A. stuartii north*	84.95	*A. stuartii north*	81.17	*A. stuartii north*	88.92	*A. stuartii north*	90.51	*A. stuartii north*	100
JM21536	*A. stuartii north*	85.38	*A. stuartii north*	92.07	*A. stuartii north*	93.57	*A. stuartii south*	68.17	*A. stuartii south*	95.16
J3810	*A. stuartii south*	62.85	*A. stuartii north*	52.21	*A. stuartii north*	60.23	*A. stuartii north*	97.48	*A. stuartii north*	70.15
CM3795	*A. stuartii south*	87.99	*A. stuartii south*	94.54	*A. stuartii south*	89.58	*A. stuartii north*	67.2	*A. stuartii north*	100
J5030	*A. stuartii south*	93.87	*A. stuartii south*	93.94	*A. stuartii south*	91.17	*A. stuartii south*	86.6	*A. stuartii south*	100
JM4432	*A. stuartii north*	82.92	*A. stuartii north*	92.35	*A. stuartii north*	92.28	*A. stuartii north*	97.41	*A. stuartii north*	100
JM3944	*A. stuartii north*	92	*A. stuartii north*	98.69	*A. stuartii north*	93.86	*A. stuartii south*	90.03	*A. stuartii north*	100
M22782	*A. stuartii north*	88.39	*A. stuartii north*	96.85	*A. stuartii north*	93.28	*A. stuartii north*	98.8	*A. stuartii north*	100
M22784	*A. stuartii north*	93.36	*A. stuartii north*	90.94	*A. stuartii north*	95.2	*A. stuartii north*	99.72	*A. stuartii north*	99.99
M22785	*A. stuartii north*	58.88	*A. subtropicus*	51.36	*A. stuartii north*	57.38	*A. stuartii north*	88.51	*A. stuartii north*	100
J15888	*A. stuartii north*	51.45	*A. stuartii south*	51.48	*A. stuartii north*	59.4	*A. stuartii north*	91.51	*A. stuartii north*	100
JM14417	*A. stuartii north*	63.98	*A. stuartii north*	81.44	*A. stuartii north*	63.89	*A. subtropicus*	61.98	*A. stuartii north*	100
RT1	*A. stuartii north*	97.79	*A. stuartii north*	99.42	*A. stuartii north*	98.32	*A. stuartii north*	98.36	*A. stuartii north*	100
JM1600	*A. stuartii north*	80.24	*A. stuartii north*	81.09	*A. stuartii north*	83.52	*A. subtropicus*	69.26	*A. stuartii south*	96.95
JM1596	*A. stuartii north*	81.93	*A. stuartii north*	87.6	*A. stuartii north*	93.59	*A. stuartii north*	74.2	*A. stuartii north*	99.28

J17400	*A. subtropicus*	100	*A. subtropicus*	100	*A. subtropicus*	100	*A. subtropicus*	100	*A. subtropicus*	100
J17401	*A. subtropicus*	82.82	*A. subtropicus*	83.59	*A. subtropicus*	97.78	*A. subtropicus*	93.8	*A. subtropicus*	100
J17402	*A. subtropicus*	100	*A. subtropicus*	100	*A. subtropicus*	100	*A. subtropicus*	99.97	*A. subtropicus*	99.49
J17403	*A. subtropicus*	99.97	*A. subtropicus*	99.99	*A. subtropicus*	99.99	*A. subtropicus*	99.96	*A. subtropicus*	100
J17406	*A. subtropicus*	99.38	*A. subtropicus*	99.72	*A. subtropicus*	99.9	*A. subtropicus*	86.44	*A. subtropicus*	100
JM1420	*A. stuartii north*	83.45	*A. stuartii north*	93.65	*A. stuartii north*	78.72	*A. stuartii north*	94.95	*A. stuartii north*	100
JM14415	*A. stuartii north*	79.26	*A. stuartii north*	86.65	*A. stuartii north*	88.53	*A. stuartii north*	64.21	*A. subtropicus*	99.84
J20265	*A. stuartii north*	95.54	*A. stuartii north*	96.88	*A. stuartii north*	97.77	*A. stuartii north*	72.12	*A. stuartii north*	93.19
MWA1	*A. stuartii north*	77.94	*A. stuartii north*	86.11	*A. stuartii north*	73.15	*A. stuartii north*	68.85	*A. stuartii north*	99.69
MWA2	*A. stuartii north*	81.1	*A. stuartii north*	83.42	*A. stuartii north*	78.38	*A. stuartii north*	98.39	*A. stuartii north*	100
CG1	*A. stuartii north*	88.79	*A. stuartii north*	76.56	*A. stuartii north*	88.12	*A. stuartii north*	75.14	*A. stuartii north*	100
JM20761	*A. stuartii north*	69.88	*A. stuartii north*	72.83	*A. stuartii north*	85.39	*A. stuartii north*	66.84	*A. stuartii north*	100
JM21357	*A. stuartii north*	83.37	*A. stuartii north*	85.79	*A. stuartii north*	88.05	*A. stuartii north*	96.27	*A. stuartii north*	100
JM14452	*A. subtropicus*	62.93	*A. stuartii north*	58.51	*A. stuartii north*	70.93	*A. subtropicus*	95.9	*A. stuartii north*	99.73
CM674	*A. subtropicus*	93.56	*A. subtropicus*	97.38	*A. subtropicus*	99.64	*A. subtropicus*	87.61	*A. subtropicus*	99.84
CM675	*A. stuartii north*	96.48	*A. stuartii north*	97.83	*A. stuartii north*	97.61	*A. stuartii north*	93.34	*A. subtropicus*	100

593