**Tables**

**Table 1:** Generalised/Akaike Information criterion scores (GIC/AIC) and W-scores of relative probabilities of modes of evolution for models of shape, log- transformed centroid size, and shape evolution with log(centroid size) as predictor variable. The most likely modes are bolded.

**Table 2:** Phylogenetically adjusted generalised least squares analysis *(PGLS)* of module shape *vs* log-tranformed centroid sizes for joint (top) and separate (bottom) GPAs. SS, sum of squares; MS, Mean Squares; R2, R-squared value; F – F values; Z, Z scores (effect sizes from F values); *p*, p-value (probability of significant association at *p* < 0.05 based on 10,000 permutations).

**Table 3:** phylogenetically informed integration, expressed as r-PLS values, between the full shape (lower triangle) and residual shape (upper triangle) of modules. All r-PLS values were significant at *p* < 0.05 except for the integration between residuals of orbital and molar modules. For P-value tables, see Supplementary Table 1. Arrows in the upper triangle indicate whether the r-PLS values of residual integration analyses are higher (↑) or lower (↓) than the full-shape r-PLS values. Numbers in the diagonal are landmark numbers for each partition.

**Table 4:** Integration strength comparisons between pairs of modules, including effect sizes (Z) and significance of strength differences (*p*).

**Table 5:** Phylogenetically informed modularity, expressed as CR coefficients, between the full shape (lower triangle) and residual shape (upper triangle) of modules. Numbers on the diagonal are landmark numbers for each partition.

**Table 6:** Modularity tests using pairwise Mantel comparisons of PCA-based distance matrices of all modules, and Mantel r statistic. An r statistic of 1 indicates a strong correlation and 0 indicates no correlation. The upper triangle reports statistics for pairwise comparisons between cranial modules of the full dataset, the lower triangle reports r statistics for the residual dataset. The values in brackets are *p* values, adjusted by Bonferroni (1936) corrections for multiple comparisons.

**Figures**

**Figure 1:** Phylo-morphological pairwise distances plots. Each point is a pairwise comparison with border and center colors corresponding to the two species’ genera. The x-axis is shared but the y-axes of morphological distances are not equivalent as they rely on different shape datasets: a full shape and b shape residual.

**Figure 2**. Shape variation related to size and after size removal.a, Plot of PC1 and PC2 for the full shape dataset and b**,** plot of log centroid size versus the projected regression score with a gray regression line indicating the common evolutionary trajectory as per Marcy et al. (2020); c,‘allometry-free’ shape residual dataset with *Notomys* genus, *Conilurus penicillatus*, and *Mastacomys fuscus* highlighted; d, shape residual dataset without *Notomys*, which mainly just switches the sign of the PC scores*.*

**Figure 3:** **a**, shape differences between the shape fitted for mean centroid size of the smallest to the largest species in the sample; **b**, shape differences between the mean shapes of these two species; **c**, differences between the hypothetical shapes captured between PC1 extremes; **d**, differences between the hypothetical shapes captured between PC1 extremes based on allometry-free data; **e**, differences between the hypothetical shapes captured between PC2 extremes on allometry-free data; **f**, differences between hypothetical shapes between PC1 extremes after removing *Notomys.* Spheres show the mean position of landmarks for the column’s dataset, vectors show landmark displacement. Colors and lengths are calculated from relative proportions of the minimum/maximum vector lengths for each comparison,and are not equivalent across individual images. **g,** comparison between the smallest *versus* largest species in the sample (reflecting the variation seen in 3a-c); **f**, examples of species on extremes of residual PC1 (*H. chrysogaster*), and residual PC1/2 (*N. fuscus, M. fuscus)*, crania aligned so the foramen magnum is approximately vertical*.* Not to scale. *P. delicatulus* specimen is from Marcy *et al.* (2018).

**Figure 4:** Modularity and integration results averaged across the whole skull **a** our five-module framework(Goswami 2006). Results from the full shape (b,d) and shape residual (c,e) datasets. Black curves are the density distribution of coefficients from . 1,000 randomly drawn modules and the arrows point to the observed coefficients, which were all significant.