

Detecting Phylogenetic Signal and Adaptation in Papionin Cranial Shape by Decomposing Variation at Different Spatial Scales

NICOLE D.S. GRUNSTRA^{1,2,3}, SILVESTER J. BARTSCH^{1,4}, ANNE LE MAÎTRE^{1,5}, AND PHILIPP MITTEROECKER^{1,2,*}

¹Department of Evolutionary Biology, University of Vienna, Vienna, Austria; ²Konrad Lorenz Institute for Evolution and Cognition Research, Klosterneuburg, Austria; ³Mammal Collection, Natural History Museum Vienna, Vienna, Austria; ⁴Department of Biomedical Imaging and Image-guided Therapy, Medical University of Vienna, Vienna, Austria; and ⁵PALEVOPRIM - UMR 7262 CNRS INEE, Université de Poitiers, Poitiers, France

*Correspondence to be sent to: Althanstrasse 14, A-1090 Vienna, Austria;

E-mail: philipp.mitteroecker@univie.ac.at

Received 29 March 2020; reviews returned 22 November 2020; accepted 26 November 2020

Associate Editor: Jacob Esselstyn

Abstract.—Phylogenetic reconstruction based on morphometric data is hampered by homoplasies. For example, many similarities in cranial form between primate taxa more strongly reflect ecological similarities rather than phylogenetic relatedness. However, the way in which the different cranial bones constitute cranial form is, if at all, of less functional relevance and thus largely hidden from selection. We propose that these “constructional details” are better indicators of phylogenetic history than any large-scale shape feature or raw form variable. Within a geometric morphometric context, we show how to analyze the relative extent of bones independently of differences in overall shape. We also show how to decompose total shape variation into small-scale and large-scale shape variation. We apply both methods to the midsagittal cranial morphology of papionin monkeys, which are well known for the discrepancy between morphological similarities and phylogenetic relationships. We study phylogenetic signal and functional adaptation using a molecular phylogeny and contextual data on feeding ecology and locomotor behavior. As expected, total cranial shape, bone outline shape, and large-scale shape features were only weakly associated with phylogenetic distance. But the relative bone contributions and small-scale shape features were both highly correlated with phylogenetic distances. By contrast, the association with ecological and behavioral variables was strongest for the outline shape and large-scale shape features. Studies of morphological adaptation and phylogenetic history thus profit from a decomposition of shape variation into different spatial scales. [Adaptation; canalization; cranial shape; geometric morphometrics; papionini; partial warps; phylogeny.]

Morphological approaches to the phylogenetic reconstruction of extant and recently extinct species have largely been superseded by molecular methods (e.g., Felsenstein 2003; Yang and Rannala 2012; Lee and Palci 2015). In paleontology and paleoanthropology, however, where organic remains are largely unavailable, phenetic and cladistic methods are still common (Sneath and Sokal 1973; Felsenstein 1982, 2003; Kitching et al. 1989). Especially in vertebrate paleontology and systematics, phenetic approaches (mostly on low taxonomic levels) have regained popularity through advancements in geometric morphometrics (Bookstein 1991, 2018; MacLeod 2002; Mitteroecker and Gunz 2009; Klingenberg 2010; Mitteroecker 2020).

The main weakness of phenetic and morphometric methods is that phylogenetic relationships are often obscured by homoplasies. Whereas random mutations of neutral genetic markers accumulate over time and thus often reliably trace the phylogenetic history, many morphological traits are functionally relevant and subject to natural selection. Hence, similarities in quantitative morphological traits (such as primate craniodental traits) between taxa can be more strongly shaped by ecological similarities and joint environmental influences on development rather than recent phylogenetic divergence (e.g., Collard and Wood 2000, 2001; Madsen et al. 2001; Collard and O’Higgins 2002).

The choice of traits is crucial to any phylogenetic approach (e.g., Poe and Wiens 2000; Caumul and Polly 2005). Like in the molecular approaches, it would be

desirable for improving phylogenetic reconstruction based on morphometric data to separate functional traits from “neutral traits” that are not functionally relevant in the studied lineages and thus subject to evolutionary drift. For a set of several neutral traits, phenotypic divergence might indeed reflect phylogenetic divergence. Based on this rationale, some researchers have tried to identify traits that more closely correlate with genetic or phylogenetic distances than other traits. For instance, some regions of the primate cranium, such as the cranial base and the temporal bone, were reported to reflect phylogenetic relationships among species or populations more reliably than facial traits (Lockwood et al. 2004; Harvati and Weaver 2006a,b; Cardini and Elton 2008; Smith 2009). Yet other studies did not find strong support of one cranial region over another (Roseman et al. 2010; von Cramon-Taubadel 2009, 2011).

However, in a complex anatomical structure such as the vertebrate cranium, which accommodates the brain, the airways, the masticatory apparatus, and various sensory organs, few morphological traits or anatomical dimensions are completely neutral; most of them contribute to one or multiple functions. Furthermore, cranial components and dimensions are tightly integrated, so that functional adaptation in one component is likely to influence other cranial regions as well (Ackermann and Cheverud 2004; Bastir and Rosas 2006; Lockwood 2007; Mitteroecker and Bookstein 2008; Hallgrímsson et al. 2009; Klingenberg 2013). Hence, cranial form represents a mosaic of adaptations and,

overall, performs poorly as a phylogenetic indicator, at least at lower taxonomic levels (e.g., genera or families).

Even though the size and shape of the functional units of the cranium are all likely to be under strong stabilizing or directional selection in different lineages, the way in which the various cranial bones constitute cranial form is, if at all, of less functional relevance. For example, the upper jaw houses the upper dentition and is important for mastication. But while the total length of the upper jaw is functionally relevant, the extent to which it is made up of premaxilla, maxilla, and palatine is likely less important. Similarly, the cranial vault protects the brain and serves as the attachment site for numerous masticatory and nuchal muscles; its shape thus is of functional relevance and likely under selection in most vertebrates. However, the extent to which the cranial vault is composed of frontal, parietal, temporal, and occipital bones may be functionally irrelevant (O'Higgins and Johnson 1988; Mitteroecker et al. 2020).

We therefore propose that the *relative contributions* of individual bones to overall shape are better indicators of phylogenetic history than any raw shape or form variable. Constraints in these relative bone dimensions more likely reflect developmental rather than direct functional constraints. Clearly, actual evolutionary rate and phylogenetic signal differ across lineages, time scales, and traits, but we expect that—averaged over different lineages and bones—the evolution of relative bone dimensions is more characterized by drift than selection when compared to overall, large-scale bone form.

Indeed, in a decomposition of human cranial shape variation into different spatial scales, Bookstein (2015) and Mitteroecker et al. (2020) found that small-scale shape feature (deformations of closely adjacent landmarks) vary considerably more across individuals than large-scale features (deformations involving more distant landmarks). In other words, the “constructional details” of cranial shape variation, including the relative dimensions of the cranial bones, seem to be less canalized than more functionally relevant large-scale features, presumably due to relaxed stabilizing selection. Hence, we expect that—largely hidden from selection—these relative contributions of the bones mainly drift during evolution (a kind of “developmental systems drift”; True and Haag 2001).

Within a geometric morphometric context, we show here how to analyze variation in the relative contributions of bones independently of differences in overall shape. This approach involves a decomposition of shape variation into the variation of overall outline shape and the variation of relative bone contributions, computed as the landmark variation after standardizing for differences in outline shape (using thin-plate spline warping). We also use a second approach to separate small-scale from large-scale shape variation based on partial warps (Bookstein 1989, 2015; Mitteroecker et al. 2020). We apply both methods to the midsagittal cranial morphology of papionin monkeys (macaques, baboons, mangabeys, and relatives). These species are well

known for the discrepancy between their morphological similarities and phylogenetic relationships resulting from strong cranial homoplasies (e.g., Strasser and Delson 1987; Delson and Dean 1993; Disotell 1994; Harris and Disotell 1998; Fleagle and McGraw 1999; Collard and Wood 2001; Collard and O'Higgins 2002; Singleton 2002; Leigh 2007). We study phylogenetic signal (i.e., the statistical correspondence between phenotypic similarity and phylogenetic relatedness; Blomberg and Garland 2002) and functional adaptation using a molecular phylogeny and contextual data on feeding ecology and locomotor behavior. We expect that the relative bone dimensions and small-scale shape features better reflect phylogenetic relatedness than overall cranial shape and large-scale shape features. By contrast, we expect that ecological and behavioral differences between these species are better reflected in overall, large-scale cranial shape. The data along with a new R package (prWarp) and a worked example for these decompositions of shape variation are available as [Supplementary material](http://dx.doi.org/10.5061/dryad.zkh189373) from Dryad (<http://dx.doi.org/10.5061/dryad.zkh189373>) and CRAN (<https://cran.r-project.org/package=prWarp>).

