doi:10.1111/evo.1398

Evolutionary integration of the frog cranium

Carla Bardua,[1,2,3] Anne-Claire Fabre,[4] Margot Bon,[4] Kalpana Das,[5] Edward L. Stanley,[6] David C. Blackburn,[7] and Anjali Goswami[4]

[1]*Department of Genetics, Evolution, and Environment, University College London, London , WC1E 6BT, United Kingdom*

[2]*Department of Life Sciences, Natural History Museum, London , SW7 5BD, United Kingdom*

[3]*E-mail: carla.bardua.15@ucl.ac.uk*

[4]*Department of Life Sciences, Natural History Museum, London , SW7 5BD, United Kingdom*

[5]*Museum für Naturkunde, Leibniz-Institut für Evolutions- und Biodiversitätsforschung, Berlin 10115, Germany*

[6]*Department of Herpetology, Florida Museum of Natural History, University of Florida, Gainesville, Florida 32610*

[7]*Department of Natural History, Florida Museum of Natural History, University of Florida, Gainesville, Florida 32611*

Received September 10, 2019

Accepted April 17, 2020

Evolutionary integration (covariation) of traits has long fascinated biologists because of its potential to elucidate factors that have shaped morphological evolution. Studies of tetrapod crania have identified patterns of evolutionary integration that reflect functional or developmental interactions among traits, but no studies to date have sampled widely across the species-rich lissamphibian order Anura (frogs). Frogs exhibit a vast range of cranial morphologies, life history strategies, and ecologies. Here, using high-density morphometrics we capture cranial morphology for 172 anuran species, sampling every extant family. We quantify the pattern of evolutionary modularity in the frog skull and compare patterns in taxa with different life history modes. Evolutionary changes across the anuran cranium are highly modular, with a well-integrated "suspensorium" involved in feeding. This pattern is strikingly similar to that identified for caecilian and salamander crania, suggesting replication of patterns of evolutionary integration across Lissamphibia. Surprisingly, possession of a feeding larval stage has no notable influence on cranial integration across frogs. However, late-ossifying bones exhibit higher integration than early-ossifying bones. Finally, anuran cranial modules show diverse morphological disparities, supporting the hypothesis that modular variation allows mosaic evolution of the cranium, but we find no consistent relationship between degree of within-module integration and disparity.

KEY WORDS: **Frogs, cranial morphology, integration, development, morphometrics, disparity.**

Trait integration is an inherent property of biological systems and manifests at many scales of analysis, from populations to large clades. Correlations among traits, and their organization into highly integrated, semi-autonomous modules, can result from genetic, developmental, or functional interactions, and it is thought that patterns of genetic and developmental integration may evolve adaptively to promote functional integration (Riedl 1978; Cheverud 1984; Wagner and Altenberg 1996). These trait associations may be replicated in macroevolutionary patterns observed across larger clades, with traits evolving in either a coordinated or a modular fashion. The modular organization of phenotypic evolution is central to many fundamental concepts in evolutionary biology, including mosaicism (De Beer 1954), which have been suggested to promote the diversification of form. Several recent studies have explored this hypothesized relationship by quantifying phenotypic (or variational) and evolutionary integration and modularity and determining how they relate to disparity and rates of evolution in various species or clades. These studies provide extensive evidence that the presence of evolutionary modules is associated with higher disparity or evolutionary rate (Goswami and Polly 2010; Claverie and Patek 2013; Randau and Goswami 2017; Felice and Goswami 2018; Larouche et al. 2018; Dellinger et al. 2019; Watanabe et al. 2019; Bardua et al. 2019b). Many studies have also assessed whether the magnitude of evolutionary integration among traits (across an entire phenotype or within a module) correlates with higher or lower disparity

