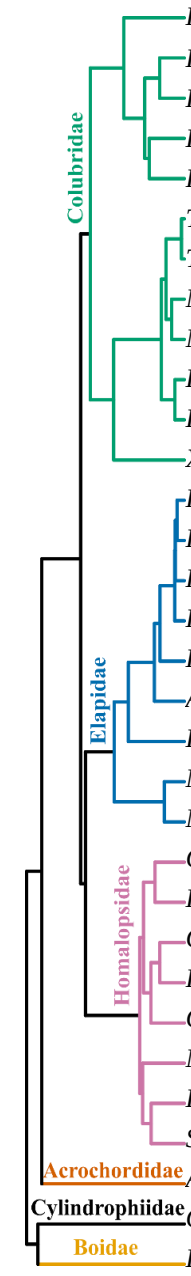
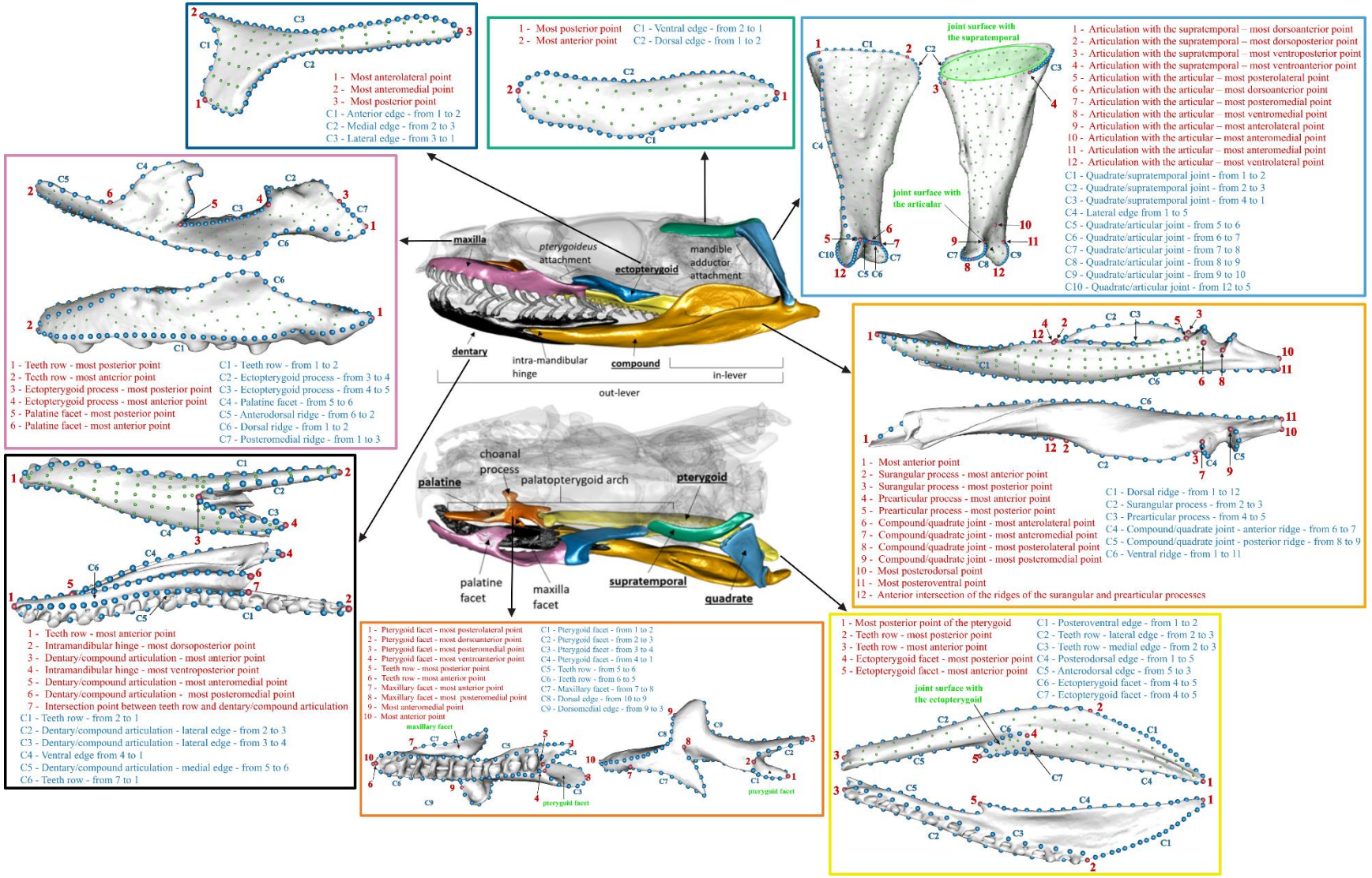
Supplementary Tables & Figures

Supplementary Table 1. Species and specimens used in this study. Museum codes: AMNH – American Museum of Natural History, FMNH – Field Museum of Natural History, CAS – California Academy of Sciences. Topology of the phylogeny used in this paper, from Pyron & Burbrink (2014).



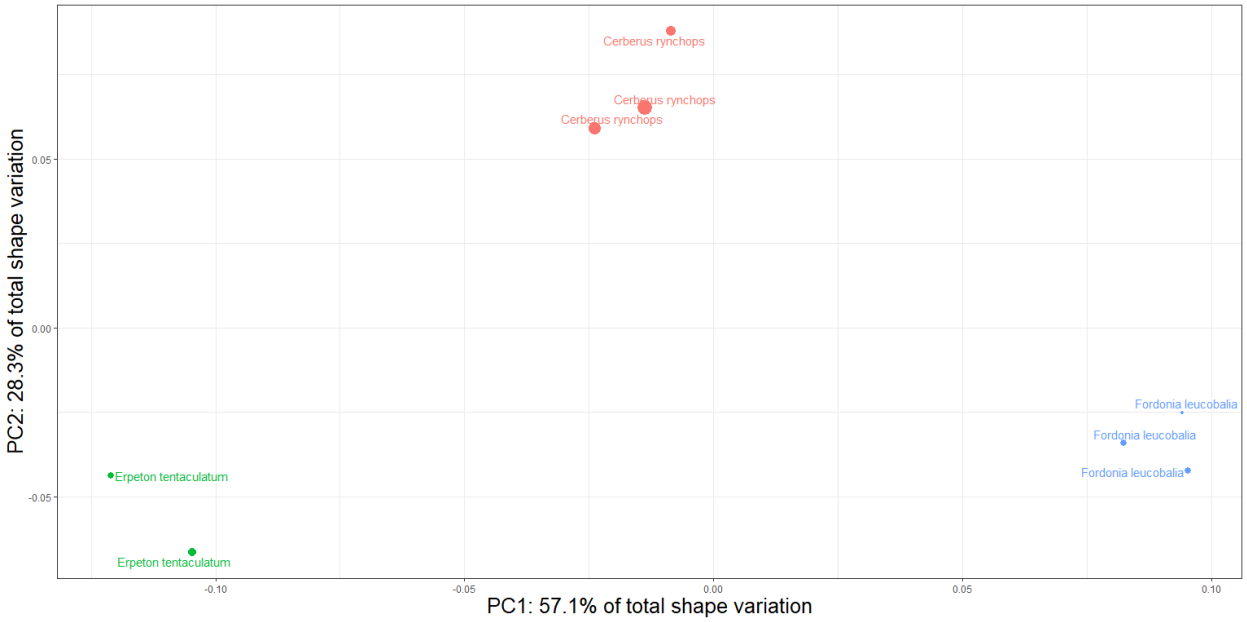
Phylogeny	*Scientific Name*	Specimen Numbers
	Hydrodynastes gigas	AMNH R60031, AMNH R93649
	Helicops carinicaudus	CAS87097
	Helicops angulatus	AMNH R130927, AMNH R18150
	Hydrops triangularis	AMNH R22449
	Pseudoeryx plicatilis	AMNH R55335
	Thamnophis atratus	AMNH R162404, AMNH R162405
	Thamnophis couchii	AMNH R09839, AMNH R09840
	Nerodia sipedon	AMNH R17086, AMNH R175012
	Nerodia rhombifer	AMNH R162256, AMNH R46754, AMNH R153729, AMNH R153727
	Liodytes rigida	AMNH R160212
	Liodytes alleni	AMNH R159304
	Xenochrophis piscator	AMNH R34085, AMNH R34086
	Hydrophis ornatus	AMNH R116013, AMNH R66588
	Hydrophis schistosus	AMNH R81854, CAS12296
	Hydrophis platurus	AMNH R19316, AMNH R19329
	Hydrophis melanocephalus	AMNH R03901, CAS22122
	Hydrelaps darwiniensis	AMNH R86169
	Aipysurus laevis	AMNH R161752, AMNH R5087
	Laticauda colubrina	AMNH R28997, AMNH R29000
	Micrurus lemniscatus	AMNH R119215
	Micrurus surinamensis	AMNH R152339
	Cerberus rynchops	AMNH R161961, FMNH199678, FMNH203432
	Homalopsis buccata	AMNH R92297, FMNH183771, FMNH229816
	Gerarda prevostiana	FMNH179104
	Fordonia leucobalia	FMNH229751, FMNH229758, FMNH229748
	Cantoria violacea	FMNH250116, CAS211909, CAS211910
	Myron richardsonii	AMNH R111792, AMNH R111793
	Erpeton tentaculatum	FMNH252609, FMNH252504
	Subsessor bocourti	FMNH191113, FMNH263528
	Acrochordus arafurae	CAS122071, CAS135488
	Cylindrophis ruffus	CAS16847
	Eunectes murinus	AMNH R56132, AMNH R46336



