View metadata, citation and similar papers at core.ac.uk

brought to you by CORE

provided by LJMU Research Online

LJMU Research Online

Giacomini, G, Scaravelli, D, Herrel, A, Veneziano, A, Russo, D, Brown, RP and Meloro, C

3D Photogrammetry of Bat Skulls: Perspectives for Macro-evolutionary Analyses

http://researchonline.ljmu.ac.uk/id/eprint/10759/

Article

Citation (please note it is advisable to refer to the publisher's version if you intend to cite from this work)

Giacomini, G, Scaravelli, D, Herrel, A, Veneziano, A, Russo, D, Brown, RP and Meloro, C (2019) 3D Photogrammetry of Bat Skulls: Perspectives for Macro-evolutionary Analyses. Evolutionary Biology. ISSN 0071-3260

LJMU has developed **LJMU Research Online** for users to access the research output of the University more effectively. Copyright © and Moral Rights for the papers on this site are retained by the individual authors and/or other copyright owners. Users may download and/or print one copy of any article(s) in LJMU Research Online to facilitate their private study or for non-commercial research. You may not engage in further distribution of the material or use it for any profit-making activities or any commercial gain.

The version presented here may differ from the published version or from the version of the record. Please see the repository URL above for details on accessing the published version and note that access may require a subscription.

For more information please contact researchonline@ljmu.ac.uk

http://researchonline.ljmu.ac.uk/

Evolutionary Biology
https://doi.org/10.1007/s11692-019-09478-6

TOOLS AND TECHNIQUES

3D Photogrammetry of Bat Skulls: Perspectives for Macro-evolutionary Analyses

Giada Giacomini[1] · Dino Scaravelli[2] · Anthony Herrel[3] · Alessio Veneziano[1] · Danilo Russo[4,5] · Richard P. Brown[1] · Carlo Meloro[1]

Received: 7 February 2019 / Accepted: 4 May 2019
© The Author(s) 2019

Abstract

Photogrammetry (PH) is relatively cheap, easy to use, flexible and portable but its power and limitations have not been fully explored for studies of small animals. Here we assessed the accuracy of PH for the reconstruction of 3D digital models of bat skulls by evaluating its potential for evolutionary morphology studies at interspecific (19 species) level. Its reliability was assessed against the performance of micro CT scan (μCT) and laser scan techniques (LS). We used 3D geometric morphometrics and comparative methods to quantify the amount of size and shape variation due to the scanning technique and assess the strength of the biological signal in relation to both the technique error and phylogenetic uncertainty. We found only minor variation among techniques. Levels of random error (repeatability and procrustes variance) were similar in all techniques and no systematic error was observed (as evidenced from principal component analysis). Similar levels of phylogenetic signal, allometries and correlations with ecological variables (frequency of maximum energy and bite force) were detected among techniques. Phylogenetic uncertainty interacted with technique error but without affecting the biological conclusions driven by the evolutionary analyses. Our study confirms the accuracy of PH for the reconstruction of challenging specimens. These results encourage the use of PH as a reliable and highly accessible tool for the study of macro evolutionary processes of small mammals.

Keywords 3D reconstruction · Geometric morphometrics · Measurement error · Technique comparison

Introduction

The use of digital 3D models in morphological studies is increasing in many scientific disciplines, including palaeontology and evolutionary biology. The digitalization of

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s11692-019-09478-6) contains supplementary material, which is available to authorized users.

✉ Giada Giacomini
g.giacomini@2016.ljmu.ac.uk

Dino Scaravelli
dino.scaravelli@gmail.com

Anthony Herrel
anthony.herrel@mnhn.fr

Alessio Veneziano
veneziano.alessio@gmail.com

Danilo Russo
danrusso@unina.it

Richard P. Brown
r.p.brown@ljmu.ac.uk

Carlo Meloro
c.meloro@ljmu.ac.uk

[1] Research Centre in Evolutionary Anthropology and Paleoecology, School of Natural Sciences and Psychology, Liverpool John Moores University, Byrom Street, Liverpool L3 3AF, UK

[2] Department of Veterinary Medical Sciences, University of Bologna, 50 Via Tolara di Sopra, Ozzano dell'Emilia, 40064 Bologna, BO, Italy

[3] Department Adaptation du Vivant, UMR 7179 C.N.R.S/M.N.H.N, 55 Rue Buffon, 75005 Paris, France

[4] Wildlife Research Unit, Agrarian Department, University of Naples Federico II, 100 Via Università, 80055 Portici, NA, Italy

[5] School of Biological Sciences, University of Bristol, Life Sciences Building, 24 Tyndall Avenue, Bristol BS8 1TQ, UK

Published online: 22 May 2019

an object not only facilitates detailed analysis of the size and shape of fragile specimens but also helps investigation of diverse evolutionary questions (e.g. Cardini et al. 2015; Cornette et al. 2013).

The use of close-range photogrammetry (PH) has grown in many fields because it is economical, portable, easy to apply and accurately reproduces the geometry and colour pattern of real and complex objects (Falkingham 2012). For this reason, it has become widely employed in a variety of disciplines such as biology (Evin et al. 2016), palaeontology (Bates et al. 2010), anthropology (Katz and Friess 2014) and medicine (Ege et al. 2004), among others.

In analyses of shape and size of objects (as in biological studies), the 3D models are often integrated with geometric morphometric methods (GMM). This approach has proved particularly useful in bats, where, for example, GMM has provided additional information on divergence of cryptic species (Sztencel-Jabłonka et al. 2009). Nevertheless, acquiring landmarks on bone sutures of bat skulls, particularly for Microchiroptera sensu Simmons and Geisler (1998), is quite difficult due to early suture ossification and their small size. This challenge often forces researchers to employ extremely precise equipment at considerable cost. However, no studies have addressed the utility of PH for this group and other similar sized mammals.

Katz and Friess (2014) and Evin et al. (2016) demonstrated the accuracy of close-range PH for large skulls (humans and wolves, respectively) relative to laser scan models. Fahlke and Autenrieth (2016) compared PH performance relative to µCT scan models for a vertebrate fossil skull (condyle-basal length = 37.5 cm) and similarly found high similarity. Very few studies have attempted to apply it to smaller specimens although Muñoz-Muñoz et al. (2016) and assessed the repeatability of PH for mice skulls (length = 45 mm) and suggested it might be appropriate for small mammals. Durão et al. (2018) suggested a protocol for 3D reconstruction of vole humerii by mean of PH. Nevertheless, no tests were conducted to assess its performance against more established 3D reconstruction techniques (e.g. µCT scan). High measurement error (random error, in particular) is well-known in small specimens and largely arises due to difficulties in landmark identification (Badawi-Fayad and Cabanis 2007; Cramon-Taubadel et al. 2007; Fourie et al. 2011; Marcy et al. 2018; Muñoz-Muñoz et al. 2016). The extent of biological variation is of paramount importance when considering the impact of technique-based error on the results (Marcy et al. 2018).