MATERIALS AND METHODS

Data

Our sample comprises 61 specimens from 16 different papionin species as well as six specimens from two nonpapionin Old World monkey taxa, *Cercopithecus mitis* and *Colobus guereza*, which represent the sister taxon of the papionini (the cercopithecini) and the sister taxon to the cercopithecinae (the colobinae), respectively (Table 1). Figure 1a illustrates the phylogenetic relationships among these taxa, based on several mitochondrial, Y-chromosomal, and autosomal markers (Arnold et al. 2010). Branch lengths are scaled to time. Only adult females were sampled in order to minimize ontogenetic variation and sexual dimorphism.

Midsagittal cranial morphology was digitized by 28 anatomical landmarks and 42 sliding landmarks (Fig. 1b, [Supplementary Table S1](#) available on Dryad) on 3D computed tomography (CT) scans using Avizo software (version 6.4.0). CT scans were freely obtained from MorphoSource, the Digital Morphology Museum (KUPRI), and the Smithsonian. Additional skulls from the Natural History Museum Vienna were scanned in the Department of Anthropology at the University of Vienna. The landmark scheme was designed to delineate the midsagittal geometry of the different cranial bones. As these midsagittal landmarks do not lie exactly on a plane, they were measured in 3D and then projected onto a least squares-fitted plane for every specimen. This was achieved by conducting, for each specimen separately, a principal component analysis (PCA) of the 3D landmark coordinates and retaining the scores of the first two PCs.

In order to represent environmental and behavioral aspects to which the primate cranium may have adapted

TABLE 1. Sample composition and sample size. Specimens include adult females only. Taxonomy according to Wilson and Reeder (2005), except for *P. hamadryas*, for which we follow Jolly (1993) and consider it a single, but polytypic species.^a

Species	(μ)CT (N)
Cercopithecinae	
Papionini	
<i>Cercocebus agilis</i>	4
<i>C. atys</i>	2
<i>C. torquatus</i>	3
<i>Lophocebus albigena</i>	6
<i>L. aterrimus</i>	1
<i>Macaca fascicularis</i>	4
<i>M. fuscata</i>	4
<i>M. mulatta</i>	4
<i>M. nemestrina</i>	4
<i>M. nigra</i>	3
<i>M. silenus</i>	3
<i>M. sinica</i>	4
<i>M. sylvanus</i>	2
<i>Mandrillus sphinx</i>	5
<i>Papio hamadryas sensu lato</i>	8
<i>Theropithecus gelada</i>	4
Cercopithecini	
<i>Cercopithecus mitis</i>	3
Colobinae	
<i>Colobus guereza</i>	3

^aSavanna baboon (*Papio*) systematics is contentious; anywhere between one and six species are recognized (e.g., Groves 2001; Jolly 1993; Zinner et al. 2013a). They show overlap in geographic ranges, diets, and continuous, clinal variation in their ecology (Frost et al. 2003; Kamilar 2006). Although several baboon populations can be distinguished genetically (Zinner et al. 2009), gene flow and introgression between populations are well-documented (Jolly et al. 2011; Zinner et al. 2009, 2011, 2013b), and ecomorphological studies have not found clear differentiation among (sub)species (Frost et al. 2003; Dunn et al. 2013). We thus follow Jolly (1993) and others (e.g., Dunn et al. 2013; Gilbert et al. 2018) and consider *Papio* a single species. Our sample does not include specimens from the Kinda baboon, which is an outlier in terms of body size.

(e.g., mechanical demands associated with foods of different hardness and toughness), we obtained data on percentages of fruits, seeds, and leaves in the diet, as well as percentages of time spent on the ground, foraging (including feeding), and traveling (Supplementary Table S2 available on Dryad). As we focus on between-species variation, we averaged available data by species for as many populations as were available from published sources (e.g., Rowe and Myers 2011; Mittermeier et al. 2013; see Supplementary Table S2 available on Dryad).

Approach 1: Relative Bone Contributions

We selected 54 of the 70 landmarks that capture the overall geometry of the cranial outline and, thus, most functionally relevant aspects of cranial form. This “outline subset” includes all the landmarks on the outline of the facial and basicranial midsection as well as the landmarks on the outer and inner shell of the cranial vault (Fig. 1b, Supplementary Fig. S1 available on Dryad). We expected the shape of these cranial structures to be functionally relevant, but not the extent to which the different bones contribute to

these structures. For this outline landmark set, we thus treated the Type I landmarks (Bookstein 1991) on the boundary between bones (such as bregma, nasion, sphenobasion) as semilandmarks (Bookstein 1997; Gunz and Mitteroecker 2013). In other words, we retained only their position perpendicular to the bone outline and discarded information on its position along the bone. After sliding these semilandmarks (by minimizing the bending energy between each individual and the sample mean shape), the landmarks and semilandmarks of the outline subset represented the overall shape of the cranium without any information on the relative contributions of the separate bones. The only anatomical landmarks in this set were the Type II landmarks (curvature maxima), such as prosthion, basion, and opisthion. All landmarks in the outline subset were then subjected to generalized Procrustes analysis (GPA) in order to standardize for differences in position, scale, and orientation among the landmark configurations (Rohlf and Slice 1990; Mittermeier et al. 2013; Fig. 2 middle panels).

Next, for each specimen, the full configuration of the original 70 landmarks (including the Type I landmarks) was warped to the average shape configuration of the outline landmarks using thin-plate spline interpolation (Bookstein 1989, 1991). This mapping from the original configuration to the average outline shape was based on the outline landmarks only. As a result, the warped landmark configurations all have exactly the same outline shape, but vary in the landmarks on sutures and synchondroses that demarcate the borders between bones (and also in the landmarks on the nasal septum and vomer, which may have little functional relevance, at least their variation in the midsagittal plane). In this “residual landmark set,” the sutures and synchondroses between the bones are represented by anatomical landmarks (Type I landmarks), and only the semilandmarks on smooth outlines were allowed to slide (Fig. 2, Supplementary Fig. S1 available on Dryad). Note that after the warping to the mean outline shape (a nonrigid registration), no further Procrustes registration is necessary for these residual data.

Approach 2: Small-Scale versus Large-Scale Shape Variation

In a second approach, we decomposed the nonaffine (i.e., nonlinear or local) shape variation of the full set of 70 landmarks into their *partial warps* (Bookstein 1989, 1991). These are shape features (orthogonal linear combinations of the shape coordinates) arising from an eigenvalue decomposition of the “bending energy matrix.” Every partial warp has two components, one for the *x* and one for the *y* coordinates (shape deformations in the horizontal and vertical direction, respectively), as well as an associated *bending energy* that serves as an inverse measure of squared spatial scale (Supplementary Fig. S3 available on Dryad). Because of their computation via an eigenvalue decomposition, the first partial warp has the highest bending energy (smallest spatial scale), the