Supplemental Figure 1. The bones analyzed in this paper (middle two skulls) and their respective landmark configurations. Red indicates anatomical landmarks, blue curve semi-landmarks and green surface semi-landmarks, along with a list of description of each landmark (red) and curves (blue) and some important anatomical features (green).

Supplementary Table 2. The influence of allometry, phylogeny, and integration on snake skull bones associated with feeding. Results of allometry (PGLS), phylogenetic signal (all significant; $0.001 < p < 0.004$), and eigenvalue dispersion analyses. Significant p-values of allometry analyses are in bold. Higher eigenvalue dispersion value corresponds to higher degrees of overall integration within each structure. That is, more integrated structures will have more correlated shape variation consolidated in the first few principal components (eigenvalues).

	Allometry		**Phylogenetic Signal**		**Integration**
Structure	R^2	P	K	Effect Size	Eigenvalue dispersion
Dentary	0.06	0.15	0.45	2.99	0.61
Compound	0.08	0.051	0.58	3.91	0.65
Quadrate	**0.26**	**0.001**	0.51	3.05	0.55
Supratemporal	**0.08**	**0.02**	0.47	3.76	0.45
Pterygoid	0.07	0.09	0.51	3.04	0.47
Ectopterygoid	0.03	0.39	0.38	3.13	0.65
Palatine	**0.08**	**0.02**	0.59	5.97	0.48
Maxilla	**0.07**	**0.03**	0.57	4.63	0.47
Common Superimposition	**0.13**	**0.001**	0.48	4.20	0.45
Local Superimpositions	**0.07**	**0.02**	0.52	4.90	0.35



Supplemental Figure 2. Principal component morphospace from a common superimposition (GPA_{All}) of individual specimens of three homalopsid snakes: *Erpeton tentaculatum* (green), *Fordonia leucobalia* (blue), and *Cerberus rynchops* (orange). Note that individual specimens of each species group together much more closely than different species.

Supplemental Text 1. Comparison of angle fixation and superimposition methods for complex articulated structures and their impact on the modularity testing, and descriptions of two alternative methods.

A common approach to analyze an articulating structure in geometric morphometrics is to remove rotational variation between the mobile components by standardizing the angle between them (Adams 1999). In three dimensions, this is implemented using the *shapeRotator* R package (Vidal-Garcia et al. 2018). The product of this standardization is, for each specimen, a landmark configuration of the whole articulating structure with the mobile elements orientated into a common position, but with their relative sizes and articulations preserved. To produce this, *shapeRotator* requires the user to define three landmarks on each of a pair of articulating bones that define a cartesian plane on each specimen. For both bones, one of these landmarks must be on the joint between the bones. In the feeding systems of snakes, however, bones do not always articulate at the same positions between species, and in most cases, the feeding bones are too disparate to reliably identify 3 landmarks that can define the same plane in all specimens. Our local superimposition dataset described in the main text fixes the angle between each bone across all specimens, however the relative size and positional variation is erased in favor of within-bone shape variation. Here, we describe two alternative approaches that fix angles within complex articulating structures (over 2 articulations) while also preserving the relative size variation and articulations unique to each specimen, by manipulating the already-fixed-angle local superimposition dataset.

As demonstrated below, both these new superimposition procedures produce results highly similar with the results from the common and local superimpositions datasets. We do not interpret these results because they do not add novel information outside of the common and local superimposition datasets, and because the common and local superimposition datasets are complementary, representing results when prioritizing relative size and positional variation (common) and within-bone pure shape variation (local).

Description of Methods

The first superimposition method begins with the local superimposition dataset, with the individually superimposed component bones of the feeding system rotated, scaled, and translated to their corresponding bone from the mean configuration of a common superimposition, and the beginning steps of the second method create an analogous dataset. In these datasets, the bones are at common angles relative to one another. Each method preserves the relative size of each bone of each specimen: Method 1 rescales each bone to its biological relative size, and Method 2 performs non-scaling Procrustes alignments from the beginning, never losing biological relative size variation. Consequently, for example, *Cylindrophis ruffus* will have a quadrate smaller than *Acrochordus arafurae's*, since the quadrate in *C. ruffus* is proportionally smaller than in *A. arafurae*. At this point, relative size variation is accounted for, and the bones in each specimen are rotated to fixed angles, but the bones do not articulate properly as they would in the result of *shapeRotator*. The two approaches presented here differ in how they recreate the biologically plausible positioning (articulations) of the bones recovered in our original μCT scans, while keeping the angles fixed. In both methods, a generalized Procrustes superimposition (GPA) was performed on the modified landmark configurations.

Method 1: Iterative rearticulation

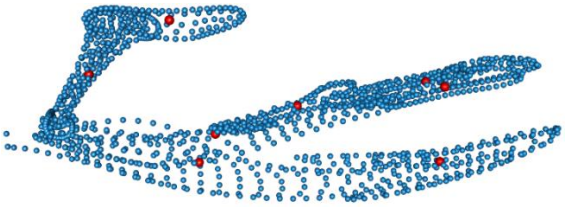
For each pair of articulating bones, choose a single landmark on either bone (*bone 1*) that is part of the articulation on every specimen (for each pair of bones in our dataset, there is at least one landmark per articulation that is common across all specimens). Next, using the landmark configurations in their μCT-scanned positions (before the common superimposition), for each articulation, calculate the Euclidean distance between the landmark you chose on *bone 1* and every landmark on the bone it is articulating with, *bone 2*. Take the landmark index of the landmark on *bone 2* closest to the landmark that was defined on *bone 1* (the smallest Euclidean distance), for each specimen. Now, for each specimen and each articulation, there are two landmarks (one for each pair of bones) that best define the joint between each pair of articulating bones. This accounts for variation in the position of joints within the bones, between the species.