An additional incentive for analysing differences between techniques is that it may lead to an understanding of when it is feasible to combine data acquired using different techniques. The introduction of random and systematic errors intrinsic to each technique is known to create unreal patterns and/or obscure biological variation (Fruciano et al. 2017; Marcy et al. 2018; Robinson and Terhune 2017).

This study was motivated by the need to assess PH as a tool for reliable analysis of bat skull morphology and assess its performance relative to µCT scan and surface laser scan (LS). We used GMM to assess the relative accuracy of PH models for quantifying size and shape via anatomical landmarks. Phylogenetic comparative methods (Cornwell and Nakagawa 2017) were used to assess the strength of biological signal against the technique error and the phylogenetic uncertainty. Our aims were to quantify the extent of measurement error introduced by the PH/GMM approach and assess the reliability of combining data extracted from different reconstruction techniques (PH, µCT, LS).

Materials and Methods

Sample

GMM and phylogenetic comparative methods were used to examine the reliability of PH for the digital reconstruction of bat skulls and assess its performance in interspecific (19 species) evolutionary analyses.

Crania from nineteen different bat species from the Natural History Museum of Paris (MNHN) were reconstructed in 3D using three different techniques: PH, LS and µCT scanning. The specimens were selected to represent bat species of small and medium size, with an average skull length of 15.62 mm (Table S1).

Data Acquisition and Model Landmarking

The 3D models have been reconstructed with three different techniques (PH, LS, CT).

The PH 3D models were obtained employing a digital SLR Nikon D5300 camera attached to a Nikkor 60 mm macro lens. The general camera lighting settings and positioning, specimen arrangement and number of pictures per specimen were adapted from Falkingham (2012) and Mallison and Wings (2014). Average mesh size was ~ 3,000,000 faces.

For the LS models, we employed a Breuckmann Laser Scan, model SmartSCAN R5/C5 5.0 MegaPixel housed at the MNHN of Paris. We used the field of view S-030 which is optimal for very small objects (240 mm length) and can achieve a maximum resolution of 10 µm.

To obtain the CT models we used a phoenix v|tome|x s housed at the MNHN of Paris. Scans resolution ranged between 18 and 28 µm (average 23 µm) in voxel.

Detailed information on devices and workflow are available in the supplementary material.

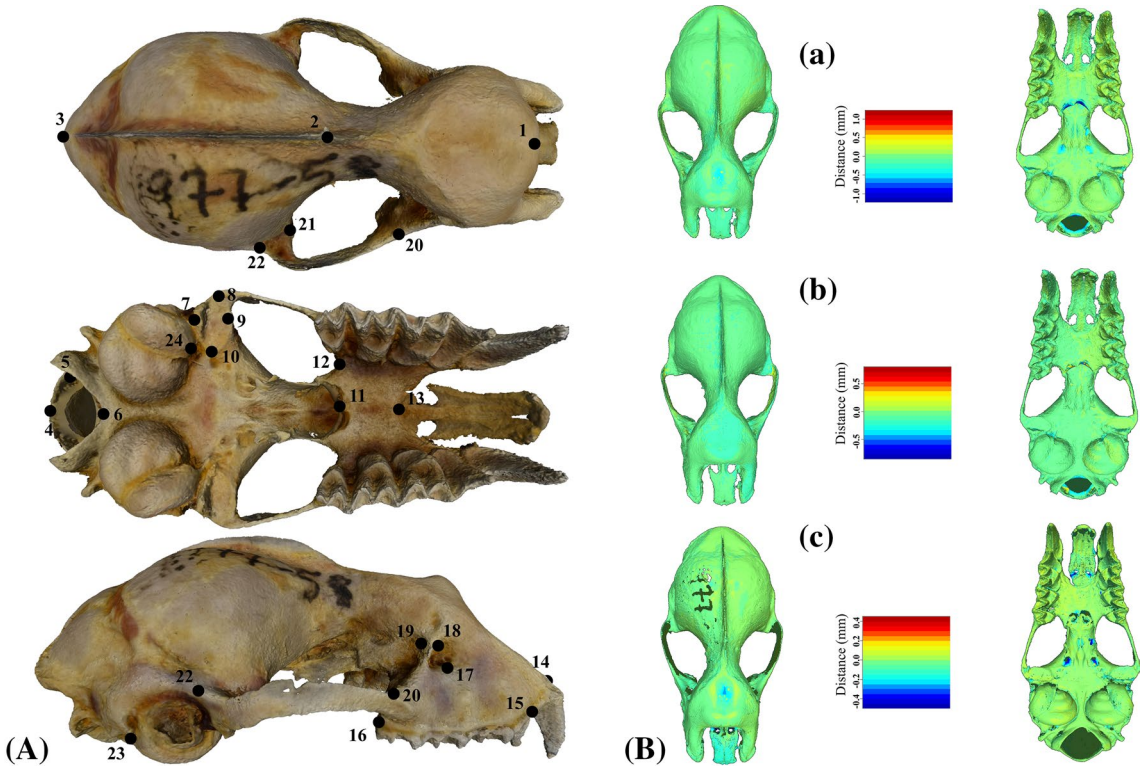

Fig. 1 **a** Landmark configuration used in the study. Species: *R. ferrumequinum*. Anatomical definition in Table S1. **b** Visualization of mesh distances on dorsal and ventral views between (a) PH and µCT; (b) PH and LS; (c) µCT and LS. The colour represents the distances in mm. Species: *R. ferrumequinum* (skull length: 18.78 mm)

The open source software Landmark Editor (Wiley et al. 2005) was used to place 24 unilateral landmarks on the dorsal, lateral and ventral side of the cranium (Fig. 1a and Table S2). See Supplementary material for details on landmarks acquisition and coordinates transformation procedure (Procrustes Shape Coordinates).

For the interspecific study (19 species), we assessed the landmarking error by recording coordinates three times on a subsample of nine species, selected to represent the morphological variation within our sample (*Carolia perspicillata, Desmodus rotundus, Glossophaga soricina, Myotis emarginatus, Myotis capaccinii, Nyctalus noctula, Rhinolophus hipposideros, Rhinolophus ferrumequinum* and *Tadarida teniiotis*) for each technique (LS, PH, µCT). Some species were morphologically very divergent, as assessed from principal component scores (see later) (e.g. *D. rotundus* and *G. soricina*) while others were very similar (e.g. *R. hipposideros* and *R. ferrumequinum*).