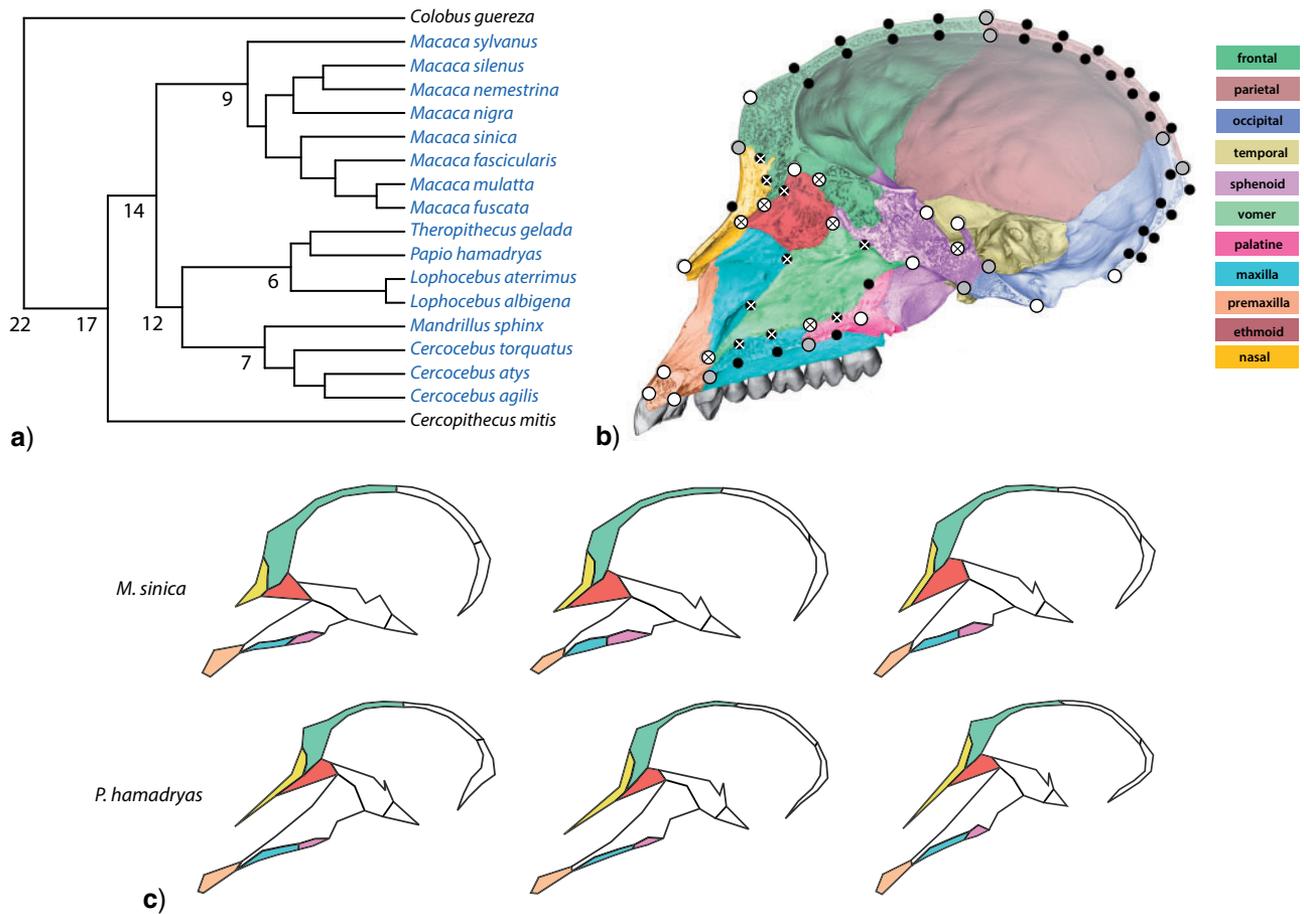


FIGURE 1. a) Molecular phylogeny, including approximate divergence times (in million years), of the papionins (blue) and the two outgroup Old World monkey taxa (black). Data are from *10kTrees* (Arnold et al. 2010). b) Landmark scheme shown for a *Macaca fuscata* specimen. The crossed-out landmarks are not part of the outline subset. White circles represent the anatomical landmarks and black circles the semilandmarks. Gray circles are landmarks that are treated as anatomical landmarks in the full set and in the residual subset but as semilandmarks in the outline subset. After standardizing outline shape variation, only the crossed-out landmarks and the gray landmarks are free to vary in the residual data (see Fig. 2). c) Three specimens of *Macaca sinica* and *Papio hamadryas*, with the different bones represented by polygons based on the measured landmarks and semilandmarks. Note how the relative dimensions of premaxilla, maxilla, and palatine vary across specimens despite relatively homogenous jaw form within species. Likewise, frontal bone, nasal bone, and ethmoid vary considerably in their contribution to the upper face.

second partial warp the second-highest bending energy, and so on. The last partial warps have the smallest bending energies and thus the largest spatial scales. Even though the partial warps are not *per se* biologically or phylogenetically meaningful, they allow for a separation of small-scale and large-scale shape variation (e.g., Rohlf 1988; MacLeod 2001, 2002; Bookstein 2018; Mitteroecker et al. 2020), which involves an (arbitrary) threshold that splits the partial warps.

For a self-similar distribution of cranial shape—as a null-model of completely unintegrated, independently varying cranial components—the variance of partial warps is expected to increase linearly with the inverse of bending energy (Mardia et al. 2006; Bookstein 2015; Mitteroecker et al. 2020). Hence, in order to study the spatial distribution of shape variation, we plotted the variance of each partial warp against inverse bending energy on a log–log scale. A slope of 1 in this plot indicates a self-similar shape distribution, whereas a

slope > 1 indicates an excess of large-scale variation and a slope < 1 an excess of small-scale shape variation (Bookstein 2015; Mitteroecker et al. 2020).

Phylogenetic and Ecological Signals

We calculated the magnitude of the phylogenetic signal by correlating pairwise phylogenetic distances with the Procrustes distances between the corresponding species mean shapes in five different data sets: all landmarks, the outline landmarks, the residual landmarks, as well as the small-scale and large-scale components of all landmarks. For each pair of species, phylogenetic distances were computed as the distance from one of the tips to the most recent ancestral node in the tree in Figure 1a (an ultrametric tree with branch lengths scaled to time). These distances are expected to be proportional to the time of independent

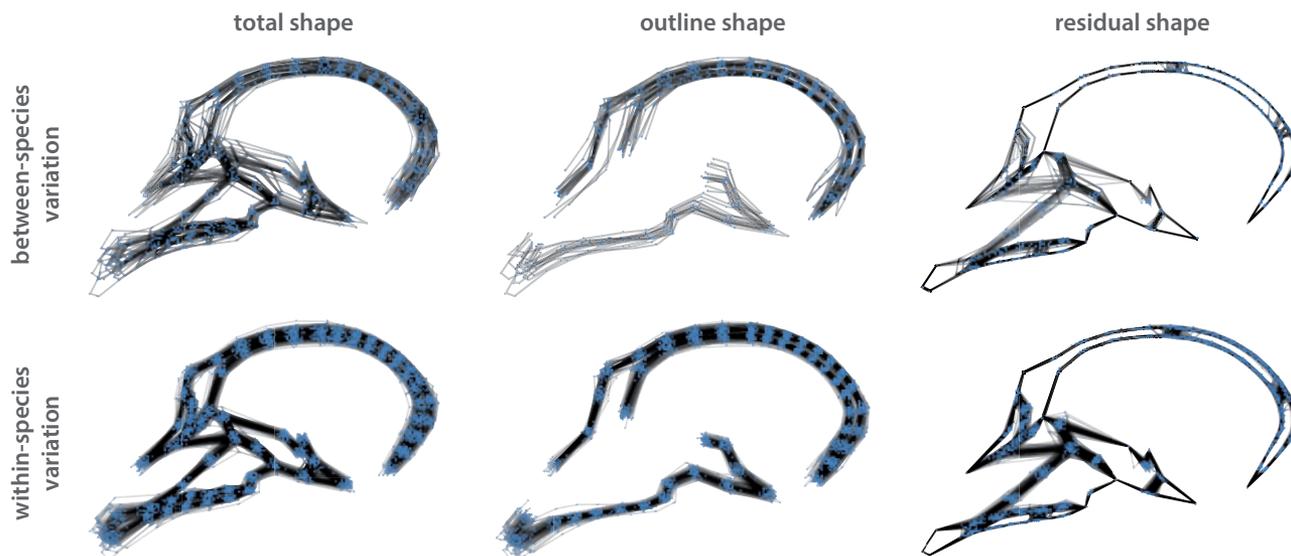


FIGURE 2. (Upper panel) Cranial shape variation *between* the mean configurations of the 18 species: total shape variation of the 70 landmarks (left), shape variation of the 54 outline landmarks (middle), and residual shape variation, that is, variation of all 70 landmarks after warping them to the same outline shape (right). Landmarks that demarcate borders between bones are treated as semilandmarks for the outline shape. For the residual shape, these landmarks are treated as anatomical landmarks and, thus, represent the differential contribution of the various bones to overall shape. (Lower panel) Individual cranial shape variation *within* the species: total shape variation (left), outline shape variation (middle), and residual shape variation (right). Note that outline shape varies considerably more between the species than within the species (total shape variation differs by a factor of 2.3), whereas residual shape variation is of more similar magnitude between and within species (a factor of 1.5).

evolution of the two species since their divergence from the last common ancestor. Statistical significance was estimated by Mantel's test using 5000 random permutations (Mantel 1967).

We built phylogenetic tree topologies from the different shape data for the African papionin genera (*Papio*, *Theropithecus*, *Mandrillus*, *Cercocebus*, *Lophocebus*), which exhibit the most pronounced craniodental homoplasies (Collard and Wood 2001; Singleton 2002; Supplementary Fig. S2 available on Dryad). We used two different tree-building algorithms: neighbor-joining (NJ) and unweighted pair group method with arithmetic mean (UPGMA). Both are agglomerative, distance-based methods. The NJ algorithm (Saitou and Nei 1987) finds sister group relationships by minimizing the sum of the tree's branch lengths and allows for different evolutionary rates in the tree and, hence, different branch lengths between sister groups. The hierarchical clustering method of UPGMA reconstructs sister relationships based on minimum average pairwise distances (Sneath and Sokal 1973). Each element is assumed to contribute equally to the average of a cluster, leading to equal branch lengths between sister groups. We reconstructed phylogenetic trees separately for outline and residual shapes as well as for small-scale and large-scale shape features. Phenotypic distance matrices were constructed from pairwise Procrustes distances between the species mean shapes.

To measure the multivariate association of the environmental variables with the five sets of shape variables, we performed a multivariate multiple

regression of each set of cranial shape variables on the six environmental variables. The resulting multivariate coefficients of determination (R^2) describe the fractions of variance in cranial shape (pooled over all variables) that can be explained by the environmental variables.

To explore a taxonomic-level effect, we carried out the matrix correlations and the multiple multivariate regressions on the total sample of 18 species as well as on the 16 papionins only. All analyses were performed in Mathematica 12.0 (Wolfram Research) and R (R Core Team 2019).