Using these pairs of 'joint' landmarks, we can translate the landmark configuration of a single bone so that the 'joint' landmark on this bone has the same cartesian coordinate value as the other 'joint' landmark on the bone it articulates with. By doing so, we can 'rearticulate' the landmark configurations. First, rearticulate the compound with the dentary, and save this rearticulated landmark configuration. Next, rearticulate the quadrate with the rearticulated dentary and compound that was just saved. Rearticulate the supratemporal with the rearticulated dentary + compound + quadrate, and so on, successively rearticulating each landmark configuration until each pair of 'joint' landmarks, for each articulation, are translated to the same cartesian coordinate value. Repeat this process for each specimen. Now, the landmark configurations for each separate bone are fixed to common orientations (the orientations of the mean configuration of the common superimposition), variation in their relative sizes is preserved, and they articulate with one another at the same places that each individual specimen does. Finally, perform a Procrustes superimposition

of all the rearticulated specimens as if they were each a single rigid structure. This product is the same as what *shapeRotator* would produce if we were able to apply their technique to our dataset. It is important to note, however, that since this is a complex articulating structure (with numerous articulations, sometimes multiple articulations per bone) where the angle of articulation is functionally relevant, fixing the angles and then rearticulating won't produce 'natural' landmark configurations that exist in nature.

Method 2: Centroids superimposition

1) Calculate the centroid positions of each bone (red spheres) from the mean shape of the common superimposition.
2) Extract the centroid positions for each bone in each species (purple spheres) in their CT-scanned positions.



Mean shape

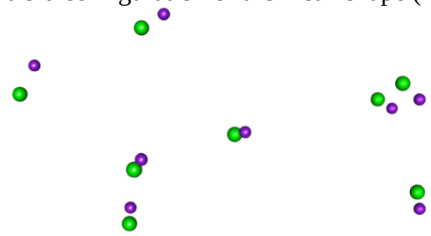
Species #4

3) Perform a non-scaling GPA (function 'ProcGPA', with scale=FALSE, *Morpho*) on the set of landmarks for each bone separately and force the alignment onto the landmarks of the corresponding bone in the mean shape of the common superimposition (e.g. blue spheres). Each bone is now aligned with its corresponding bone in the mean shape, the angles between the bones are homologous but, as the mean shape comes from a GPA, the scaling does not match each species' biological (real) size.

4) Translate the aligned bones into their original centroid configuration to preserve the respective position of the bones in the real skull (calculated in step 2).

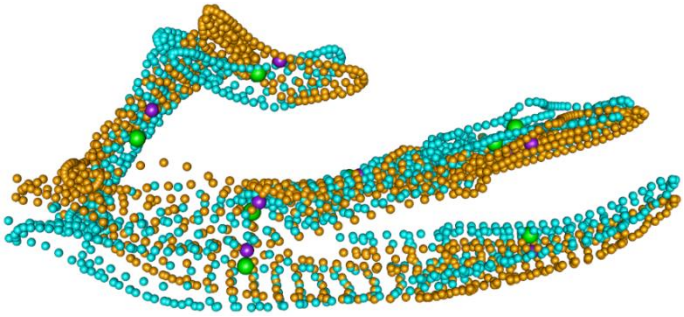
5) To isolate positional variation of bones between specimens, we performed a non-scaling GPA of the centroid configurations calculated in step 2 and forced the alignment onto the centroid configuration of the mean shape (in red in the figure above).

Purple spheres: original position of the centroids of species #4 (as in previous figure)
Green spheres: aligned centroids for species #4



6) Translate the superimposed landmark configurations for each bone (from step 3) into their corresponding aligned centroid configuration (calculated in step 5) to rescale the whole landmark configuration to its biological size without changing the fixed angles. This preserves variation in the respective position of the bones in the real skull (calculated in step 2).

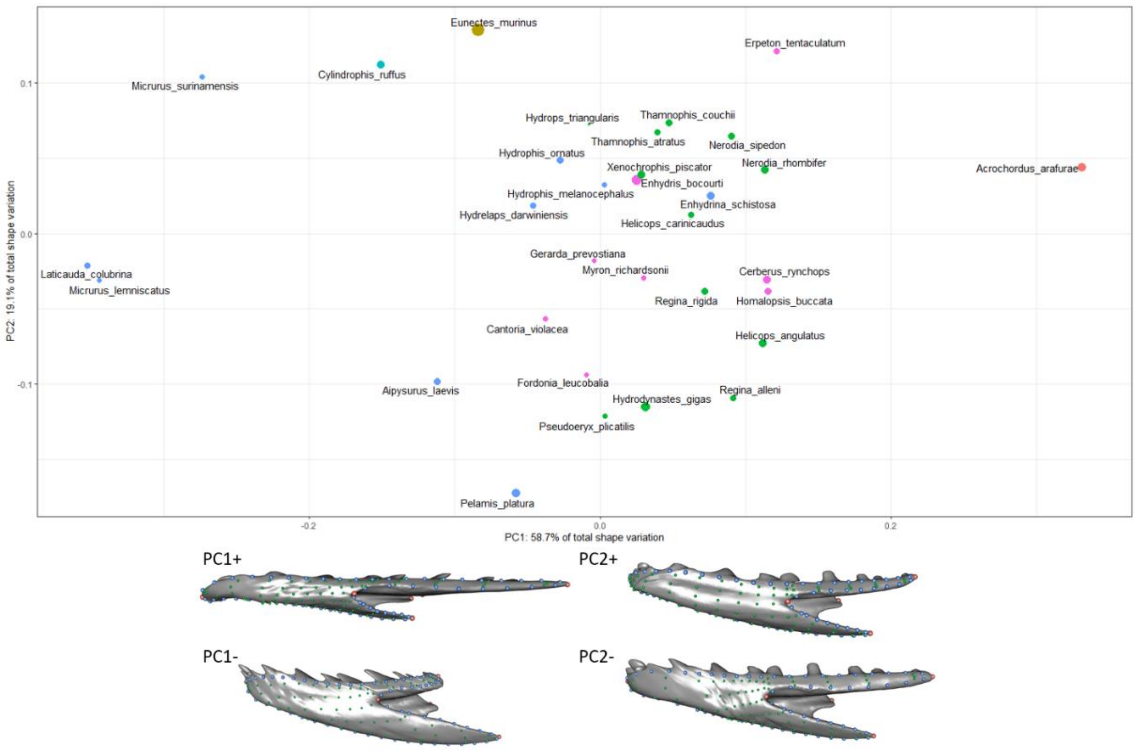
Original landmarks in orange, original centroid positions in purple, translated and rotated landmarks in cyan and aligned centroid positions in green