Measurement error evaluation

Mesh Distances

The average distances between the 19 paired models were calculated in R software (R Core Team 2018) using the meshDist function in the "Morpho" package (Schlager 2013). This distance is defined as an average of the shortest distances between every triangle of a mesh and the closest triangle of the other (Bærentzen and Henrik 2002). It returns the average distance and a coloured scale model that visually represents the differences between each pair of meshes.

Shape Visualization

The preliminary visual analysis of the shape differences between the specimens was achieved using a Principal Component Analysis (PCA) for the interspecific dataset. We used the variance–covariance matrix of the Procrustes coordinates to extract orthogonal vectors (PCs) that summarise variation within our sample. Shape changes in 3D skulls were visualised by warping the 3D coordinates along the PC axes. This was achieved applying a Thin-Plate-Spline (Bookstein 1989) algorithm on the mean shape of the morphospace. The 3D bat skull with lowest deviation from the mean shape was chosen for the visualisation. This model was warped along the positive and the negative sides of PC axes to display the shape variation in the sample (Drake and Klingenberg 2010).

Error in Geometric Morphometrics

Pearson and Mantel tests were employed to assess the similarity between the centroid size vectors produced by each technique, and their shape coordinates matrices, respectively (Cardini 2014). Procrustes and standard ANOVAs were used to quantify the variance explained by the different techniques for shape and size, respectively. Nested ANOVAs were used to analyse replicate measurements to assess the landmarking error in a subsample of the data (nine species, see above), with repeatability computed using the intraclass correlation coefficient, i.e., among individual-variance divided by within-individual variance components (see Fruciano 2016). The variability of Procrustes variance computed for each triplet of replicate was used as a further indicator of random measurement error within each technique (Marcy et al. 2018). The Procrustes variance, also known as morphological disparity, measures the magnitude of morphological variation for each triplet by technique (Zelditch et al. 2012). Kruskal–Wallis tests were used to compare median Procrustes variances between techniques (greater variation suggests lower precision in landmark identification). Pearson correlation tests between Procrustes variance and centroid size assessed whether errors in landmark identification were greater for smaller specimens.

Error in Evolutionary Analysis

Additional analyses were performed on the interspecific dataset to assess the use of PH-generated data in evolutionary studies.

Phylogenetic trees for the 19 selected species were inferred by Bayesian inference, as implemented in MrBayes version 3.2 (Huelsenbeck and Ronquist 2001). Input data consisted of an alignment of 20,364 base pairs of nuclear and mitochondrial DNA from Shi and Rabosky (2015). The alignment was divided into 29 partitions (for details see Shi and Rabosky 2015) to allow for evolutionary differences between partitions. The GTR + G model was applied to each partition. The MCMC chain was run for 5 million generations, with trees saved every 500 generations and the first 5×10^3 trees discarded as burn-in. The remaining posterior sample of 1000 trees and the 50% majority rule consensus tree was used for subsequent analyses.

The R packages "ape" (Paradis et al. 2004) and "geomorph" (Adams and Otárola-Castillo 2013) were used to test for the presence of evolutionary allometry (Cardini and Polly 2013) in the three datasets using the log10 transformed centroid size as the independent variable and Procrustes shape coordinates as the dependent variable. Phylogenetic generalised least squares (PGLS) analyses with 999 permutations was employed on the three datasets separately to test for the presence of evolutionary allometry after taking the phylogenetic variance–covariance matrix into account, with the phylogeny represented by the Bayesian consensus tree (Adams and Collyer 2015; Rohlf 2007).

The presence of phylogenetic signal (quantified by the K statistic, Blomberg et al. 2003) in the three datasets and the degree of congruence for size and shape (Adams 2014) were also analysed using the consensus tree. The *K* statistic reflects the degree of congruence between phenotypic data and the phylogeny (Blomberg et al. 2003). Statistical significance of *K* and its multivariate extension K_{multiv} was assessed using randomization (Adams 2014).

To examine whether the same evolutionary conclusions were obtained using different techniques, we computed a series of ANOVAs with morphological data (log10 transformed centroid size and shape coordinates) as the dependent variable and ecological data (log10 transformed frequency peak and log10 transformed bite force) as the independent variable for all species in the study except *Pipistrellus nathusii* (no data on bite force were available for the species). Frequency peak data were extracted from the literature (Brinkløv et al. 2011; Kalko et al. 1998; Rodríguez-San Pedro and Allendes 2017; Russo and Jones 2002; Siemers et al. 2001; Siemers and Schnitzler 2004). We obtained unpublished (collected by AH) and published bite force data (Aguirre et al. 2002) for these analyses. The same analyses were repeated under a phylogenetic comparative approach using PGLS.

To assess whether the same results were obtained from mixed datasets acquired from the three different 3D reconstruction techniques, we constructed 1000 morphological datasets in which data for each of the nineteen species were randomly selected from one of the three techniques (PH, µCT, LS). Allometry, phylogenetic signal and correlation with bite force and frequency peak (assessed as previously described) were analysed for each dataset using standard and phylogenetic comparative approaches. The mean, standard deviation, minimum and maximum of the parameter distributions were used as statistical descriptors of the variables distributions and were compared to the original results obtained with singular-technique datasets (PH, µCT, LS).

Fruciano et al.'s (2017) approach was used to assess the error due to phylogenetic uncertainty in the evolutionary analyses. The 1000 posterior trees represented the phylogenetic uncertainty in these analyses. Three common evolutionary analyses were performed: assessment of phylogenetic signal, relation between peak frequency and morphological data, relation between bite force and morphological data. For each technique-tree combination we performed the three analyses for both size and shape obtaining a distribution of 1000 estimates for each analysis. ANOVAs were performed on each distribution to assess the variance explained by the both phylogenetic uncertainty and reconstruction technique.

Table 1 Average distances (mm) between the surface of the models

Specimen	PH-LS	µCT-PH	LS-µCT
Carollia perspicillata	0.070	0.090	0.001
Desmodus rotundus	0.007	0.013	0.012
Eptesicus serotinus	0.028	0.035	0.020
Glossophaga soricina	0.051	0.071	0.023
Hypugo savii	0.032	0.034	0.004
Myotis daubentonii	0.058	0.092	0.016
Miniopterus schreibersi	0.040	0.039	0.002
Myotis capaccinii	0.173	0.188	0.012
Myotis emarginatus	0.069	0.065	0.000
Myotis nigricans	0.040	0.083	0.029
Myotis dasycneme	0.026	0.046	0.060
Noctilio albiventris	0.001	0.002	0.003
Nyctalus noctula	0.004	0.058	0.043
Pipistrells pipistrellus	0.027	0.037	0.016
Pipistrellus nathusii	0.036	0.042	0.012
Plecotus austriacus	0.075	0.076	0.002
Rhinolophus ferrumequinum	0.001	0.007	0.004
Rhinolophus hipposideros	0.030	0.021	0.011
Tadarida teniotis	0.016	0.033	0.015
Mean	0.041	0.054	0.015
SD	0.039	0.042	0.015