RESULTS

Shape Variation

The variation in cranial shape between the 18 species mean configurations and also within the species (pooled individual within-species variation) differed among total cranial shape (all landmarks), outline shape, and residual shape (Fig. 2). The TPS warping successfully standardized for variation in outline shape as only the variation in relative bone contributions was visible in the residual landmarks (right panel of Fig. 2). Outline shape varied considerably more between the species than within species (total variances of 0.0036 and 0.0015), whereas residual shape variation was more similar in magnitude (total variances of 0.0027 between-species and 0.0018 within-species).

We considered the first 50 partial warps (PWs) as small-scale shape features and the remaining PWs 51–67 as large-scale features because this separation yielded the greatest difference in phylogenetic signal (see below;

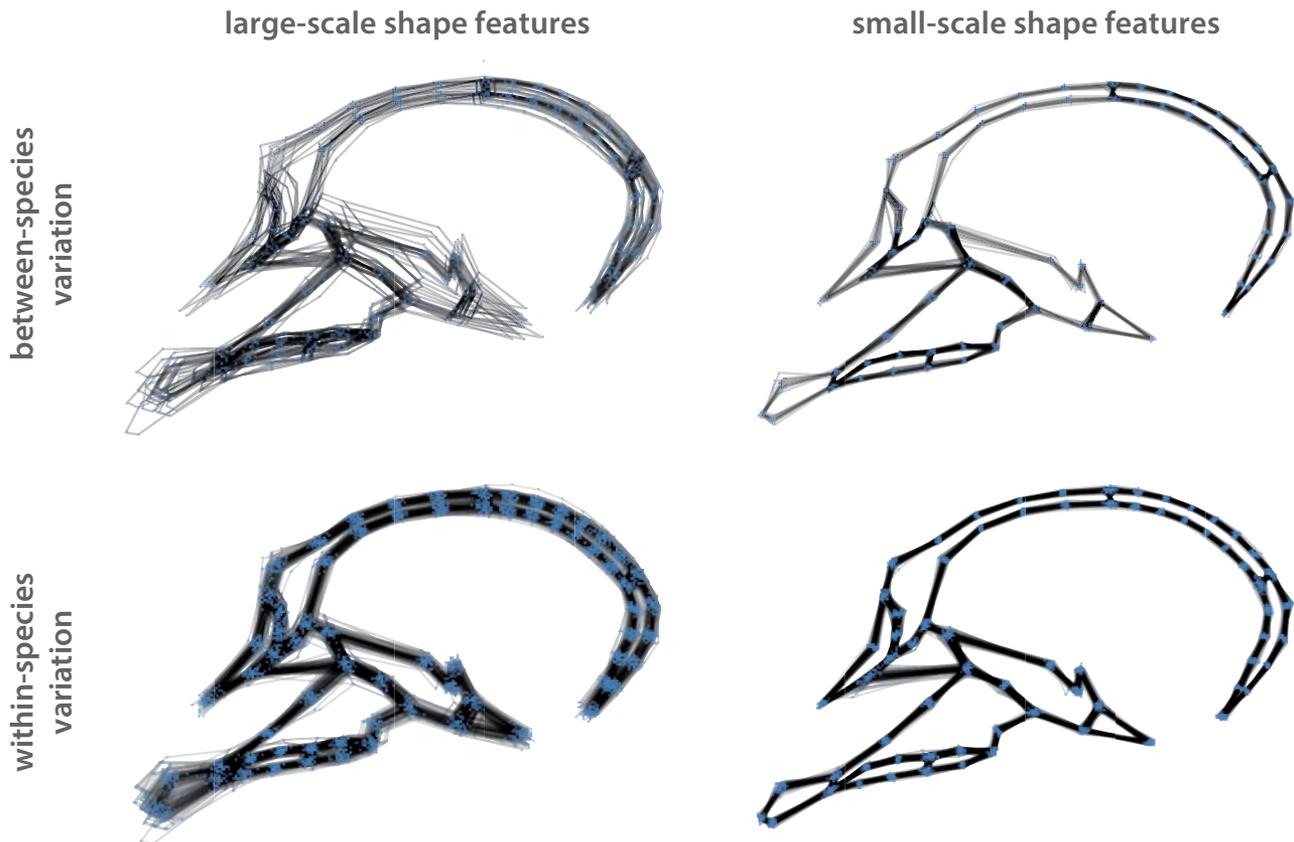


FIGURE 3. Large-scale and small-scale shape variation *between* the 18 species (upper panel) and *within* the species (lower panel). Small-scale variation was captured by the first 50 partial warps, and large-scale variation by the remaining 17 partial warps. Large-scale shape features vary 2.2 times more between species than within species, whereas small-scale shape shows about the same magnitude of variation within and between species.

but different thresholds led to qualitatively comparable results). Similar to outline shape variation, large-scale shape variation primarily comprised variation in relative facial length and orientation as well as in the shape of the brow ridge and the cranial base (Fig. 3). Small-scale shape variation comprised differences in the relative bone contributions but, unlike residual shape variation, also variation in cranial vault thickness. Variation in cranial base geometry was captured by the residual shape but was relatively constant for the small-scale shape (Figs. 2 and 3). Small-scale shape features showed about the same amount of variation within species and between species (total variances of 0.00028 and 0.00027, respectively). Large-scale features, by contrast, were more than twice as variable between species than within species (total variances of 0.0029 and 0.0013). As a result, individuals belonging to the different species showed a better group separation in a PCA of large-scale shape whereas they widely overlapped in a PCA of small-scale shape (Supplementary Fig. S3 available on Dryad).

This unequal distribution of shape variation across spatial scales is also apparent from the plots of partial warp variance against inverse bending energy (Fig. 4). The linear regression slope in the log-log plots was

0.74 for within-species variation and 0.93 for between-species variation. Hence, large-scale features were more variable than small-scale shape features both within and between species. But shape variation between the species means was close to a self-similar shape distribution, indicating that *relative to scale* all shape features showed a similar degree of evolutionary divergence. Within-species variation, by contrast, was characterized by an excess of small-scale shape variation, not only absolutely but also relative to scale.

The first two principal components (PCs) of outline shape mainly represented facial length and orientation, along with neurocranial globularity. The first PC of residual shape corresponded to the relative contributions of the frontal, parietal, and occipital bones to the cranial vault as well as to the position of the frontonasal suture (Fig. 5). PC 2 of residual shape mainly represented the relative dimensions of the ethmoid, frontal, and nasal bones as well as the relative contributions of the premaxilla, maxilla, and palatine to the upper jaw. The species mean shapes of the major phylogenetic lineages clustered better in the PCs of residual shape as compared with those of outline shape.

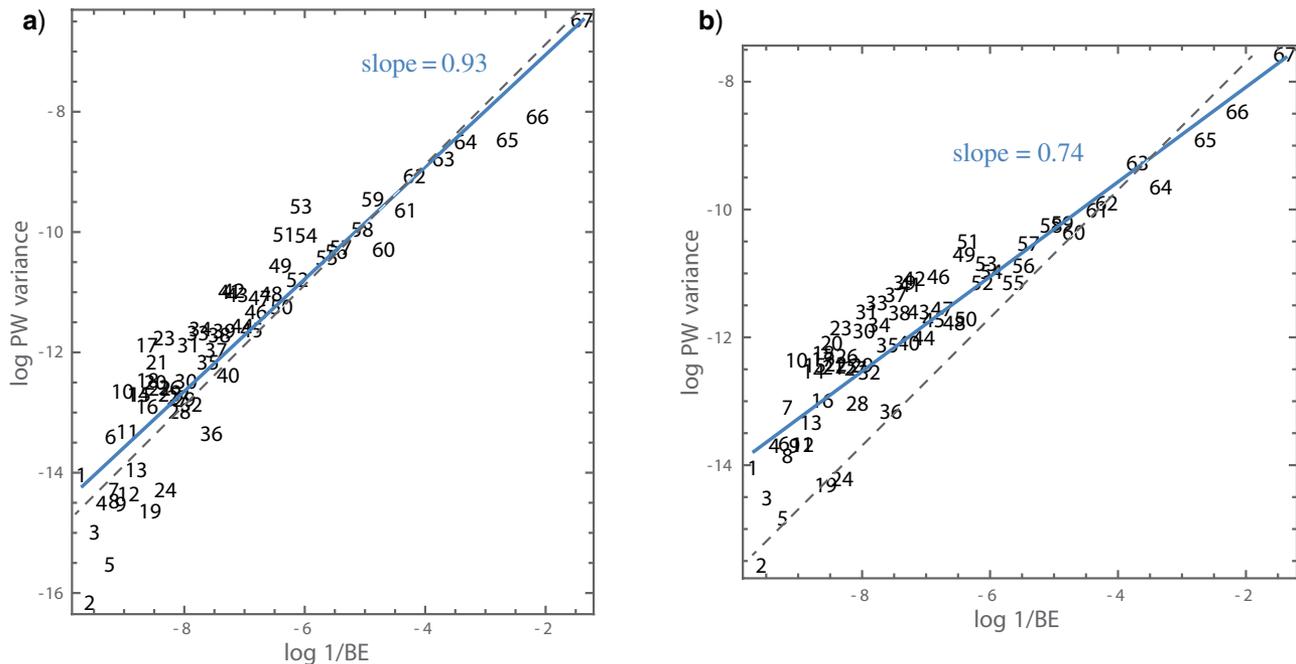


FIGURE 4. Plots of partial warp (PW) variance versus inverse bending energy (BE) for cranial shape variation between the 18 species (a) and within the species (b) on a log-log scale. The 70 landmarks give rise to 67 partial warps (indicated by the numbers in the plots), which are shape features of increasing spatial scale (measured by the inverse bending energy associated with every PW). The blue lines represent the linear regressions through these data. A slope of 1 (dashed line) in these log-log plots would indicate a linear scaling of nonaffine landmark variation and spatial scale, that is, *relative to scale*, the shape variation would be the same at every spatial scale (a self-similar shape distribution; see Materials and Methods). A slope clearly below 1, as for the within-species variation in panel b, indicates an excess of small-scale shape variation over large-scale shape variation. In both plots, the small-scale PWs with very small variance correspond to the positions of the semilandmarks along their curves, which were standardized by the semilandmark algorithm.