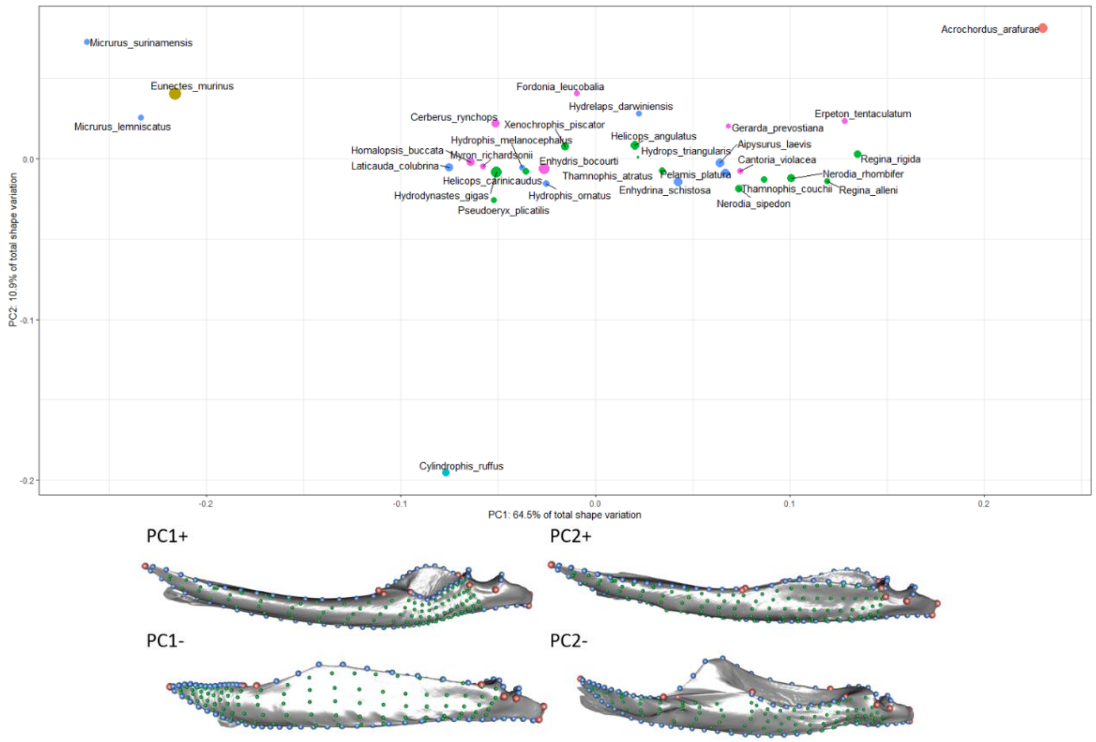
This method allows us to preserve both the respective size and position of the bones in the original skull while also fixing the angles. By superimposing the centroid positions of each bone, rather than just extracting the centroid positions of each bone for each species from the product of the common superimposition, we ensure that each bone has an equal weight during alignement so that we can isolate variation in the (biologically-relevant) interspecific positional variation of the bones. However, bones do not articulate correctly following this procedure as they would in the product of *shapeRotator*.

Supplementary Table 3. Explanations for each alternative hypothesis of modularity.

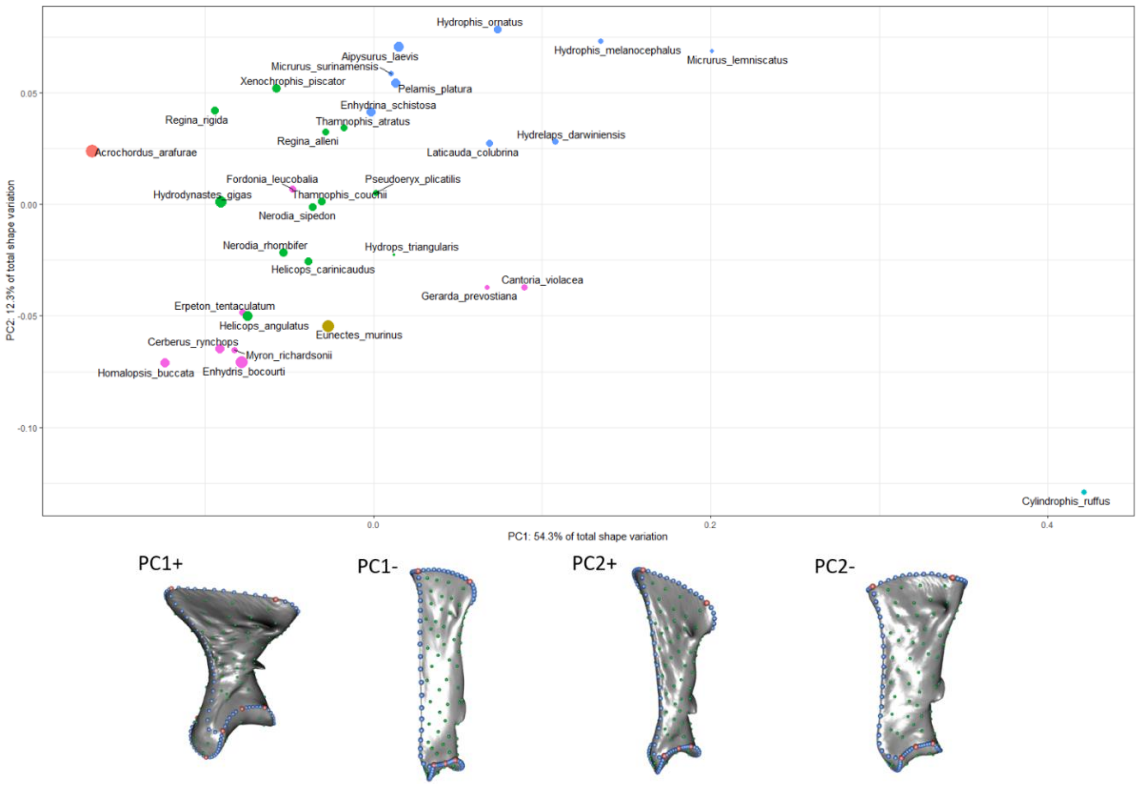
#	Hypothesis	Explanation
1	[dentary, compound] + [else]	Cranial elements and mandibular elements as separate evolutionary modules. Dorsal-ventral modularity.
2	[dentary, compound, quadrate, supratemporal] + [pterygoid, palatine, ectopterygoid, maxilla]	Suspensorium module consisting of the mandible, quadrate, and supratemporal and an upper jaw module. The mandible, quadrate, and supratemporal cumulatively contribute to gape size, which dictates the maximum size of prey (Cundall & Greene 2000). The elements of the suspensorium must act in concert to successfully strike at prey (Cundall & Greene 2000; Moon et al. 2019).
3	[maxilla, ectopterygoid] + [else]	The maxilla is not functionally integrated as strongly as the other bones (Cundall 1983), however the presence of fangs restructures its morphology and function (Kardong 1979; Vonk et al. 2008), and it may be integrated with the ectopterygoid because they articulate.
4	[dentary, compound, quadrate, maxilla] + [supratemporal, pterygoid, palatine, ectopterygoid]	Lateral vs. medial modules
5	[dentary, anterior compound] + [else]	Same as hypothesis #1 but with developmental modularity in compound. The posterior compound and quadrate are splanchnocranial.
6	[pterygoid, palatine] + [else]	Two-module hypothesis with the palatopterygoid arch as an individual module. The pterygoid and palatine articulate to make the palatopterygoid arch, which is necessary for the successfully functioning of the 'pterygoid walk' (Boltt and Ewer 1964).
7	[quadrate, supratemporal] + [else]	The quadrate's size and relative orientation varies greatly among species, and it is the only bone involved in prey detection as it connects to the stapes. The supratemporal is articulated with the quadrate and its relative position also varies greatly across the dataset.
8	[posterior compound, quadrate, supratemporal] + [else]	Developmental modularity; the quadrate and articular of the compound (posterior compound) are splanchnocranial, the other bones are dermatocranial.
9	[dentary, compound] + [quadrate, supratemporal] + [pterygoid, palatine, ectopterygoid, maxilla]	Same as hypothesis #1 and #2, but with the quadrate and supratemporal as an individual module. The relative orientation of the quadrate varies greatly among species, which has biomechanical implications (Scanferla 2016).
10	[dentary, compound, quadrate, supratemporal] + [maxilla] + [pterygoid, palatine, ectopterygoid]	Same as hypothesis #2, but with the mandible (suspensorium) as an individual module.
11	[dentary, compound, quadrate, supratemporal] + [maxilla, ectopterygoid] + [pterygoid, palatine]	Same as hypothesis #11, but with the ectopterygoid integrated with the maxilla, because they articulate.
12	[dentary, maxilla] + [compound, quadrate, supratemporal] + [pterygoid, palatine, ectopterygoid]	The dentary and maxilla are tooth bearing bones which are in direct physical contact with prey during prey capture (Cundall 1983; Cundall & Greene 2000; Moon et al. 2019). The compound and quadrate are connected via the mandibular adductor, which originates on the anterior quadrate and inserts on the posterolateral compound (Johnston 2014). The palatopterygoid arch articulates with the ectopterygoid.
13	[dentary, compound, maxilla, ectopterygoid] + [quadrate, supratemporal] + [pterygoid, palatine]	Same as hypothesis #11, but with the compound and ectopterygoid in the same module as the dentary and maxilla because they articulate with the dentary and maxilla, respectively.
14	[dentary, anterior compound] + [posterior compound, quadrate, supratemporal] + [pterygoid, palatine, ectopterygoid, maxilla]	Same as hypothesis #9 but incorporating developmental modularity in the compound.
15	[dentary, compound] + [quadrate, supratemporal] + [pterygoid, palatine, ectopterygoid] +[maxilla]	Same as hypothesis #9, but with the maxilla as an individual module. The maxilla has little direct role in any functional module other than prey capture in many species (Cundall 1983).
16	[dentary, compound] + [quadrate, supratemporal] + [pterygoid, palatine] + [maxilla, ectopterygoid]	Same as hypothesis #15, but with the ectopterygoid in the maxilla module.
17	[dentary, anterior compound] + [posterior compound, quadrate, supratemporal] + [pterygoid, palatine] + [ectopterygoid, maxilla]	Same as hypothesis #16, but modularity within compound.
18	Complete modularity except [pterygoid, palatine] + [dentary, compound]	The mandible and palatopterygoid arch are hypothesized to be two of the most integrated pairs of bones. This hypothesis considers modularity except for these two structures.
19	Complete modularity except [pterygoid, palatine]	Same as hypothesis #5, but with complete modularity outside of the palatopterygoid arch.
20	Complete modularity except [dentary, compound]	The mandibular elements are expected to be integrated because of their articulation and functional relationships; are mandibular, but not cranial elements integrated?
21	Complete modularity	The separate bones are unfused with one another, potentially allowing one bone to evolve without the necessary evolution of another. Does cranial kinesis promote modularity?