PH photogrammetry, *LS* laser scan, *µCT* micro CT scan

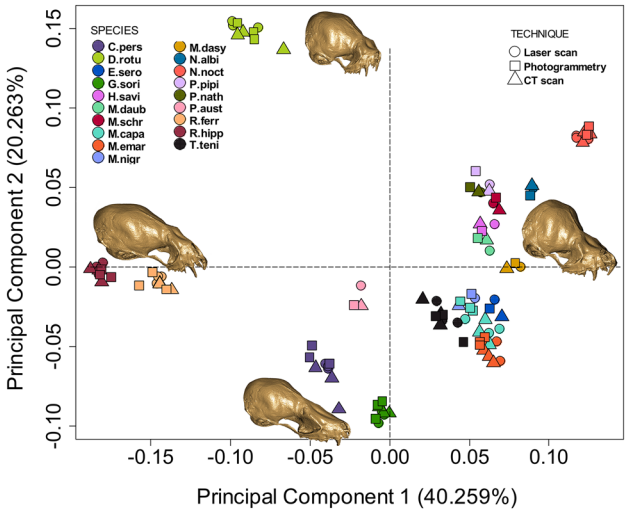

Fig. 2 PCA of 57 models (19 specimens × 3 techniques) and 54 replicas (9 specimens × 2 replicas × 3 techniques). Each skull was reconstructed with three different techniques (filled circle LS, filled square PH and filled triangle µCT). For nine specimens we recorded the landmarks three times (1,2,3 inside the symbols): C.pers, Drotu, G.sori, M.emar, M.capa, N.noct, R.hipp, R.ferr and T.teni. The four skull images on the two axes represent the extreme shapes of the morphospace for PC1 and PC2 (species used as reference model for the warping: *Plecotus austriacus*). *Carolia perspicillata* (C.pers), *Desmodus rotundus* (D.rotu), *Eptesicus serotinus* (E.sero), *Glossophaga soricina* (G.sori), *Hypsugo savii* (H.savi), *Miniopterus schreibersi* (M.schr), *Myotis daubentonii* (M.daub), *Myotis dasycname* (M.dasy), *Myotis emarginatus* (M.emar), *Myotis capaccinii* (M.capa), *Myotis nigricans* (M.nigr), *Noctilio albiventris* (N.albi), *Nyctalus noctula* (N.noct), *Pipistrellus nathusii* (P.nath), *Pipistrellus pipistrellus* (P.pipi), *Plecotus austriacus* (P.aust), *Rhinolophus hipposideros* (R.hipp), *Rhinolophus ferrumequinum* (R.ferr) and *Tadarida teniotis* (T.teni)

Results

The nineteen models were reconstructed in 3D with the three different technique and the photogrammetric 3D model of *Rhinolophus ferrumequinum* (MNHN-ZM-MO-1977-58) can be downloaded as an example from Morphosource (model ID = M30222; https://www.morphosource.org).

Mesh Distances

Visual examination of the meshes revealed strong general similarity between the three data sets except in certain specific areas (Fig. S1). There were small distances between the surfaces of the models as shown for *Rhinolophus ferrumequinum* (Fig. 1b). The average distance between PH and LS models was 0.041 mm, in agreement with that found by Evin et al. (2016) for 5 wolf skulls (0.088 mm) (Table 1). The average distance between the PH and µCT models was 0.054 mm. Finally, the µCT and LS models were extremely similar with an average distance of 0.015 mm (Table 1).

Shape Visualization

The morphospace of the 111 specimens (i.e., 57 models plus 54 replicates) displays the shape variability in the sample (Fig. 2). The first principal component (PC1) explains 40.26% of the total variance while PC2 explains 20.26%. PC1 shows shape variation mainly related to the length of the supra-occipital bone, while PC2 represents variation in palate length (warped skulls in Fig. 2). Samples clearly cluster according to the species/individuals and not to the technique employed. Replicates were also tightly clustered, except for *M. capaccinii* (which had some cartilage tissue still attached to the bone making landmark identification difficult). Replicate clusters indicated no evidence of explicit random or systematic (i.e., bias) errors: none of the technique showed greater variability relative to the others nor was there evidence of differences in mean positioning due to replicate/technique.

Error in Geometric Morphometrics

Correlations between centroid size vectors obtained from the different models provided coefficients greater than 0.99 for all combinations (PH-LS: $R = 0.997$, $p < 0.001$; µCT-PH: $R = 0.996$, $p < 0.001$; LS-µCT: $R = 0.998$,

Table 2 (A) ANOVA on size and (B) Procrustes ANOVA on shape for 57 models (19 specimens × 3 techniques) and 54 replicas (9 specimens × 2 replicas × 3 techniques)

(A) Size	Df	SS	MS	Rsq	F	Z	Pr(> F)
Species	18	4016.129	223.118	0.997	2632.394	17.688	0.001
Technique	2	5.744	2.872	0.001	33.887	7.539	0.001
Residuals	90	7.628	0.085	0.002			
Total	110	4029.502					
(B) Shape	Df	SS	MS	Rsq	F	Z	Pr(> F)
Species	18	2.117	0.118	0.945	90.535	20.517	0.001
Technique	2	0.006	0.003	0.003	2.274	13.769	0.001
Residuals	90	0.117	0.001	0.052			
Total	110	2.240					

Table 3 Landmarking error and repeatability for replicas only. (A) ANOVA on size and (B) Procrustes ANOVA on shape for 81 models (9 specimens × 3 replicas × 3 techniques)

(A) Size	Df	SS	MS	Rsq	F	Z	Pr(> F)	Repeatability
Species	8	2645.123	330.640	0.939	12,842.461	10.497	0.001	0.99
Species:Replicas	18	0.463	0.026	0.000	0.008	− 6.129	0.999	
Residuals	54	170.862	3.164	0.061				
Total	80	2816.449						
(B) Shape	Df	SS	MS	Rsq	F	Z	Pr(> F)	Repeatability
Species	8	1.688	0.211	0.941	104.374	7.384	0.001	0.97
Species:Replicas	18	0.036	0.002	0.020	1.553	23.080	0.001	
Residuals	54	0.070	0.001	0.039				
Total	80	1.795						

$p < 0.001$). Similarly, high associations were obtained from Mantel matrix correlations on the Procrustes distances between individual specimens across the techniques (PH-LS: $R = 0.988$, $p < 0.001$; µCT-PH: $R = 0.988$, $p < 0.001$; LS-µCT: $R = 0.992$, $p < 0.001$). Furthermore, the ANOVA test on size showed that 99.67% ($p = 0.001$) of the variance can be explained by biological differences between specimens, with only 0.14% attributable to the technique ($p = 0.001$) (Table 2). In terms of shape, 94.52% ($p < 0.001$) of the shape variance was explained by the specimen differences while only the 0.26% was represented by the different techniques ($p = 0.001$).