The PCs of the large-scale shape features (Fig. 6) were similar to those of the outline shape and did not separate the clades well. The first three PCs of small-scale shape mainly corresponded to the relative extensions of the nasal, frontal, ethmoid, and premaxillary bones. Taxonomic groups clustered relatively well along PCs 1 and 3 of small-scale shape (Fig. 6, bottom).

Phylogenetic Signal

For the full set of 70 landmarks, the matrix correlation of phylogenetic distance and Procrustes distance between species mean shapes was only 0.28 across all 18 species and 0.20 across the 16 papionins, indicating a weak phylogenetic signal of overall cranial shape. The matrix correlations were similarly low for the outline shape and for the large-scale shape features (Table 2). Residual shape, however, which captures the relative bone contributions, showed a matrix correlation of 0.60 ($P < 0.001$) across all species and of 0.28 ($P < 0.001$) across papionins. Small-scale shape features even showed a matrix correlation of 0.69 ($P < 0.001$) across all species and 0.38 ($P < 0.001$) across papionins (Table 2, Supplementary Fig. S5 available on Dryad).

These differences in phylogenetic signal were also reflected by the tree topologies obtained from outline and residual shape. Although the results differed slightly between the NJ and UPGMA algorithms, for

both methods the trees based on residual shape were more congruent with molecular genetic data than those based on outline shape (Fig. 7). However, the trees obtained from small-scale and large-scale shape were relatively similar and incongruent with the molecular genetic data (not shown).

Ecological Signal

In contrast to the phylogenetic signal, the association with ecological and behavioral variables was consistently stronger for the outline shape as compared with the residual shape (R^2 of 0.51 vs. 0.33 across all 18 species). This association was also stronger for large-scale shape features than for small-scale features (R^2 of 0.49 vs. 0.40; Table 3).

DISCUSSION

As predicted, phylogenetic signal differed considerably across the spatial scales of papionin cranial shape. Outline shape, which presumably captures most of the functionally relevant aspects of cranial morphology, showed substantially more variation between the species than within species. This may have resulted from adaptive divergence between the species and stabilizing selection within species. Indeed, outline

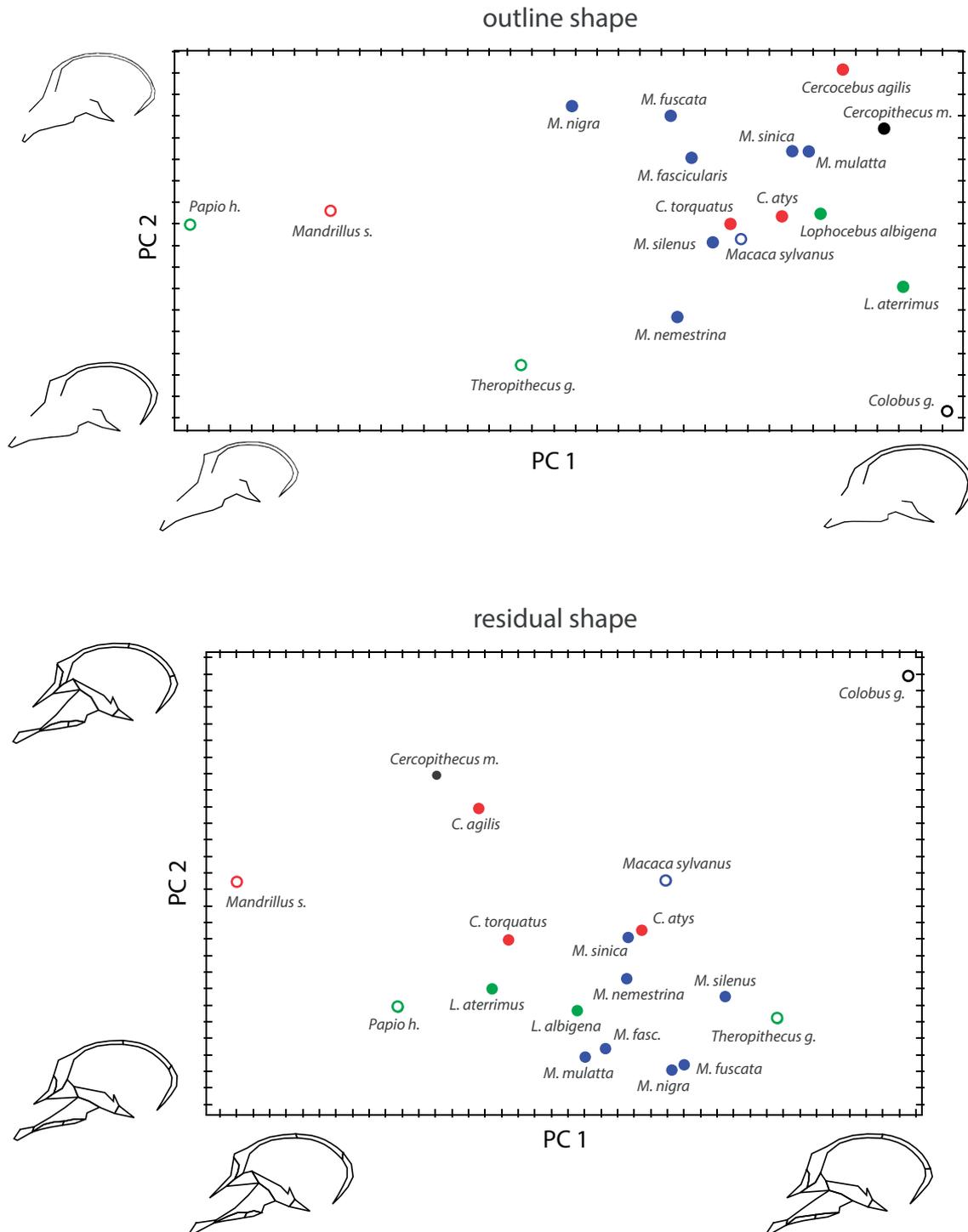


FIGURE 5. Principal components (PCs) of the species mean shapes for cranial outline shape (upper panel) and residual shape (lower panel). The shape features corresponding to the first two PCs are visualized by the reconstructed configurations along the axes; they correspond to the minimal and maximal scores along the respective axis. The species are colored according to the major phylogenetic lineages (cf. Fig. 1a). Note that these lineages cluster better for the residual shape than for the outline shape.

shape correlated only weakly with phylogenetic distance between the taxa but correlated strongly with ecological and behavioral variables. By contrast, residual shape, which captures the relative contributions of the different bones to overall shape, showed about the same amount

of variance between species and within species. Due to their limited functional relevance, these relative bone contributions do not seem to be strongly canalized in individual development, and they showed a relatively high phylogenetic signal across the 18 species.