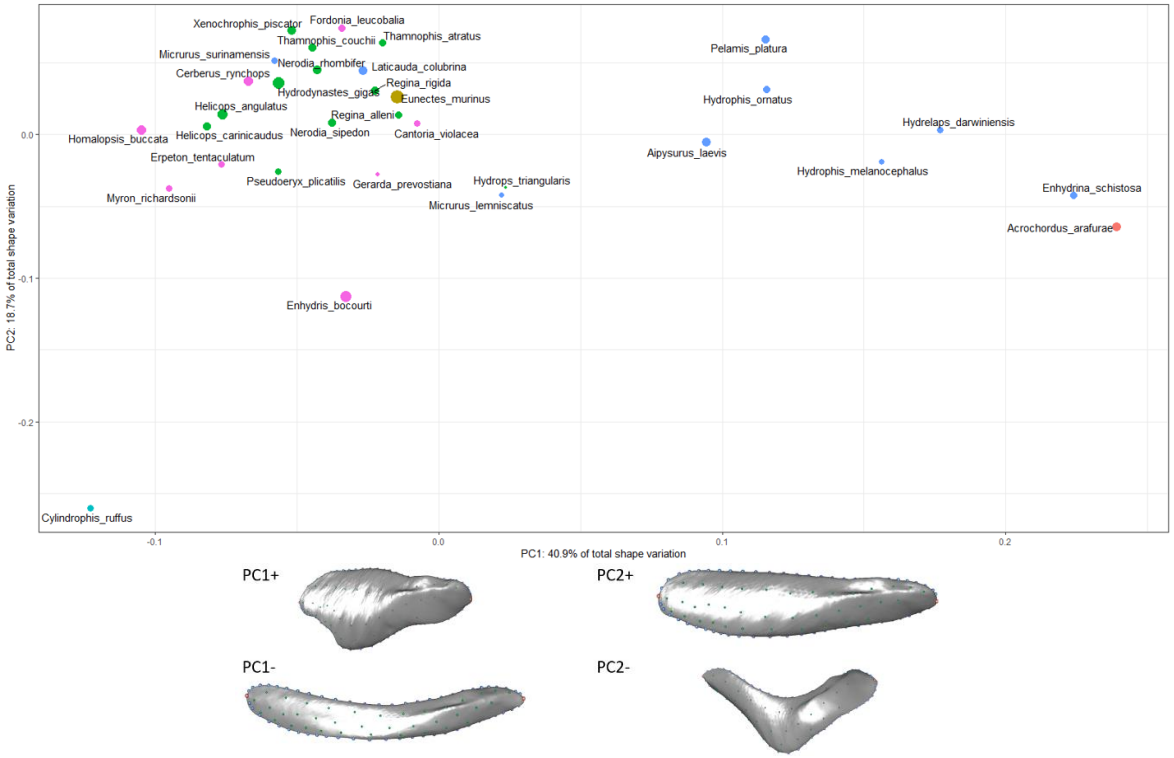
Supplementary Figure 3. Morphospace of the first two principal components of the dentary (above), and axes of shape variation along the first two principal components in lateral view (below). Size of each point (species) corresponds to centroid size of the dentary and color corresponds to taxonomic family as in Figure 1. This scheme is consistent for Supplementary figures 3-10.



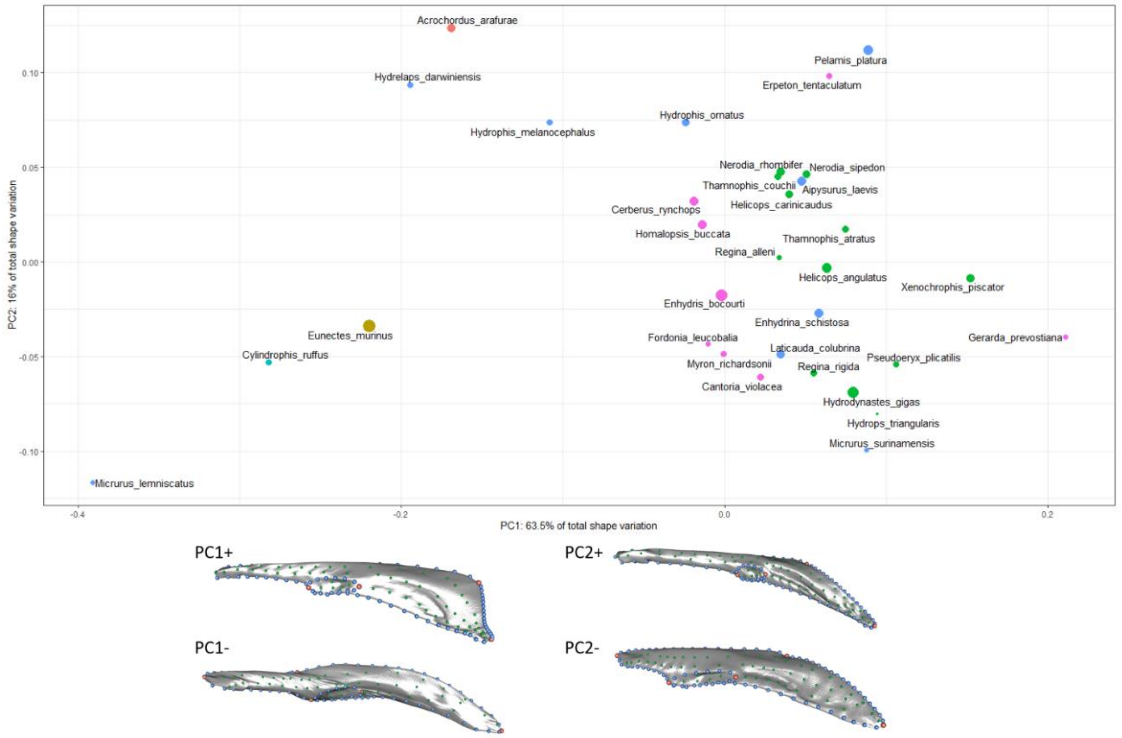
Supplementary Figure 4. Morphospace of the first two principal components of the compound (above), and axes of shape variation along the first two principal components in lateral view (below).