or evolutionary rate. For this question, the answer is less clear than for the one above. Some studies find that high evolutionary integration among traits is correlated with low disparity and slow evolutionary rate (Claverie and Patek 2013; Felice and Goswami 2018; Martín-Serra et al. 2019). Conversely, others have found that highly integrated traits show greater disparity than less integrated ones. In addition, some have found that there is no relationship between evolutionary modularity and disparity or rate of evolution at this scale (Goswami et al. 2014; Watanabe et al. 2019; Bardua et al. 2019b). However, our understanding of the factors shaping the evolutionary integration among traits and its relevance for morphological evolution is incomplete, with most studies focused on a few clades, such as mammals and birds, both of which have relatively little developmental diversity. Here, we have expanded the study of evolutionary modularity to one of the most taxonomically, developmentally, ecologically, and morphologically diverse clades of tetrapods: the lissamphibian order Anura (frogs).

Lissamphibians constitute over 25% of extant tetrapod species (totalling 8,160 species; AmphibiaWeb 2020), and yet patterns of trait integration at any scale across this diverse clade remain largely unexplored compared to amniotes. Previous studies of both phenotypic and evolutionary integration in lissamphibian crania have found limited support for modular structure, with studies recovering either few modules or no modular structure at all (Ivanović and Kalezić 2010; Sherratt 2011; Ivanović and Arntzen 2014; Simon and Marroig 2017; Vidal-García and Keogh 2017). With the application of high-dimensional geometric morphometric data, strong modular structure has been identified at the intraspecific (phenotypic) and evolutionary levels across caecilian (Marshall et al. 2019; Bardua et al. 2019b) and salamander (Bon et al. 2020; Fabre et al. 2020) crania, in analyses exploring much wider ranges of models. However, the most diverse lissamphibians, frogs, have not yet been incorporated into these studies. The few studies of cranial modularity in frogs to date vary widely in taxonomic and morphological coverage as well as data type, hindering our understanding of anuran cranial modularity and preventing direct comparison of modular structures across Lissamphibia (e.g., Simon and Marroig 2017; Vidal-García and Keogh 2017). Beyond their taxonomic diversity, anurans provide a unique opportunity to investigate patterns of cranial integration across lineages varying markedly in life history, including the repeated loss of a free-living and feeding larval stage. The presence of such a larval stage has been hypothesized to drive the decoupling of genetic and developmental traits with functional traits, as the distinct, divergent selection pressures associated with larval and adult ecological niches may drive low genetic correlations between these two life history stages ("adaptive decoupling hypothesis"; Ebenman 1992; Moran 1994). Indeed, tadpoles and frogs can evolve independently, responding to antagonistic selection pressures (Sherratt et al. 2017), and different tadpole morphologies can converge to the same adult morphology (Bragg and Bragg 1958; Pfennig 1990). Frogs with a free-living, feeding larval stage may therefore experience fewer developmental or genetic constraints, which may lead to a decrease in strength of phenotypic integration and a greater partitioning of traits into modules, or to greater variation in patterns of phenotypic integration across taxa, both of which would be reflected in greater evolutionary modularity. Some support for this hypothesis is offered by recent analysis of evolutionary modularity in salamanders, with species undergoing complete metamorphosis exhibiting elevated cranial modularity compared with pedomorphic taxa (Fabre et al., 2020). Life history may therefore have a significant and persistent influence on the strength and pattern of cranial integration, and this effect is expected to be particularly strong in lineages with disparate free-living developmental stages, such as frogs.