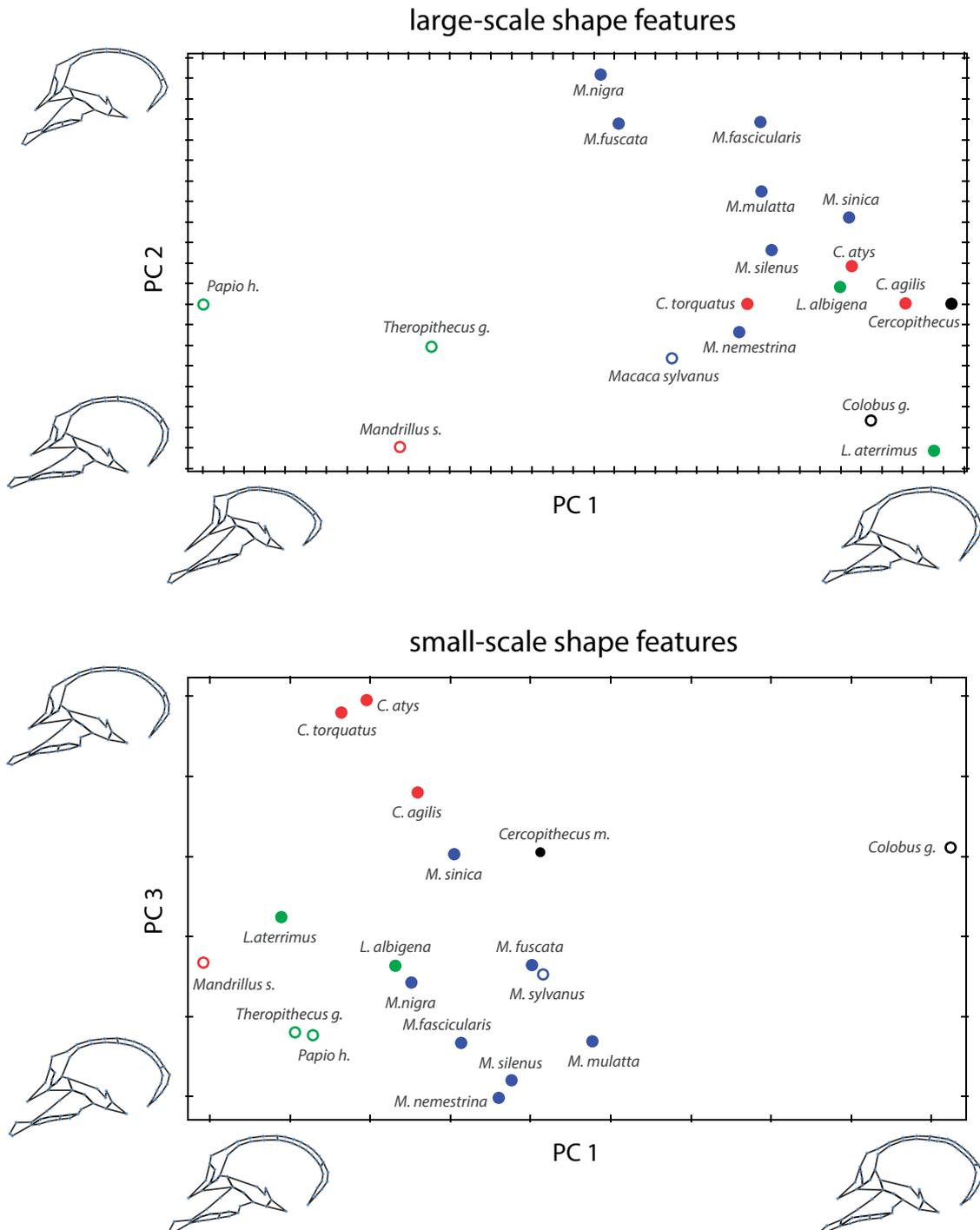


FIGURE 6. Principal components (PCs) of large-scale and small-scale shape features (upper and lower panel, respectively). The shape features corresponding to the PCs are visualized by the reconstructed configurations along the axes; they correspond to the minimal and maximal scores along the respective axis. Note that the major phylogenetic lineages (represented by the different colors) cluster better for small-scale shape than for the large-scale shape.

Very similar results were obtained from the decomposition of cranial shape variation into large-scale and small-scale components based on partial warps. The first 50 partial warps yielded the highest phylogenetic signal and a nice separation of phylogenetic lineages in the PCA. Large-scale shape

features (PWs 51–67) showed more evolutionary divergence than small-scale shape but did not separate the lineages well; they yielded a similar ordination of the species mean shapes and also a similar correlation with the ecological variables as the outline shape.

TABLE 2. Phylogenetic signal (correlation between pairwise morphological and phylogenetic distances) for the five different data sets, along with P -values estimated by Mantel's tests (uncorrected for multiple testing).

	All landmarks	Outline shape	Residual shape	Large-scale shape variation	Small-scale shape variation
All 18 species	$r=0.28$ ($P=0.01$)	$r=0.28$ ($P=0.01$)	$r=0.60$ ($P<0.001$)	$r=0.26$ ($P=0.04$)	$r=0.69$ ($P<0.001$)
16 papionins	$r=0.20$ ($P=0.01$)	$r=0.16$ ($P=0.04$)	$r=0.28$ ($P<0.001$)	$r=0.21$ ($P=0.01$)	$r=0.38$ ($P<0.001$)

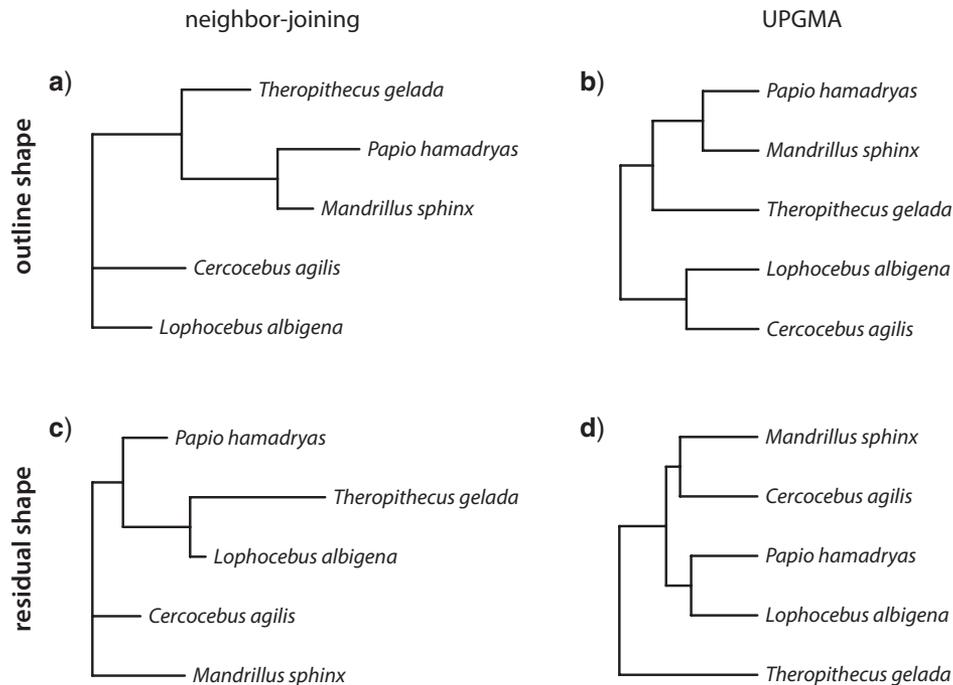


FIGURE 7. Trees illustrating the phylogenetic relationships recovered for the African papionin genera based on outline shape (a,b) and residual shape (c, d) using two different algorithms, neighbor-joining (a, c) and UPGMA (b, d). The relationships in a and b resemble those previously recovered based on craniodental morphology, whereas those in c and d are more consistent with molecular genetic data (compare Fig. 1a and Supplementary Fig. S2 available at Dryad).

TABLE 3. Ecological signal, estimated by multivariate R^2 for the five different data sets and the six ecological and behavioral variables.

	All landmarks	Outline shape	Residual shape	Large-scale shape variation	Small-scale shape variation
All 18 species	$R^2=0.50$	$R^2=0.51$	$R^2=0.33$	$R^2=0.49$	$R^2=0.40$
16 papionins	$R^2=0.48$	$R^2=0.45$	$R^2=0.38$	$R^2=0.48$	$R^2=0.41$

For the African papionin genera, which show the most pronounced homoplasies in cranial shape, we constructed phylogenetic trees based on the four different shape data sets. Whereas the tree topologies inferred from outline shape resembled those previously recovered based on craniodental morphology, the trees inferred from residual shape were relatively consistent with molecular genetic data. However, the trees computed from small-scale and large-scale shape features both resembled the overall morphological similarities (as outline shape did) and were incongruent with the molecular data.

Residual shape and especially small-scale shape showed weak associations with the ecological variables,

indicating that some functional aspects of cranial shape were still present in these data sets. Indeed, a PLS analysis between outline shape and residual shape revealed that muzzle length, which differs considerably across the 18 species, covaries with the organization of the anterior cranial base and also with the relative size of the premaxilla (Supplementary Fig. S6 available on Dryad). As a typical allometric pattern in primates and other placentals, muzzle length is correlated with body size and is of functional relevance (e.g., Singleton 2002; Frost et al. 2003; Mitteroecker et al. 2004; Cardini and Polly 2013; Grunstra et al. 2018; van der Geer et al. 2018). Regressing out cranial size (centroid size) from the small-scale shape data thus partially increased the consistency

of the recovered phylogenetic tree with the molecular data (Supplementary Fig. S7 available on Dryad).