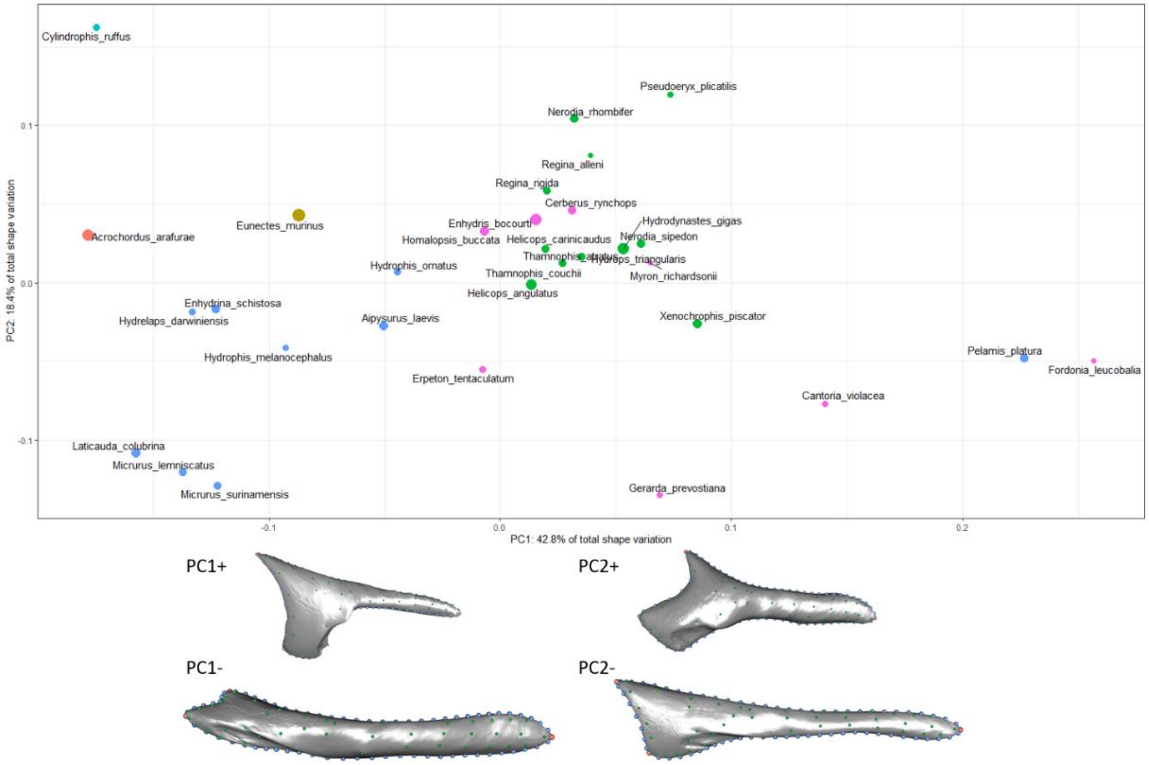
Supplementary Figure 5. Morphospace of the first two principal components of the quadrate (above), and axes of shape variation along the first two principal components in lateral view (below).



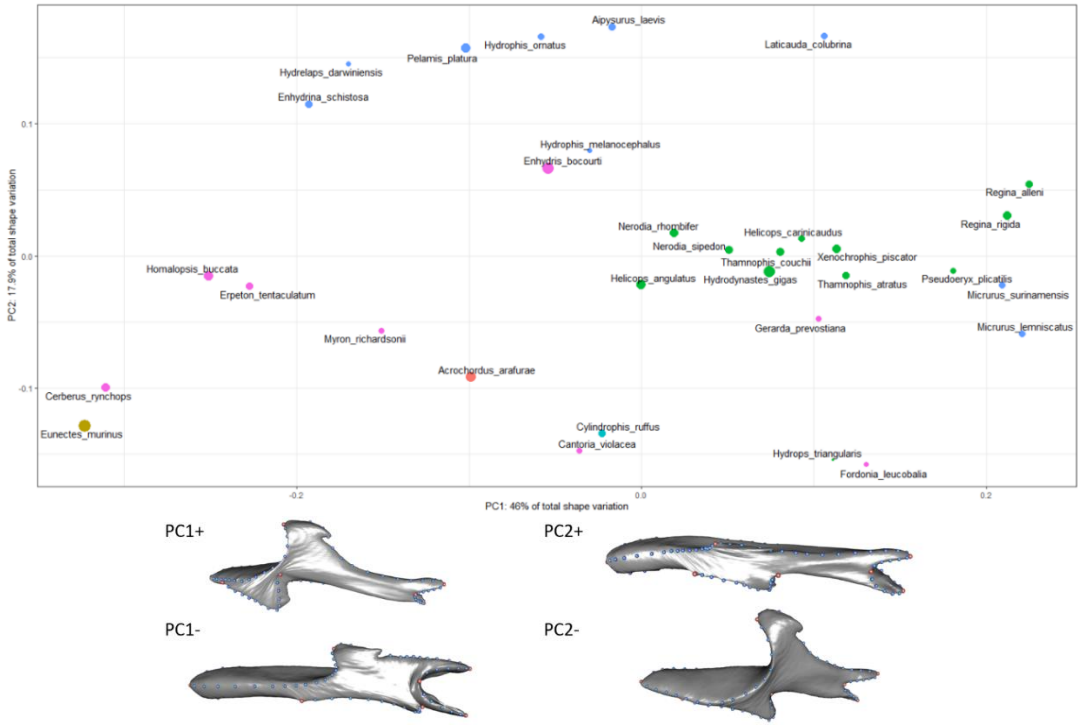
Supplementary Figure 6. Morphospace of the first two principal components of the supratemporal (above), and axes of shape variation along the first two principal components in lateral view (below).



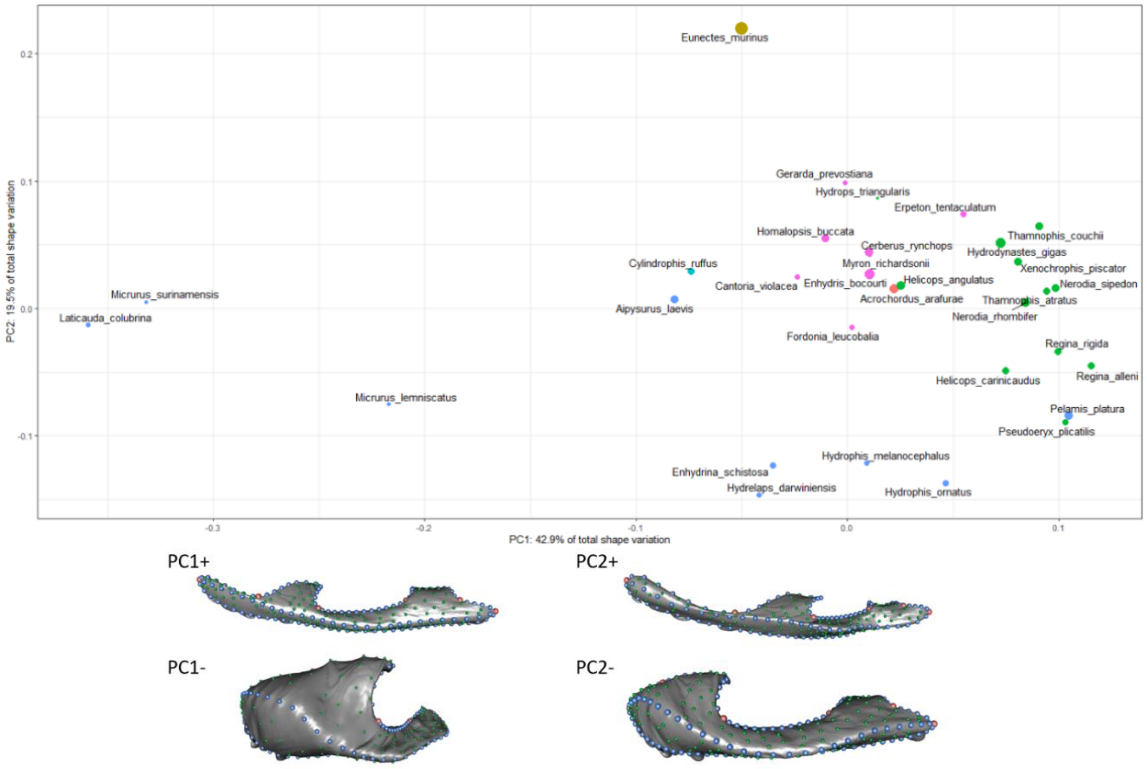
Supplementary Figure 7. Morphospace of the first two principal components of the pterygoid (above), and axes of shape variation along the first two principal components in dorsal view (below).