Other aspects of development may also influence evolutionary integration across frogs. For example, derivation from different cranial neural crest (CNC) streams may generate developmental modules that may be expected to evolve in a coordinated manner (Felice and Goswami 2018). The pattern of the contribution of CNC streams to the osteocranium is considerably different for the frog *Xenopus* than for the axolotl (a salamander) or for amniotes (Hanken and Gross 2005; Gross and Hanken 2008; Piekarski et al. 2014), suggesting a possible deep divergence in the pattern of cranial development between frogs and other tetrapod clades. Furthermore, frogs differ from other lissamphibian clades in cranial ossification sequence timing, with a heterochronic shift whereby frogs exhibit relatively later ossification of bones associated with adult feeding (Harrington et al. 2013). Larval and adult feeding modes differ more for frogs than for salamanders or caecilians, so larval frog mouthparts would likely be more impeded by early ossification of bones involved in adult feeding. However, direct-developing frogs, which feed like adults immediately following hatching, partially reverse this trend (Hanken et al. 1992; Kerney et al. 2007; Harrington et al. 2013). Cranial ossification sequences therefore vary across Anura (Weisbecker and Mitgutsch 2010), and can even vary intraspecifically (Reiss 2002; Moore and Townsend Jr 2003), which may result in greater developmental modularity between bones (Weisbecker and Mitgutsch 2010). Moreover, because early-ossifying bones are generally conservative in timing and function (protecting the brain and otic capsules, Duellman and Trueb 1986; Heatwole and Davies 2003), these elements may be expected to display higher evolutionary integration than later ossifying elements. Ossification sequence timing and derivation from CNC streams may therefore both influence the pattern and

strength of evolutionary integration in frog crania and contribute to differences in evolutionary integration between frogs and other tetrapod clades.

Here, we quantify cranial morphology across 172 anuran species, sampling every family and an extensive range of ecologies and developmental strategies. We implement a high-density data approach (detailed in Bardua et al. 2019a), which has proved successful in capturing morphology across a range of structures and clades (Dumont et al. 2015; Parr et al. 2016; Felice and Goswami 2018; Watanabe et al. 2019; Bardua et al. 2019b). This work represents the most comprehensive analysis of amphibian cranial integration to date, in terms of both taxonomic sampling and density of shape data. With these shape data, we investigate patterns of evolutionary integration and modularity in the frog skull, testing a range of developmental and functional models. We further investigate developmental influences on cranial integration, first by comparing pattern and magnitude of integration in taxa with and without a feeding larval stage, and second by assessing the relationship between magnitude of integration and relative timing of ossification for each cranial bone. Finally, we address the macroevolutionary significance of anuran cranial integration by quantifying the relationship between magnitude of integration and morphological diversity (disparity).

Materials and Methods

SPECIMENS

Our dataset consists of the crania of 172 extant anuran species (one specimen per species), encompassing representatives from all extant families of frogs (Table S1). All specimens were spirit preserved. Determining juvenile from adult specimens can be difficult as adult crania are variably ossified across Anura (Trueb 1973; Nishikawa 2000), so the largest specimen available in the collections visited for each species was selected. Due to lack of data and/or the availability of specimens we did not control for sex in data collection, but it is unlikely that sex-related shape dimorphism would significantly affect results at this scale of macroevolutionary analysis (see Sherratt et al. 2014; Bardua et al. 2019b). Specimens were micro-CT scanned and three-dimensional isosurface models ("meshes") were created from segmenting the volumes using Avizo version 9.3 (FEI, Hillsboro, OR, USA) and VG Studio MAX version 2.0 (Volume Graphics 2001). Meshes were prepared and cleaned in Geomagic Wrap (3D Systems) by removing the mandible from each specimen and removing small foramina texturing the surface, because these can hinder the collection of surface semilandmarks (Bardua et al. 2019a). Data were taken from the right side of each cranium, so 11 specimens were mirrored when this side was incomplete or damaged (Table S1).

PHYLOGENY

Anuran phylogenetic relationships have been the subject of extensive recent study. To incorporate evolutionary relationships into analyses, we used the most recent, most comprehensive phylogeny of Anura (Jetz and Pyron 2018) (Fig. 1). One hundred and sixty-eight anuran species that we sampled are present in this phylogeny, and 140 of our specimens have a phylogenetic position based on direct analysis of molecular data. The consensus phylogeny from Jetz and Pyron (2018) was pruned using the "drop.tip" function in the R package *ape* (Paradis et al. 2004). The four remaining specimens did not have species assignment, so these were added to the phylogeny at the appropriate position for their genus, to allow inclusion in the phylogenetically informed analyses (*Raorchestes* sp., *Dendrobates* sp., *Capensibufo* sp., and *Xenorhina* sp. were added as *Raorchestes anili*, *Dendrobates auratus*, *Capensibufo rosei*, and *Xenorhina varia*, respectively).