The landmarking error represented a small portion of the variance in both size (between-replicate variance: 0.02%, $p = 0.999$) and shape (between-replicate variance: 2.03%, $p = 0.001$). The repeatability was 0.99 for size and 0.97 for shape (Table 3a-b).

The mean Procrustes variance was not statistically different between techniques ($p = 0.979$) suggesting difficulty in landmark identification is similar between the techniques (Fig. 3a). Correlations between Procrustes variances (for each technique) and centroid size showed no significant associations (PH: $R = 0.16$, $p = 0.683$; CT: $R = 0.48$, $p = 0.187$; LS: $R = 0.052$, $p = 0.894$).

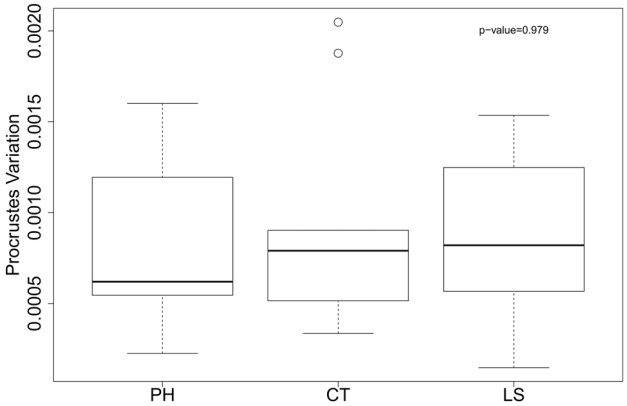

Fig. 3 Procrustes variation computed for replicates/technique for the interspecific dataset

Table 4 Results of K_{multiv} phylogenetic signal and R^2 for allometry and correlation with ecological variables for the interspecific dataset

	PS size		PS shape			Allometry			
						no-Phy		with-Phy	
	Kmultiv	pvalue	Kmultiv	pvalue		Rsq	pvalue	Rsq	pvalue
PH	0.818	0.027	0.919	0.001	PH	0.062	0.297	0.098	0.160
CT	0.857	0.019	0.938	0.001	µCT	0.068	0.226	0.105	0.124
LS	0.868	0.018	0.972	0.001	LS	0.072	0.196	0.099	0.145
Mean	0.848	0.021	0.943	0.001	Mean	0.067	0.240	0.101	0.143
SD	0.021	0.004	0.022	0.000	SD	0.004	0.042	0.003	0.015
	Size ~ BF					Shape ~ BF			
	no-Phy		with-Phy			no-Phy		with-Phy	
	Rsq	pvalue	Rsq	pvalue		Rsq	pvalue	Rsq	pvalue
PH	0.780	0.001	0.846	0.001	PH	0.080	0.196	0.037	0.724
µCT	0.771	0.001	0.826	0.001	µCT	0.082	0.18	0.039	0.801
LS	0.774	0.001	0.835	0.001	LS	0.097	0.103	0.051	0.577
Mean	0.775	0.001	0.836	0.001	Mean	0.087	0.160	0.042	0.701
SD	0.004	0.000	0.008	0.000	SD	0.008	0.041	0.006	0.093
	Size ~ FP					Shape ~ FP			
	no-Phy		with-Phy			no-Phy		with-Phy	
	Rsq	pvalue	Rsq	pvalue		Rsq	pvalue	Rsq	pvalue
PH	0.012	0.680	0.316	0.004	PH	0.152	0.013	0.093	0.051
µCT	0.013	0.672	0.331	0.001	µCT	0.156	0.014	0.093	0.052
LS	0.012	0.681	0.329	0.002	LS	0.158	0.013	0.092	0.056
Mean	0.012	0.678	0.325	0.002	Mean	0.155	0.013	0.092	0.053
SD	0.000	0.004	0.007	0.001	SD	0.002	0.000	0.001	0.002

Results are computed by technique with (with-Phy) and without (no-Phy) phylogenetic correction

PS phylogenetic signal, *BF* bite force, *FP* frequency peak

Error in Evolutionary Analysis

Comparisons between the three different scanning techniques for all 19 species identified consistent (although non-significant) evolutionary allometry patterns (Table 4). These were validated by PGLS analyses (Table 4). When testing for phylogenetic signal across the three datasets using the consensus tree, we obtained K_{multiv} values that were highly significant and close to one (Table 4). The signal was less strong for size but equally significant regardless of the technique (Table 4). The results for the association between morphological data and ecological data (peak frequency and bite force) are reported in Table 4 for each technique, with and without phylogenetic correction, and show a high degree of concordance between techniques.

Comparisons of parameter values obtained with the single-techniques (PH, µCT, LS) against our 1000 mixed datasets revealed similar means and standard deviations. Nevertheless, in most of the cases standard deviations were slightly greater for multi-technique datasets (Table S3).

When testing for variation due to the phylogenetic uncertainty and technique error the distributions of parameters estimates displayed similar shapes between techniques but in some cases the technique caused a shift in their location (i.e., allometry, phylogenetic signal for shape and correlation between shape and bite force) (Fig. S2). In particular, means of R^2 distributions for allometry differed between each technique (PH = 0.098; µCT = 0.105; LS = 0.099) but standard deviations did not (PH = µCT = LS = 0.004). A similar pattern is observed for the K_{multiv} of shape (mean: PH = 0.916, µCT = 0.936, LS = 0.969; standard deviation: PH = 0.024, µCT = 0.026 LS = 0.025) and R^2 for correlations between shape and bite force (mean: PH = 0.100, µCT = 0.105, LS = 0.107; standard deviation PH = µCT = LS = 0.004). Nevertheless, the p-values for K_{multiv} of shape were smaller than 0.001 for all combinations of trees/techniques. P-values for allometry and shape correlation with bite force equally resulted in coherent non-significant patterns ($p > 0.15$ in all cases). The ANOVA on the allometry estimates revealed that 36.35% ($p < 0.001$) of the variance in allometry was

explained by the technique employed, while 62.54% ($p < 0.001$) by the phylogenetic uncertainty. The ANOVA on the phylogenetic signal for size demonstrated that the majority of the variance was due to the phylogenetic uncertainty in the dataset (Table 5). The phylogenetic signal variance for shape is still mainly represented by the phylogenetic uncertainty (55.75%, $p < 0.001$) but a significant portion of the variance is due to the different technique employed (43.75%, $p < 0.001$). When the correlation between morphological data and frequency peak is computed the variance due to the technique error is significant but small (size: 1.15%, $p < 0.001$; shape: 2.04%, $p < 0.001$). Similar results were obtained for the correlation between bite force and size (0.35%, $p < 0.001$). Nevertheless, 37.00% of the correlation between bite force and shape was explained by the technique ($p < 0.001$) and 61.65% was explained by phylogenetic uncertainty ($p < 0.001$) (Table 5).