We did detect a taxonomic-level effect: The phylogenetic signal in each landmark set was greatest when considering the total sample of 18 species. Omitting the two nonpapionin taxa from the analyses resulted in a reduction of the phylogenetic signal in each landmark set due to the reduced evolutionary time scale. The total sample captures 22 Myr of evolution, whereas the papionin sample represents only 14 Myr of evolution (Fig. 1a). The effects of genetic drift, and thus the strength of phylogenetic signal, are expected to play out more strongly over longer evolutionary time scales.

Among the five data sets, small-scale shape showed the highest matrix correlation for the 18-species sample as well as for the papionins only. However, the PCA ordinations and tree topologies indicate that small-scale shape better reflects phylogenetic relationships at higher taxonomic levels and residual shape at lower taxonomic levels. For instance, small-scale shape separated the major papionin clades well but performed poorly in recovering the phylogeny of the African papionins. Future research on taxonomically wider samples are necessary to verify this impression.

Our results of within-species shape variation are in agreement with those of Bookstein (2015) and Mitteroecker et al. (2020), who found that cranial shape in modern humans is characterized by considerably more small-scale variation than large-scale variation. Mitteroecker et al. suggested that mechanically induced growth of cranial bones compensates for the substantial individual variation in the relative dimensions of cranial bones and, thus, canalizes shape variation at larger scales. This may be a general pattern in primates, perhaps even in vertebrates. The differences in relative bone contributions *between* species, however, must have some genetic basis and are thus subject to genetic and developmental drift.

Our findings imply that morphometric studies of morphological adaptation and phylogenetic history profit from a decomposition of shape variation into outline and residual shape or into shape variation at different spatial scales. Phylogenetic signal is reduced for the functional aspects of cranial shape, as captured by bone outline shape and large-scale shape features, due to homoplasies resulting from similar adaptive processes in different species or due to stabilizing selection that prevents any (substantial) evolutionary change. Because of their functional importance, these large-scale aspects tend to be strongly canalized within a species and usually lead to a better group separation of individuals in a PCA ordination as compared with small-scale shape features (Supplementary Fig. S4 available on Dryad). By contrast, phylogenetic relationships *between* species are better represented by small-scale aspects of cranial shape. Hence, clustering individuals in order to identify populations or species, as is common in paleontology and paleoanthropology, is likely to be more effective for outline shape or large-scale features, whereas phylogenetic reconstruction based on species

mean shapes tends to be more successful for small-scale shape features or the relative bone contributions. Using larger samples and wider taxonomic scopes, future research will help to understand how relative bone contributions and other constructional aspects have evolved in different lineages alongside the large-scale, functional aspects of cranial shape, and how these associations affect phylogenetic inference or permit insights into past selective regimes.

Our approaches only apply to morphological regions comprised of multiple anatomical structures (e.g., bones, cartilages), such as crania, mandibles (except in mammals, where the mandible consists of only the dentary), and pelvises. We expect that the discrepancy in signal between the outline shape and relative bone contributions (or small-scale vs. large-scale features) increases with the number of different structures. The separations into outline/residual and small-scale/large-scale shapes are unaffected by sample size and covariance structure as they are only based on the mean shapes. Also the number of variables *per se* does not affect this decomposition, but the distribution of landmarks can do so. An anatomical region covered by many closely adjacent landmarks gives rise to more localized shape features than regions covered only by a few distant landmarks. The observed variances for these features, however, are not constrained by the distribution of landmarks. The approaches also extend to 3D landmark data. Semilandmarks can be slid along 3D curves and surfaces, and the outline landmarks can be warped to their mean configuration by the appropriate 3D TPS function (Bookstein 1991; Gunz et al. 2005). Likewise, partial warps and their bending energies can be computed for 3D landmarks.

SUPPLEMENTARY MATERIAL

Data, R code, and Supplementary Figures and Tables are available from the Dryad Digital Repository (<https://doi.org/10.5061/dryad.zkh189373>). The R package is also available from CRAN (<https://cran.r-project.org/package=prWarp>).

FUNDING

This work was supported by the Austrian Science Fund [FWF P29397].

REFERENCES

- Ackermann R.R., Cheverud J.M. 2004. Morphological integration in primate evolution. In: Pigliucci M., Preston K., editors. Phenotypic integration: studying the ecology and evolution of complex phenotypes. Oxford University Press.
- Bastir M., Rosas A. 2006. Correlated variation between the lateral basicranium and the face: a geometric morphometric study in different human groups. *Arch. Oral Biol.* 51:814–824.

- Blomberg S.P., Garland T. 2002. Tempo and mode in evolution: phylogenetic inertia, adaptation and comparative methods. *J. Evol. Biol.* 15:899–910.
- Bookstein F.L. 1989. Principal warps: thin plate splines and the decomposition of deformations. *IEEE Trans. Pattern Anal. Mach. Intell.*, 11:567–585.
- Bookstein F. 1991. *Morphometric tools for landmark data: geometry and biology*. Cambridge (UK); New York: Cambridge University Press.
- Bookstein F.L. 1997. Landmark methods for forms without landmarks: morphometrics of group differences in outline shape. *Med. Image Anal.* 1(3):225–243.
- Bookstein F.L. 2015. Integration, disintegration, and self-similarity: characterizing the scales of shape variation in landmark data. *Evol. Biol.* 42(4):395–426.
- Bookstein F.L. 2018. *A course in morphometrics for biologists: geometry and statistics for studies of organismal form*. Cambridge, UK: Cambridge University Press.
- Cardini A., Elton S. 2008. Does the skull carry a phylogenetic signal? Evolution and modularity in the guenons. *Biol. J. Linn. Soc.* 93:813–834.
- Cardini A., Polly P.D. 2013. Larger mammals have longer faces because of size-related constraints on skull form. *Nat. Commun.* 4:2458.
- Caumul R., Polly P.D. 2005. Phylogenetic and environmental components of morphological variation: skull, mandible and molar shape in marmots (*Marmota*, Rodentia). *Evolution* 59:2460–2472.
- Collard M., Wood B. 2000. How reliable are human phylogenetic hypotheses? *Proc. Natl. Acad. Sci. USA* 97(9):5003–5006.
- Collard M., Wood B. 2001. Homoplasy and the early hominid masticatory system: inferences from analyses of extant hominoids and papionins. *J. Hum. Evol.* 41(3):167–194.
- Collard M., O'Higgins P. 2002. Ontogeny and homoplasy in the papionin monkey face. *Evol. Dev.* 3(5):322–331.
- Delson E., Dean D. 1993. Are *Papio baringensis* R. Leakey, 1969, and *P. quadratiostris* Iwamoto, 1982, species of *Papio* or *Theropithecus*? In: Jablonski N.G., editor. *Theropithecus: the rise and fall of a primate genus*. Cambridge: Cambridge University Press. p. 126–156.
- Felsenstein J. 1982. Numerical methods for inferring evolutionary trees. *Q. Rev. Biol.* 57(4):379–404.
- Felsenstein J. 2003. *Inferring phylogenies*. Sunderland, MA: Sinauer.
- Frost S.R., Marcus L.F., Bookstein F.L., et al. 2003. Cranial allometry, phylogeography, and systematics of large-bodied papionins (primates: Cercopithecoidea) inferred from geometric morphometric analysis of landmark data. *Anat. Rec. A* 275A:1048–1072.
- Disotell T.R. 1994. Generic level relationships of the Papionini (Cercopithecoidea). *Am. J. Phys. Anthropol.* 94:47–57.
- Fleagle J.G., McGraw W.S. 1999. Skeletal and dental morphology supports diphyletic origin of baboons and mandrills. *Proc. Natl. Acad. Sci. USA* 96:1157–1161.
- Gilbert C.C., Frost S.R., Pugh K.D., Anderson M., Delson, E. 2018. Evolution of the modern baboon (*Papio hamadryas*): a reassessment of the African Plio-Pleistocene record. *J. Hum. Evol.* 122:38–69.
- Grunstra N.D.S., Mitteroecker P., Foley R.A. 2018. A multivariate ecogeographic analysis of macaque craniodental variation. *Am. J. Phys. Anthropol.* 166(2):386–400.
- Gunz P., Mitteroecker P., Bookstein F.L. 2005. Semilandmarks in three dimensions. In: Slice D., editor. *Modern morphometrics in physical anthropology*. Kluwer Press.
- Gunz P., Mitteroecker P. 2013. Semilandmarks: a method for quantifying curves and surfaces. *Hystrix* 24(1):103–109.
- Hallgrímsson B., Jamniczky H., Young N.M., Rolian C., Parson T.E., Boughner J.C., Marcuccio R.S. 2009. Deciphering the palimpsest: studying the relationship between morphological integration and phenotypic covariation. *Evol. Biol.* 36(4):355–376.
- Harris E., Disotell T.R. 1998. Nuclear gene trees and the phylogenetic relationships of the mangabeys (Primates: Papionini). *Mol. Biol. Evol.* 15:892–900.
- Harvati K., Weaver T.D. 2006a. Reliability of cranial morphology in reconstructing Neanderthal phylogeny. In: Harvati K. and Harrison T., editors. *Neanderthals revisited: new approaches and perspectives*. Dordrecht: Springer Netherlands. p. 239–254.
- Harvati K., Weaver T.D. 2006b. Human cranial anatomy and the differential preservation of population history and climate signatures. *Anat. Rec. A Discov. Mol. Cell. Evol. Biol.* 288(12):1225–1233.
- Jolly C.J. 1993. Species, subspecies, and baboon systematics. In: Kimbel W.H., Martin L.B., editors. *Species, species concepts, and primate evolution*. New York: Plenum. p. 67–107.
- Jolly C.J., Burrell A.S., Phillips-Conroy, J.E., Bergey C., Rogers J. 2011. Kinda baboons (*Papio kindae*) and grayfoot chacma baboons (*P. ursinus griseipes*) hybridize in the Kafue river valley, Zambia. *Am. J. Primatol.* 73:291e303.
- Kamilar J. 2006. Geographic variation in savanna baboon (*papio*) ecology and its taxonomic and evolutionary implications. In: Lehman S.M. and Fleagle J.G., editors. *Primate biogeography*. New York: Springer. p. 169–200.
- Kitching I.J., Forey P.L., Humphries C.J., Williams D.M. 1989. *Cladistics: theory and practice of parsimony analysis: the theory and practice of parsimony analysis*. Oxford: Oxford University Press.
- Klingenberg C.P. 2010. Evolution and development of shape: integrating quantitative approaches. *Nat. Rev. Genetics* 11(9):623–35.
- Klingenberg C.P. 2013. Cranial integration and modularity: insights into evolution and development from morphometric data. *Hystrix* 24(1):43–58.
- Lee M.S.Y., Palci A. 2015. Morphological phylogenetics in the genomic age. *Curr. Biol.* 25(19):R922–R929.
- Leigh S.R. 2007. Homoplasy and the evolution of ontogeny in papionin primates. *J. Hum. Evol.* 52(5):536–558.
- Lockwood C.A., Kimbel W.H., Lynch J.M. 2004. Morphometrics and hominoid phylogeny: support for a chimpanzee-human clade and differentiation among great ape subspecies. *Proc. Natl. Acad. Sci. USA* 101(13):4356–60.
- Lockwood C.A. 2007. Adaptation and functional integration in primate phylogenetics. *J. Hum. Evol.* 52(5):490–503.
- MacLeod N. 2001. Landmarks, localization, and the use of morphometrics in phylogenetic analysis. In: Adrain J.M., Edgecombe G.D., Lieberman B.S., editors. *Fossils, phylogeny, and form*. Boston: Springer.
- MacLeod N. 2002. Phylogenetic signals in morphometric data. In: MacLeod N., Forey P.L., editors. *Morphology, shape and phylogeny*. London: Taylor & Francis. p. 100–138.
- Madsen O., Scally M., Douady C.J., Kao D.J., DeBry R.W., Adkins R., Amrine H.M., Stanhope M.J., de Jong Wilfried W., Springer M.S. 2001. Parallel adaptive radiations in two major clades of placental mammals. *Nature* 409:610–614.
- Mantel N. 1967. The detection of disease clustering and a generalized regression approach. *Cancer Res.* 27:209–220.
- Mardia K.V., Bookstein F.L., Kent J.T., Meyer C.R. 2006. Intrinsic random fields and image deformations. *J. Math. Imaging Vis.* 26(1–2):59–71.
- Mittermeier R., Rylands A., Wilson D., editors. 2013. *Handbook of the mammals of the world*. Vol. 3. Barcelona: Primates. Lynx Edicions.
- van der Geer A., Lyras G., Mitteroecker P., MacPhee R. 2018. From jumbo to dumbo: cranial shape changes in elephants and hippos during phyletic dwarfing. *Evol. Biol.* 45(3):303–317.
- Mitteroecker P., Bookstein F.L. 2007. The conceptual and statistical relationship between modularity and morphological integration. *Syst. Biol.* 56(5):818–836.
- Mitteroecker P., Bookstein F.L. 2008. The evolutionary role of modularity and integration in the hominoid cranium. *Evolution* 62(4):943–958.
- Mitteroecker P., Gunz P., Bernhard M., Schaefer K., Bookstein F.L. 2004. Comparison of cranial ontogenetic trajectories among great apes and humans. *J. Hum. Evol.* 46:679–697.
- Mitteroecker P., Gunz P. 2009. Advances in geometric morphometrics. *Evol. Biol.* 36(2):235–247.
- Mitteroecker P. 2020. Morphometrics in evolutionary developmental biology. In: Nuno de la Rosa L, Müller G., editors. *Evolutionary developmental biology*. Cham: Springer.
- Mitteroecker P., Bartsch S., Erkingen C., Grunstra N.D.S., Le Maître A., Bookstein F.L. 2020. Morphometric variation at different