LIFE HISTORY

There is a high degree of variation in life history across Anura (Duellman and Trueb 1986; Wells 2010), with nearly 40 reproductive modes defined based on the site of egg development (Haddad and Prado 2005). We coded specimens by life history in terms of possession of feeding larval stage (Table S1). Direct-developing taxa, and some metamorphic species, have nonfeeding larvae provisioned solely with yolk ("endotrophic"), whereas most metamorphic taxa have feeding larvae that feed on external food sources ("exotrophic") (McDiarmid and Altig 1999). This categorization distinguishes larvae that interact with their environment (and presumably experience strong environmental pressures as larvae) from those that do not or do so in a limited way.

MORPHOMETRIC DATA COLLECTION

Among living amphibians, anurans have posed challenges for collecting high-dimensional morphological data. With currently 7,204 species (AmphibiaWeb 2020), frogs are nearly 10 times more species rich than salamanders (742 species), and over 30 times more speciose than caecilians (214 species). Anurans are also the most diverse in larval and adult cranial structure of the extant amphibian orders (Trueb 1993). Many cranial bones are lost as discernible elements repeatedly across this clade, including through fusion between elements or failure to ossify, as well as novel bones evolving in specific subclades (e.g., see Trueb, 1970, 1973; Hall and Larsen 1998; Campos et al. 2010; Schoch 2014; Pereyra et al. 2016; Lambert et al. 2017). Traditional morphometric approaches would fail to capture the morphology of these variably present bones. A high-dimensional approach allows inclusion of this extreme variation in morphology and in bone presence, as this variation may have a crucial influence on patterns of trait integration and modularity across Anura.