Discussion

Performance of Photogrammetry Technique

Analyses of mesh distances, shape visualisation (i.e., PCA graphs) and geometric morphometric error in the interspecific dataset demonstrated that PH, µCT and LS provide comparable raw material (i.e., centroid size and Procrustes coordinates) for GMM analyses. This was supported by high correlation coefficients for centroid size and Procrustes coordinates between the techniques and low proportion of variance explained by the techniques for both size and shape. This was in accordance with previous studies of much larger skulls, for example humans (Katz and Friess 2014) and wolves (Evin et al. 2016).

High intraclass correlation coefficients indicated high repeatability and reflected low random measurement error, which suggested that landmarking error was not important for our interspecific dataset. These coefficients (0.97–0.99) were similar to values previously obtained for human skulls (0.99; Badawi-Fayad and Cabanis 2007), kangaroo-size skulls (0.95; Fruciano et al. 2017) and small rodent skulls (0.75; Marcy et al. 2018). No technique-related differences in landmarking difficulties were found, based on Procrustes variance, which contrasts Marcy et al. (2018) finding of systematically better µCT relative to laser scans. This difference might be due to their use of a fast data collection scheme (10 min/sample) without employing additional measures to ensure quality of the models. Alternatively it could be linked to intrinsic differences in the LS and PH devices that were employed.

Experience plays an important role in identification and placement of landmarks (Osis et al. 2015; Sholts et al. 2011) and different approaches can induce different levels of

Table 5 ANOVAs on parameter estimates for allometry (R^2), phylogenetic signal (K_{multiV}) for size (PS Size) and shape (PS shape), and correlation (R^2) with ecological variables (*BF* bite force, *FP* frequency peak) computed by technique and using 1000 trees from the posterior distribution

		Df	SS	MS	Rsq	F value	Pr(> F)			Df	SS	MS	Rsq	F value	Pr(> F)
PS Size	Technique	2	1.133	0.567	0.068	34,190.410	< 0.001	PS Shape	Technique	2	1.447	0.724	0.438	90,648.232	< 0.001
	Tree	999	15.588	0.016	0.930	941.694	< 0.001		Tree	999	1.844	0.002	0.558	231.240	< 0.001
	Residuals	1998	0.033	0.000	0.002				Residuals	1998	0.016	0.000	0.005		
Size ~ BF	Technique	2	0.004	0.002	0.003	11,421.015	< 0.001	Shape ~ BF	Technique	2	0.028	0.014	0.37002	27,415.477	< 0.001
	Tree	999	1.052	0.001	0.996	6568.066	< 0.001		Tree	999	0.046	0.000	0.61649	91.444	< 0.001
	Residuals	1998	0.000	0.000	0.000				Residuals	1998	0.001	0.000	0.013		
Size ~ FP	Technique	2	0.004	0.002	0.011	16,665.770	< 0.001	Shape ~ FP	Technique	2	0.002	0.001	0.020	4069.128	< 0.001
	Tree	999	0.325	0.000	0.988	2876.151	< 0.001		Tree	999	0.078	0.000	0.975	388.964	< 0.001
	Residuals	1998	0.000	0.000	0.001				Residuals	1998	0.000	0.000	0.005		
Allometry	Technique	2	0.028	0.014	0.363	32,714.085	< 0.001								
	Tree	999	0.049	0.000	0.625	112.698	< 0.001								
	Residuals	1998	0.001	0.000	0.011										

systematic error (Marcy et al. 2018). In the current study, we did not specifically test for operator bias as previous studies reported inter-operator error being similar across different techniques (Robinson and Terhune 2017).

We also showed that centroid size and Procrustes coordinates extracted from PH models are suitable for subsequent macro-evolutionary analyses such as size-shape correlations (allometry), calculation of phylogenetic signal and correlation between morphological (size and shape) and ecological (frequency peak and bite force) data: parameters estimates were similar among techniques even when accounted for phylogenetic relatedness. All methods lead to the same biological interpretation, further confirming that PH provides suitable raw data for evolutionary analysis.

PH has several advantages in addition to being affordable and easy to use. It is particularly suitable when access to more expensive equipment is limited, where specimens cannot easily be transported, and/or where data collection has to take place in a remote location. Nevertheless, a significant down-side was the lack of detail achieved for teeth reconstruction and difficulties in reproducing thin structures (such as the zygomatic arch). Future studies may explore the use of focused stacking techniques in order to achieve a greater level of detail (Brecko et al. 2014; Nguyen et al. 2014; Santella and Milner 2017).

Mixed Data from Different Reconstruction Techniques

Our examination of multi-technique datasets revealed increases in standard deviations for allometry, phylogenetic signal and correlation with ecological variables compared with single-technique datasets, although this had no impact on the biological interpretation of the results. This suggests that multi-technique datasets could be potentially used (with caution and following exploratory studies), at least for interspecific analysis as long as the use of different techniques was relatively balanced across different groups (such as species, populations or sex). Mixing data from different devices is not recommended when researchers suspect a relatively small portion of biological variance in the sample (e.g. in populational studies).

When the same analyses were performed using the set of posterior trees, the interaction between phylogenetic uncertainty and technique became significant. However, the amount of parameter variation was relatively small and mainly due to the phylogenetic variation rather than technique error. Also, the general biological conclusions are essentially the same for almost all analyses (i.e., degrees of allometry and phylogenetic signal for size and variance explained by ecological variables). For instance, under the different techniques, bite force predicts between 8.85 and 11.94% of the shape variance supporting the inference that bite force moderately influences the shape evolution in bats. Fruciano et al. (2017) has pointed out that the phylogenetic signal in shape (as reflected by K statistics) is strongly influenced by both phylogenetic uncertainty and technique. In our sample, K_{multiv} varies from 0.85 to 1.05 between techniques which would lead to different evolutionary conclusions (Adams 2014; Blomberg et al. 2003), but the significance of K is unaffected. Revell et al. (2008) noted that K is indicative of statistical dependence between traits and phylogenetic relatedness but no inference on evolutionary rate and mode of evolution should be drawn from its value alone. Therefore, while we suggest that researchers should be cautious about inferring biological meaning from the magnitude of K for shape on mixed technique datasets, its significance can provide a reliable indicator of the presence of phylogenetic signal.