Supplementary Figure 8. Morphospace of the first two principal components of the ectopterygoid (above), and axes of shape variation along the first two principal components in dorsal view (below).



Supplementary Figure 9. Morphospace of the first two principal components of the palatine (above), and axes of shape variation along the first two principal components in dorsal view (below).



Supplementary Figure 10. Morphospace of the first two principal components of the maxilla (above), and axes of shape variation along the first two principal components in dorsal view (below).

Supplementary Table 4. The r-PLS values of PLS1 (above diagonal) and p-values (below diagonal) of each 2BPLS analysis. Significant values are in bold.

	Dentary	Compound	Quadrate	Supratemporal	Pterygoid	Ectopterygoid	Palatine	Maxilla
Dentary		**0.634**	**0.619**	**0.665**	0.476	**0.822**	0.651	**0.711**
Compound	**0.009**		0.589	0.465	0.407	**0.573**	0.561	0.578
Quadrate	**0.043**	0.053		**0.674**	**0.647**	**0.686**	0.651	**0.65**
Supratemporal	**0.029**	0.49	**0.013**		0.560	**0.73**	**0.692**	**0.686**
Pterygoid	0.158	0.312	**0.022**	0.111		0.453	**0.74**	0.606
Ectopterygoid	**0.001**	**0.043**	**0.028**	**0.012**	0.197		0.611	**0.806**
Palatine	0.051	0.123	0.051	**0.02**	**0.003**	0.105		**0.747**
Maxilla	**0.009**	0.115	**0.045**	**0.041**	0.087	**0.001**	**0.003**	

Supplementary Table 5. The z-PLS values of PLS1 (above diagonal) and p-values (below diagonal) of each 2BPLS analysis. Significant values are in bold.

	Dentary	Compound	Quadrate	Supratemporal	Pterygoid	Ectopterygoid	Palatine	Maxilla
Dentary		**2.905**	**1.910**	**1.980**	1.035	**4.058**	1.818	**2.435**
Compound	**0.009**		1.827	0	0.389	**2.073**	1.192	1.262
Quadrate	**0.043**	0.053		**2.362**	**2.342**	**2.348**	1.852	**1.874**
Supratemporal	**0.029**	0.49	**0.013**		1.255	**2.512**	**2.138**	**1.766**
Pterygoid	0.158	0.312	**0.022**	0.111		0.8233	**3.040**	1.452
Ectopterygoid	**0.001**	**0.043**	**0.028**	**0.012**	0.197		1.4378	**3.272**
Palatine	0.051	0.123	0.051	**0.02**	**0.003**	0.105		**2.944**
Maxilla	**0.009**	0.115	**0.045**	**0.041**	0.087	**0.001**	**0.003**	

Supplementary Table 6: a. Alternative hypotheses of modularity and their support, measured as phylogenetic-corrected Z_{CR}, for both common and local superimposition procedures. The lower the Z_{CR} and CR values, the stronger the modular signal. 'Else' refers to all the other bones not yet mentioned as part of their own module. Results from CR analyses for each alternative superimposition method, as well as the average Z_{CR} value for each hypothesis. Lower effect sizes are green, higher effect sizes (less support) are colored in red. b. Summary of the pros and cons of the superimposition methods along with their most supported modularity hypotheses.

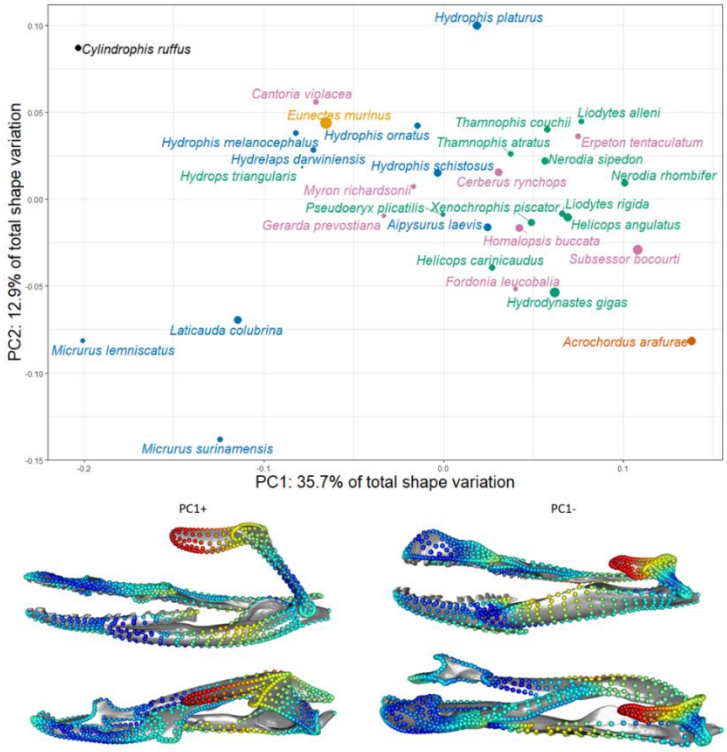
Table 6a.

		Common Superimposition		Local Superimpositions		Method 1 Iterative Rearticulations		Method 2 Superimposed Centroids		Average
Hypothesis	Number of Modules	CR	Zcr	CR	Zcr	CR	Zcr	CR	Zcr	Zcr
H1	2	0.63	-31.18	0.65	-27.02	0.68	-30.96	0.6	-31.16	-30.08
H2	2	0.76	-31.03	0.77	-26.08	0.7	-31.19	0.71	-31.07	-29.84
H3	2	0.85	-29.13	0.85	-18.27	0.79	-30.17	0.85	-28.62	-26.55
H4	2	0.82	-30.54	0.68	-28.14	0.74	-31	0.77	-30.78	-30.11
H5	2	0.74	-30.11	0.74	-22.45	0.8	-29.05	0.71	-30.18	-27.95
H6	2	0.90	-25.91	0.83	-17.53	0.84	-28.33	0.88	-26.09	-24.46
H7	2	0.98	-12.00	0.98	-3.36	0.99	-7.22	0.97	-13.26	-8.96
H8	2	0.9	-28.12	0.98	-3.57	0.9	-27.77	0.89	-28.2	-21.92
H9	3	0.76	-30.86	0.74	-25.52	0.76	-30.72	0.73	-30.86	-29.49
H10	3	0.74	-30.68	0.58	-28.12	0.67	-30.95	0.7	-30.73	-30.12
H11	3	0.81	-30.15	0.81	-21.25	0.76	-30.48	0.78	-30.11	-28.00
H12	3	0.81	-30.53	0.73	-26.69	0.77	-30.74	0.79	-30.5	-29.62
H13	3	0.92	-25.26	0.86	-18.04	0.89	-27.27	0.91	-25.44	-24.00
H14	3	0.78	-30.47	0.77	-24.24	0.76	-30.52	0.74	-30.58	-28.95
H15	4	0.72	-30.93	0.62	-28.13	0.7	-30.97	0.68	-30.95	-30.24
H16	4	0.76	-30.75	0.76	-24.79	0.76	-30.7	0.74	-30.72	-29.24
H17	4	0.77	-30.61	0.76	-24.5	0.74	-30.69	0.74	-30.64	-29.11
H18	6	0.75	-29.78	0.68	-23.66	0.72	-29.96	0.73	-29.71	-28.28
H19	7	0.71	-30.37	0.65	-24.35	0.69	-30.39	0.69	-30.33	-28.86
H20	7	0.77	-29.42	0.72	-22.47	0.74	-29.67	0.75	-29.3	-27.72
H21	8	0.73	-30.21	0.67	-24.72	0.71	-30.27	0.7	-30.15	-28.84