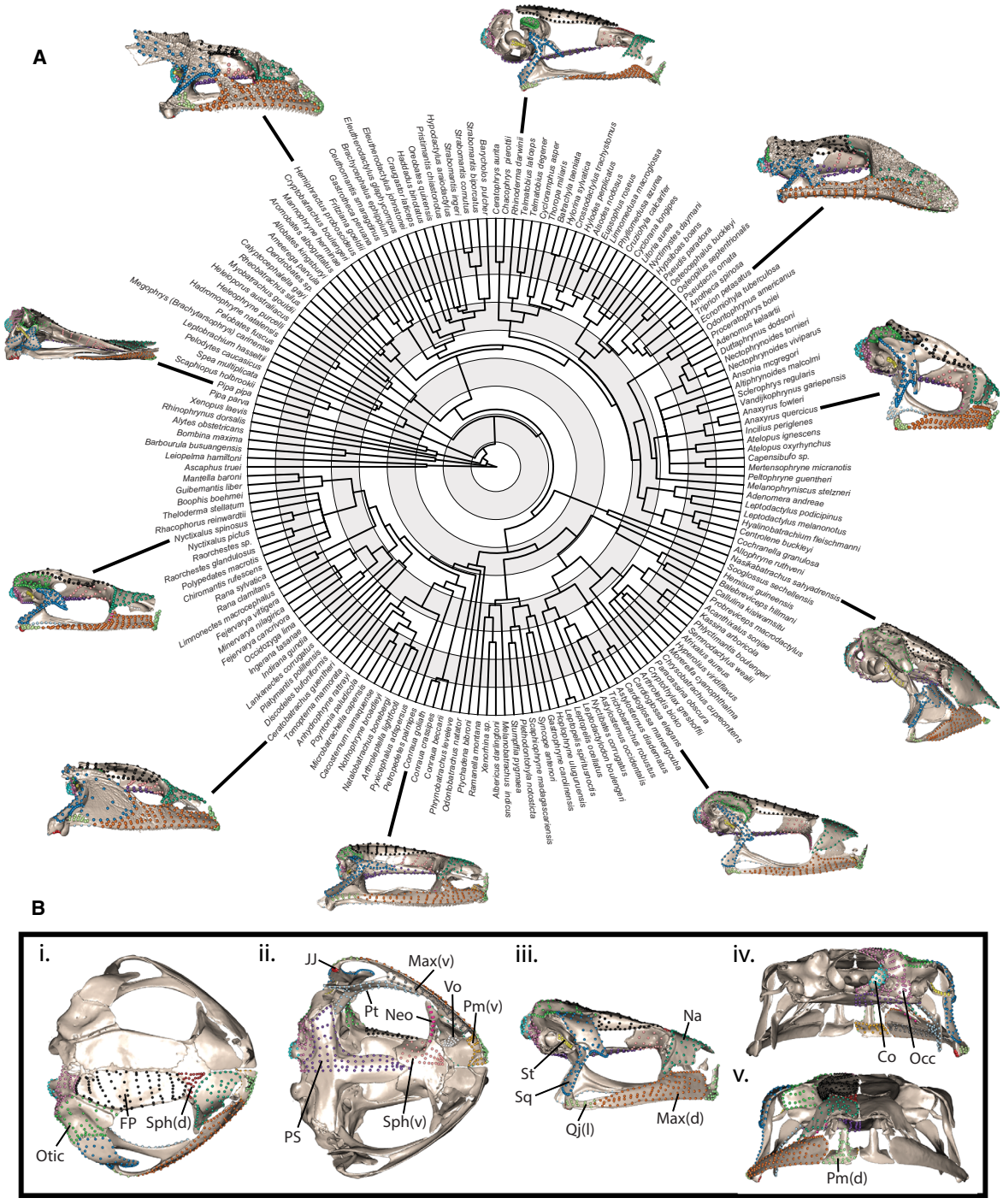

Figure 1. Defined cranial regions displayed on representative skulls across the phylogeny (skulls not to scale). (A) Phylogeny modified from Jetz and Pyron (2018), for the 172 species included in this study. Landmarks and semilandmarks, coloured by cranial region, visualized on 10 skulls (lateral aspects) to illustrate the diversity of cranial shape across Anura. (B) Landmarks and semilandmarks, coloured by cranial region, displayed on *Adenomus kelaartii* (FMNH 1580) in (i) dorsal, (ii) ventral, (iii) lateral, (iv) posterior, and (v) anterior aspect. Regions are as follows: Co (aqua): occipital condyle; Occ (light purple): otooccipital, excluding condyle; FP (black): frontoparietal; JJ (red): quadratojugal (jaw joint articulation); Max (d) (orange): maxilla (dorsal/lateral surface); Max (v) (gray): maxilla (ventral/palatal surface); Na (green): nasal; Neo (hot pink): Neopalatine; Otic (lime green): otic region; Pm (d) (pale green): premaxilla (dorsal/anterior surface, i.e., alary process); Pm (v) (gold): premaxilla (ventral/palatal surface); PS (purple): parasphenoid; Pt (light blue): pterygoid; Qu (l) (light green): quadratojugal (lateral surface, maxillary process); Sph (d) (brown): sphenethmoid (dorsal surface); Sph (v) (light pink): sphenethmoid (ventral surface); Sq (blue): squamosal; St (yellow): stapes; Vo (white): vomer. Branch lengths are scaled to time, with alternate shaded bands indicating periods of 25 million years (outer extreme at Recent).

Regions

We divided the right side of each frog cranium into 19 regions, which represented the maximum reasonable partitioning of the cranial morphology (Fig. 1). In most cases, regions were simply whole bone surfaces, but some bones could be further divided into two distinct regions each (premaxilla, maxilla, otooccipital, quadratojugal, and sphenethmoid; see Table S2 for details). All regions were defined by clear anatomical structures and mostly represented homologous regions, although functionally analogous regions had to be combined in rare cases, for example, when novel bones were present (Table S2). In addition, the otooccipital (exoccipital and opisthotic) can in some cases indistinguishably fuse with the prootic, so for these specimens the "occipital" (∼ otooccipital) and "otic" (∼ prootic) regions were divided along the posterior epiotic ridge.