- spatial scales: coordination and compensation in the emergence of organismal form. *Syst. Biol.* 69(5):913–926.
- O’Higgins P., Johnson D.R. 1988. The quantitative description and comparison of biological forms. *CRC Crit. Rev. Anat. Sci.* 1(2):149–170.
- Poe S., Wiens J.J. 2000. Character selection and the methodology of morphological phylogenetics. In: Wiens J.J., editor. *Phylogenetic analysis of morphological data*. Washington, D.C.: Smithsonian Institution Press. p. 20–36.
- R Core Team. 2019. R: a language and environment for statistical computing, Version 3.6.2. Vienna (Austria): R Foundation for Statistical Computing. Available from: URL <https://cran.R-project.org/>.
- Rohlf F.J. 1988. On applications of geometric morphometrics to studies of ontogeny and phylogeny. *Syst. Biol.* 47(1):147–158.
- Rohlf F.J., Corti M. 2000. The use of two-block partial least-squares to study covariation in shape. *Syst. Biol.* 49:740–753.
- Rohlf F.J., Slice D.E. 1990. Extensions of the Procrustes method for the optimal superimposition of landmarks. *Syst. Zool.* 39:40–59.
- Roseman C.C., Willmore K.E., Rogers J., Hildebolt C., Sadler B.E., Richtsmeier J.T., Cheverud J.M. 2010. Genetic and environmental contributions to variation in baboon cranial morphology. *Am. J. Phys. Anthropol.* 143(1):1–12.
- Rowe N., Myers M., editors. *All the world’s primates*, Charlestown (RI): Primate Conservation Inc. Available from: www.alltheworldsprimates.org. Downloaded on Wed March 7 2018.
- Saitou N., Nei M. 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.* 4(4):406–425.
- Singleton M. 2002. Patterns of cranial shape variation in the Papionini (Primates: Cercopithecinae). *J. Hum. Evol.* 42(5):547–578.
- Smith H.F. 2009. Which cranial regions reflect molecular distances reliably in humans? Evidence from three-dimensional morphology. *Am. J. Hum. Biol.* 21(1):36–47.
- Sneath P., Sokal R.R. 1973. *Numerical taxonomy: the principles and practice of numerical classification*. San Francisco: W. H. Freeman.
- Strasser E., Delson E. 1987. Cladistic analysis of cercopithecoid relationships. *J. Hum. Evol.* 16:81–99.
- True J.R., Haag E.S. 2001. Developmental systems drift and flexibility in evolutionary trajectories. *Evol. Dev.* 3:109–119.
- von Cramon-Taubadel N. 2009. Congruence of individual cranial bone morphology and neutral molecular affinity patterns in modern humans. *Am. J. Phys. Anthropol.* 140(2):205–215.
- von Cramon-Taubadel N. 2011. The relative efficacy of functional and developmental cranial modules for reconstructing global human population history. *Am. J. Phys. Anthropol.* 146(1):83–93.
- Wilson D.E., Reeder, D.M., editors. 2005. *Mammal Species of the World: A Taxonomic and Geographic Reference*. 3rd ed. Baltimore: Johns Hopkins University Press.
- Yang Z., Rannala B. 2012. Molecular phylogenetics: principles and practice. *Nat. Rev. Genetics* 13:303–314.
- Zinner D., Groeneveld L.F., Keller C., et al. 2009. Mitochondrial phylogeography of baboons (*Papio* spp.) – indication for introgressive hybridization? *BMC Evol. Biol.* 9:83. doi: 10.1186/1471-2148-9-83.
- Zinner D., Buba U., Nash S., et al. 2011. Pan-African voyagers: the phylogeography of baboons. In: Sommer V., Ross C., editors. *Primates of gashaka*. New York, NY: Springer. p. 319–358.
- Zinner D., Fickenscher G.H., Roos C. et al. 2013a. Family Cercopithecidae (Old World monkeys). In: Mittermeier R.A., Rylands A.B., Wilson D.E., editors. *Handbook of the mammals of the world – Vol. 3*. Barcelona: Primates, Lynx Edicions. p. 550–754.
- Zinner D., Wertheimer J., Liedigk R., Groeneveld L.F., Roos C. 2013b. Baboon phylogeny as inferred from complete mitochondrial genomes. *Am. J. Phys. Anthropol.* 150:133–140.