In conclusion, combining data acquired from model reconstructed with different techniques inevitably introduces an additional source of error. Its impact needs to be assessed according to whether it has an effect on the biological conclusions. Phylogenetic uncertainty can interact with other source of error (e.g. technique employed) suggesting preliminary test on phylogenetic comparative analysis are essential to identify possible not negligible source of error.

Acknowledgements The authors are grateful to the staff of the MNHN (including Jean-Marc Pons, Virginie Bouetel and Maité Adam) for providing access to the collections and infrastructures. This research received support from the SYNTHESYS 3 Project http://synthesys3.myspecies.info/ which is financed by the European Community Research Infrastructure Action under the FP7 (FR-TAF-6924 awarded to G. Giacomini). We wish to thank Isabelle De Groote and Peter Falkingham who provided initial guidance on photogrammetry techniques.

Data Accessibility 3D model available from the Morphosource repository: https://www.morphosource.org/Detail/MediaDetail/Show/media_id/30222.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

Adams, D. C. (2014). A generalized K statistic for estimating phylogenetic signal from shape and other high-dimensional multivariate

data. *Systematic Biology, 63*(5), 685–697. https://doi.org/10.1093/sysbio/syu030.

Adams, D. C., & Collyer, M. L. (2015). Permutation tests for phylogenetic comparative analyses of high-dimensional shape data: What you shuffle matters. *Evolution, 69*(3), 823–829. https://doi.org/10.1111/evo.12596.

Adams, D. C., & Otárola-Castillo, E. (2013). geomorph: An r package for the collection and analysis of geometric morphometric shape data. *Methods in Ecology and Evolution, 4*(4), 393–399. https://doi.org/10.1111/2041-210X.12035.

Aguirre, L. F., Herrel, A., van Damme, R., & Matthysen, E. (2002). Ecomorphological analysis of trophic niche partitioning in a tropical savannah bat community. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, *269*(1497), 1271–1278. http://rspb.royalsocietypublishing.org/content/269/1497/1271.abstract

Badawi-Fayad, J., & Cabanis, E. A. (2007). Three-dimensional procrustes analysis of modern human craniofacial form. *Anatomical Record, 290*(3), 268–276. https://doi.org/10.1002/ar.20442.

Bærentzen, J. A., & Henrik, A. (2002). Generating signed distance fields from triangle meshes. Informatics and mathematical modelling. Imm-Techni Report 2002–21.

Bates, K., Falkingham, P., & Rarity, F. (2010). Application of high-resolution laser scanning and photogrammetric techniques to data acquisition, analysis and interpretation in palaeontology. *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, *38*(5), 68–73. http://www.isprs.org/proceedings/xxxvIII/part5/papers/193.pdf

Blomberg, S. P., Garland, T., & Ives, A. R. (2003). Testing for phylogenetic signal in comparative data: Behavioral traits are more labile. *Evolution, 57*(4), 717–745. https://doi.org/10.1111/j.0014-3820.2003.tb00285.x.

Bookstein, F. L. (1989). "Size and shape": A comment on semantics. *Systematic Zoology, 38*(2), 173–180. https://doi.org/10.2307/2992387.

Brecko, J., Mathys, A., Dekoninck, W., Leponce, M., VandenSpiegel, D., & Semal, P. (2014). Focus stacking: Comparing commercial top-end set-ups with a semi-automatic low budget approach. A possible solution for mass digitization of type specimens. *ZooKeys, 464,* 1–23. https://doi.org/10.3897/zookeys.464.8615.

Brinkløv, S., Jakobsen, L., Ratcliffe, J. M., Kalko, E. K. V., & Surlykke, A. (2011). Echolocation call intensity and directionality in flying short-tailed fruit bats, *Carollia perspicillata* (Phyllostomidae). *The Journal of the Acoustical Society of America, 129*(1), 427–435. https://doi.org/10.1121/1.3519396.

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, the Italian Journal of Mammalogy, 25*(2), 73–81. https://doi.org/10.4404/hystrix-25.2-10993.

Cardini, A., & Polly, P. D. (2013). Larger mammals have longer faces because of size-related constraints on skull form. *Nature Communications, 4,* 2458. https://doi.org/10.1038/ncomms3458.

Cardini, A., Polly, 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.

Cornette, R., Baylac, M., Souter, T., & Herrel, A. (2013). Does shape co-variation between the skull and the mandible have functional consequences? A 3D approach for a 3D problem. *Journal of Anatomy, 223*(4), 329–336. https://doi.org/10.1111/joa.12086.

Cornwell, W., & Nakagawa, S. (2017). Phylogenetic comparative methods. *Current Biology, 27*(9), R333–R336. https://doi.org/10.1016/j.cub.2017.03.049.

Drake, A. G., & Klingenberg, C. P. (2010). Large-scale diversification of skull shape in domestic dogs: Disparity and modularity. *The American Naturalist, 175*(3), 289–301. https://doi.org/10.1086/650372.

Durão, A. F., Muñoz-Muñoz, F., Martínez-Vargas, J., & Ventura, J. (2018). Obtaining three-dimensional models of limb long bones from small mammals: A photogrammetric approach. In C. Rissech, L. Lloveras, J. Nadal, & J. M. Fullola (Eds.), *MONOGRAFIES 14. Geometric morphometrics. Trends in biology, paleobiology and archaeology*. Barcelona: Seminari d'Estudis i Recerques Prehistòriques.

Ege, A., Seker, D. Z., Tuncay, I., & Duran, Z. (2004). Photogrammetric analysis of the articular surface of the distal radius. *Journal of International Medical Research, 32*(4), 406–410. https://doi.org/10.1177/147323000403200409.

Evin, A., Souter, T., Hulme-Beaman, A., Ameen, C., Allen, R., Viacava, P., et al. (2016). The use of close-range photogrammetry in zooarchaeology: Creating accurate 3D models of wolf crania to study dog domestication. *Journal of Archaeological Science: Reports, 9,* 87–93. https://doi.org/10.1016/j.jasrep.2016.06.028.

Fahlke, J., & Autenrieth, M. (2016). Photogrammetry VS. Micro-CT scanning for 3D surface generation of typical vertebrate fossil—a case study. *Journal of Paleontological Techniques, 14,* 1–18.

Falkingham, P. L. (2012). Acquisition of high resolution three-dimensional models using free, open-source, photogrammetric software. *Palaeontologia Electronica, 15*(1), 1–15. https://doi.org/10.1371/journal.pone.0004532.

Fourie, Z., Damstra, J., Gerrits, P. O., & Ren, Y. (2011). Evaluation of anthropometric accuracy and reliability using different three-dimensional scanning systems. *Forensic Science International, 207*(1), 127–134. https://doi.org/10.1016/j.forsciint.2010.09.018.