Approximately half of our specimens had at least one absent region, and over one third of the regions were variably present. Removing regions that were variably present (or specimens lacking these regions) would greatly impact either our quantification of morphology or our sample size (Table S1). To incorporate the full range of skull variation in our analyses, we represented absent regions as "negligible regions," that is, a position of zero size (or near-zero following Procrustes Alignment) described below and previously (Bardua et al. 2019b).

Landmarks and curve semilandmarks

Each cranial region was defined by Type I and Type II landmarks (Bookstein 1991) linked by "curves" (consisting of sliding semilandmarks; Gunz et al. 2005) (Fig. S1). Landmarks and curves were digitized in IDAV Landmark Editor version 3.6 (Wiley et al. 2005). A total of 58 landmarks and 59 curves (ranging from two to 12 semilandmarks, Tables S3–S4) were placed on the right side of each cranium to define all regions present across all specimens. These curves were resampled (for a description and code, see Supporting Information in Botton-Divet et al. 2016), resulting in a total of 410 curve semilandmarks. An additional 24 landmarks and 24 curves were used to define variably present regions in specimens that had those regions represented.

Surface semilandmarks

A semi-automated procedure in the R package *Morpho* version 2.5.1 (Schlager 2017) was used to project surface semilandmarks from a template onto each present cranial region, as described (Schlager 2017) and previously implemented (Felice and Goswami 2018; Marshall et al. 2019; Bardua et al. 2019b). Variably present regions were patched as normal when present, with the additional landmarks and curves defining these regions. These landmarks and curves were then removed prior to analyses, so that these regions were represented only by surface semilandmarks. When variably present regions were absent, these regions were represented by one coordinate position (that best represented the location of each missing region), replicated to achieve an array of dimensions matching that of present regions. This method has been previously implemented (Bardua et al. 2019b), and a similar method was suggested for incorporating novel structures (see fig 1B from Klingenberg 2008). Variably present regions were thus represented only by surface points. A total of 527 surface semilandmarks were applied evenly across each cranium (see Table S5), so that each cranium was represented by a total of 995 landmarks and semilandmarks.

DATA ANALYSES

Procrustes alignment

Non-shape aspects of our data (translation, rotation, and scale) were removed through Procrustes Alignment, using the "gpagen" function in the R package *geomorph* version 3.1.3 (Adams and Otárola-Castillo 2013; Adams et al. 2017). Our data were mirrored prior to alignment, as a bilaterally symmetrical structure aligns more successfully (Cardini 2016). Anuran crania only have two midline landmarks (anterior and posterior extremes of the parasphenoid), so additional midline positions were created by finding the midpoint of two bilaterally symmetrical landmarks (anteromedial and posteromedial extremes of the frontoparietal). Data were mirrored using the "mirrorfill" function in *paleomorph* version 0.1.4 (Lucas and Goswami 2017). Following alignment, the mirrored data were removed, as these data were redundant.

Phylogenetic signal and correction

We calculated K_{mult}, the phylogenetic signal in our shape data under the assumption of Brownian motion, using the "physignal" function in the R package *geomorph* (Adams 2014). A phylogenetic correction was applied to our shape data to account for shared evolutionary history and presumed increased similarity between more closely related species. We computed phylogenetic independent contrasts (Felsenstein 1985) for our shape data and used these in further analyses. We also generated a morphospace for the first two phylogenetic principal components to visualize the spread of cranial shapes within our dataset (Fig. S2).

Allometric signal and correction

Size-related shape changes were investigated by quantifying evolutionary allometry in our shape data. We used the "procD.pgls" function in *geomorph*, which conducts a phylogenetic generalized least squares analysis using log centroid size. Centroid size is the square root of the sum of squared distances of all landmarks from the center (centroid) of a structure. For the centroid size of each cranium, see Table S6. We also visualized size-related shape changes using the "procD.allometry" function in *geomorph* v3.0.5.