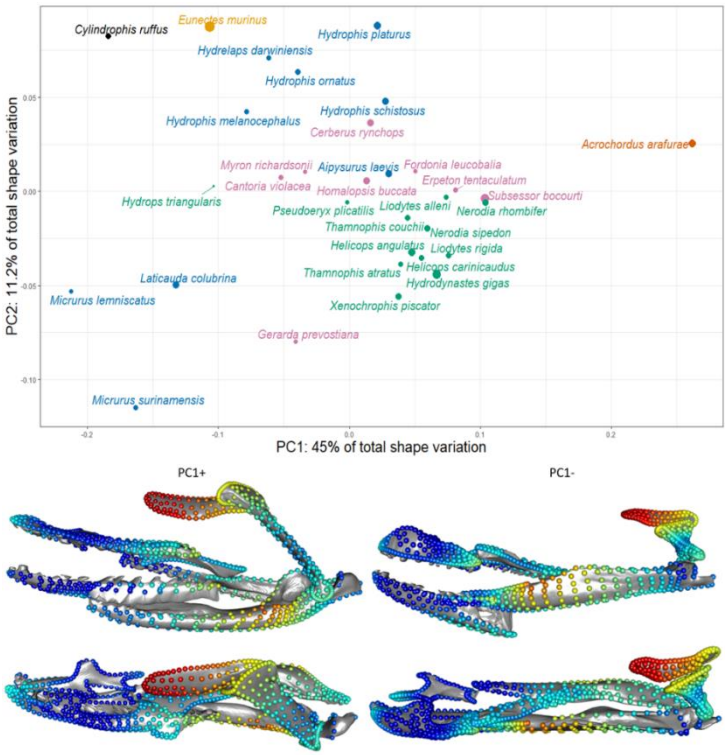
Table 6b.

		Common Superimposition	Local Superimposition	Method 1 Iterative Rearticulations	Method 2 Superimposed Centroids
Fixed angles		**no**	**yes**	**yes**	**yes**
Biological size		**yes**	**no**	**yes**	**yes**
Biological position		**yes**	**no**	**no**	**yes**
Articulation		**yes**	**no**	**yes**	**no**
Intra-bone variation		**no**	**yes**	**no**	**no**
Most supported hypotheses	1st	1	4	2	1
	2nd	2	15	4	2
	3rd	15	10	15	15

a. Iterative Rearticulations

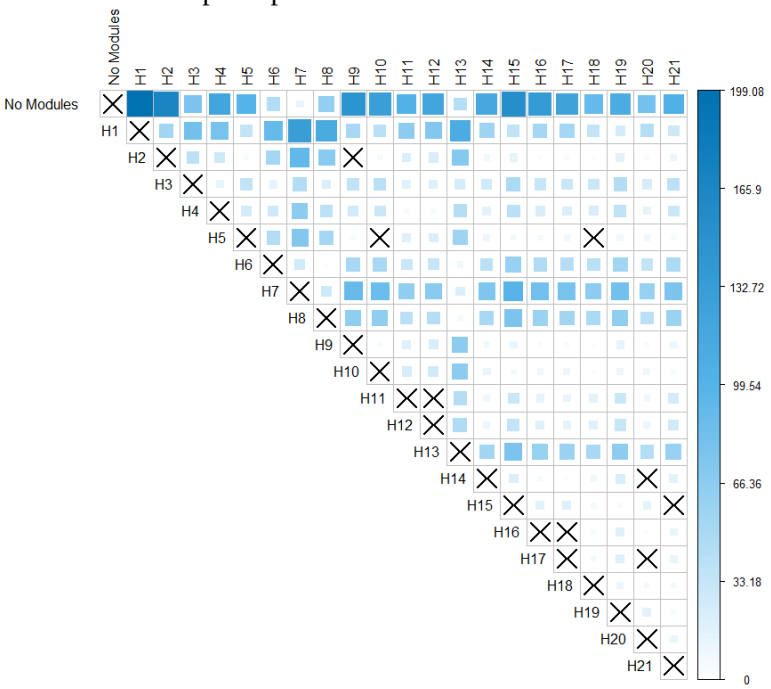


b. Superimposed Centroids

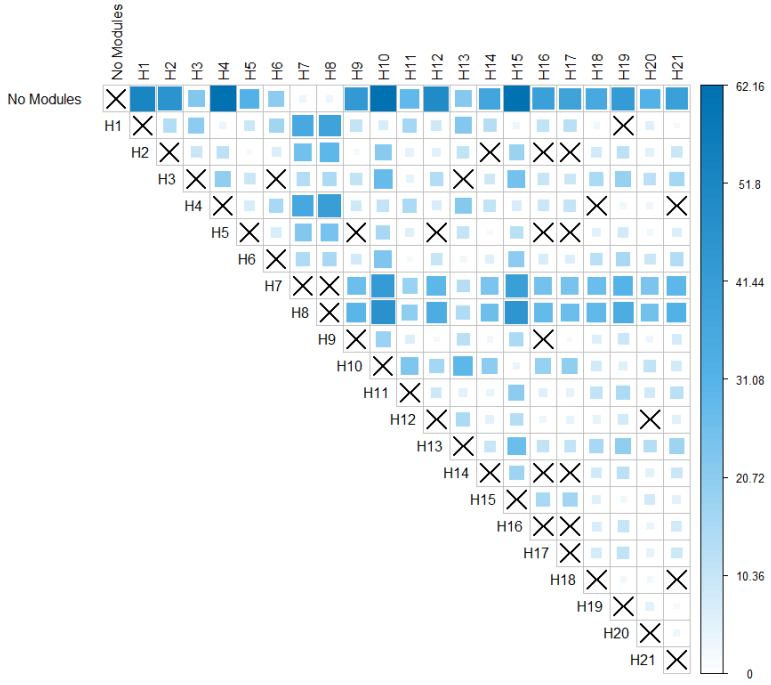


Supplementary Figure 11: (Above) Principal component analysis morphospace of landmark configurations generated by the Iterative Rearticulations (Method 1, a.) and Superimposed Centroids (Method 2, b.). Colors correspond to taxonomic family, as in Figure 1. (Below) Shape variation along PC1 in lateral (top) and dorsal (bottom) views. The color of each dot corresponds to its variation along PC1.

a. Common superimposition



b. Local superimposition



Supplementary Figure 12. Pairwise effect sizes of alternative hypotheses of modularity of the common superimposition (a.) and local superimposition (b.) datasets. Darker blue cells correspond to higher effect sizes and larger differences between strength of modular signal between alternative hypotheses. Insignificant (i.e., not significantly different support) values are marked by 'X's. 'No Modules' refers to the null hypothesis of complete integration.