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.

Huelsenbeck, J. P., & Ronquist, F. (2001). MRBAYES: Bayesian inference of phylogenetic trees. *Bioinformatics, 17*(8), 754–755. https://doi.org/10.1093/bioinformatics/17.8.754.

Kalko, E. K. V., Schnitzler, H., Kaipf, I., & Grinnell, A. D. (1998). Echolocation and foraging behavior of the lesser bulldog bat, *Noctilio albiventris*: Preadaptations for piscivory? *Behavioral Ecology and Sociobiology, 42*(5), 305–319.

Katz, D., & Friess, M. (2014). Technical note: 3D from standard digital photography of human crania: A preliminary assessment. *American Journal of Physical Anthropology, 154*(1), 152–158. https://doi.org/10.1002/ajpa.22468.

Mallison, H., & Wings, O. (2014). Photogrammetry in paleontology: A practical guide. *Journal of Paleontological Techniques, 12,* 1–31.

Marcy, A. E., Fruciano, C., Phillips, M. J., Mardon, K., & Weisbecker, V. (2018). Low resolution scans can provide a sufficiently accurate, cost- and time-effective alternative to high resolution scans for 3D shape analyses. *PeerJ, 6,* e5032. https://doi.org/10.7717/peerj.5032.

Muñoz-Muñoz, F., Quinto-Sánchez, M., & González-José, R. (2016). Photogrammetry: A useful tool for three-dimensional morphometric analysis of small mammals. *Journal of Zoological Systematics and Evolutionary Research, 54*(4), 318–325. https://doi.org/10.1111/jzs.12137.

Nguyen, C. V., Lovell, D. R., Adcock, M., & La Salle, J. (2014). Capturing natural-colour 3D models of insects for species discovery and diagnostics. *PLoS ONE, 9*(4), e94346. https://doi.org/10.1371/journal.pone.0094346.

Osis, S. T., Hettinga, B. A., Macdonald, S. L., & Ferber, R. (2015). A novel method to evaluate error in anatomical marker placement using a modified generalized Procrustes analysis. *Computer Methods in Biomechanics and Biomedical Engineering, 18*(10), 1108–1116. https://doi.org/10.1080/10255842.2013.873034.

Paradis, E., Claude, J., & Strimmer, K. (2004). APE: Analyses of phylogenetics and evolution in R language. *Bioinformatics, 20*(2), 289–290. https://doi.org/10.1093/bioinformatics/btg412.

R Core Team. (2018). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna. https://www.R-project.org/.

Revell, L. J., Harmon, L. J., & Collar, D. C. (2008). Phylogenetic signal, evolutionary process, and rate. *Systematic Biology, 57*(4), 591–601. https://doi.org/10.1080/10635150802302427.

Robinson, C., & Terhune, C. E. (2017). Error in geometric morphometric data collection: Combining data from multiple sources. *American Journal of Physical Anthropology, 164*(1), 62–75. https://doi.org/10.1002/ajpa.23257.

Rodríguez-San Pedro, A., & Allendes, J. L. (2017). Echolocation calls of free-flying common vampire bats *Desmodus rotundus* (Chiroptera: Phyllostomidae) in Chile. *Bioacoustics, 26*(2), 153–160. https://doi.org/10.1080/09524622.2016.1231079.

Rohlf, F. J. (2007). Comparative methods for the analysis of continuous variables: geometric interpretations. *Evolution, 55*(11), 2143–2160. https://doi.org/10.1111/j.0014-3820.2001.tb00731.x.

Russo, D., & Jones, G. (2002). Identification of twenty-two bat species (Mammalia: Chiroptera) from Italy by analysis of time-expanded recordings of echolocation calls. *Journal of Zoology, 258*(1), S0952836902001231. https://doi.org/10.1017/S0952836902001231.

Santella, M., & Milner, A. R. (2017). Coupling focus stacking with photogrammetry to illustrate small fossil teeth. *Journal of Paleontological Techniques, 18,* 1–17.

Schlager, S. (2013). Morpho: Calculations and visualizations related to Geometric Morphometrics. R package version 0.17. http://sourceforge.net/projects/morpho-rpackage/.

Shi, J. J., & Rabosky, D. L. (2015). Speciation dynamics during the global radiation of extant bats. *Evolution, 69*(6), 1528–1545. https://doi.org/10.1111/evo.12681.

Sholts, S. B., Flores, L., Walker, P. L., & Wärmländer, S. K. T. S. (2011). Comparison of coordinate measurement precision of different landmark types on human crania using a 3D laser scanner and a 3D digitiser: Implications for applications of digital morphometrics. *International Journal of Osteoarchaeology, 21*(5), 535–543. https://doi.org/10.1002/oa.1156.

Siemers, B. M., Kalko, E. K. V., & Schnitzler, H. U. (2001). Echolocation behavior and signal plasticity in the Neotropical bat *Myotis nigricans* (Schinz, 1821) (Vespertilionidae): A convergent case with European species of Pipistrellus? *Behavioral Ecology and Sociobiology, 50*(4), 317–328. https://doi.org/10.1007/s002650100379.

Siemers, B. M., & Schnitzler, H.-U. (2004). Echolocation signals reflect niche differentiation in five sympatric congeneric bat species. *Nature, 429,* 657. https://doi.org/10.1038/nature02547.

Simmons, N. B., & Geisler, J. H. (1998). Phylogenetic relationships of Icaronycteris, Archaeonycteris, Hassianycteris, and Palaeochiropteryx to extant bat lineages, with comments on the evolution of echolocation and foraging strategies in Microchiroptera. *Bulletin of the American Museum of Natural History, 235,* 1–182.

Sztencel-Jabłonka, A., Jones, G., & Bogdanowicz, W. (2009). Skull morphology of two cryptic bat species: *Pipistrellus pipistrellus* and *P. pygmaeus*—a 3D geometric morphometrics approach with landmark reconstruction. *Acta Chiropterologica, 11*(1), 113–126. https://doi.org/10.3161/150811009x465730.

Von Cramon-Taubadel, N., Frazier, B. C., & Mirazo, M. (2007). The problem of assessing landmark error in geometric morphometrics: Theory, methods, and modifications. *American Journal of Physical Anthropology, 134*(1), 24–35. https://doi.org/10.1002/ajpa.

Wiley, D. F., Amenta, N., Alcantara, D. A., Ghosh, D., Kil, Y. J., Delson, E., et al. (2005). Evolutionary Morphing. In *Proceedings of IEEE Visualization* (pp. 431–438). https://doi.org/10.1109/visual.2005.1532826

Zelditch, M., Swiderski, D., & Sheets, H. (2012). *Geometric morphometrics for biologists a primer* (2nd ed., p. 488). Cambridge: Academic Press.