Table 1. Alternative models of modular organization tested in EMMLi analysis. Module hypotheses are as follows (parentheses after each hypothesis denotes number of modules): A, fully integrated; B, variable presence; C, functional (snout/rest); D, functional (dorsal/ventral); E, ossification sequence timing (early/mid/late); F, developmental (hyoid/mandibular/branchial CNC); G, functional (snout/rest/squamosal); H, functional (snout/posterior/medial); I, functional (snout/rest/brain region); J, functional (snout/neurocranium/suspensorium/roof); K, functional (facial/sphenoid/cheek/cranial vault/palate/occipital); L, functional (caecilian model); M, by bone (but one large occipital region); N, by bone; O, by region. Hypotheses C, D, G, H, and I modified from Vidal-García and Keogh (2017). Hypotheses E and J modified from Simon and Marroig (2017). Hypothesis F modified from Piekarski et al. (2014). †Tested six versions of this model, allowing for different module designations of the frontoparietal and parasphenoid/sphenethmoid. Occipital and otic region could have also been coded as "none" but this was not investigated. ‡Tested six versions of this model, allowing for different module designations of the nasal and frontoparietal. §Also tested the frontoparietal region in the parietal module of the caecilian model.

	A (1)	B (2)	C (2)	D (2)	E (3)	F† (3)	G (3)	H (3)	I (3)	J‡ (4)	K (6)	L§ (9)	M (13)	N (14)	O (19)
Premaxilla (d)	1	1	1	1	2	1	1	1	1	1	1	1	1	1	1
Maxilla (d)	1	1	1	2	2	2	1	1	1	1	1	2	2	2	2
Parasphenoid	1	1	2	1	1	2	2	2	2	2	2	3	3	3	3
Squamosal	1	1	2	2	2	2	3	2	2	3	3	4	4	4	4
Pterygoid	1	1	2	2	3	2	2	2	2	3	3	5	5	5	5
Frontoparietal	1	1	2	1	1	1	2	3	3	4	4	6	6	6	6
Premaxilla (v)	1	1	1	2	2	2	1	1	1	1	5	7	1	1	7
Maxilla (v)	1	1	1	2	2	2	1	1	1	1	5	2	2	2	8
Nasal	1	1	1	1	2	1	1	1	1	4	1	1	7	7	9
Occipital	1	1	2	1	1	3	2	2	3	4	6	8	8	8	10
Occipital condyle	1	1	2	1	1	3	2	2	3	4	6	8	8	8	11
Otic	1	1	2	1	1	3	2	2	2	2	6	8	8	9	12
Neopalatine	1	2	1	1	3	NA	1	1	1	1	5	2	9	10	13
Stapes	1	2	2	2	NA	NA	2	2	2	2	3	9	10	11	14
Quadratojugal (l)	1	2	2	2	3	2	2	2	2	3	3	4	11	12	15
Quadratojugal (JJ)	1	2	2	2	3	2	2	2	2	3	3	4	11	12	16
Sphenethmoid (d)	1	2	2	1	3	NA	2	3	3	4	4	6	12	13	17
Vomer	1	2	1	1	3	1	1	1	1	1	5	3	13	14	18
Sphenethmoid (v)	1	2	1	1	3	2	2	3	2	2	2	3	12	13	19

To account for size-related shape changes, we corrected our shape data for allometry. We performed a Procrustes ANOVA using the "procD.lm" function in the R package *geomorph*, with log centroid size as a factor, and used the residuals from this analysis in modularity analyses.

Modularity

We hypothesized 27 different model structures, ranging from a fully integrated cranium (one "module"), to every bone or every region as its own module (14 or 19 modules) (see Table 1). We compared a range of models based on function, development, and ossification sequence rank, including models modified from previous studies (Simon and Marroig 2017; Vidal-García and Keogh 2017). These include models based on the contribution of CNC streams to cranial bones, and various divisions of the cranium based on hypothesized functional units. The ossification sequence rank model uses the median rank positions of cranial bones from Weisbecker and Mitgutsch (2010). We also included a model analogous to the 10-module model recovered across caecilian crania (Bardua et al. 2019b). For details on the models, see Text S1.

We investigated patterns of trait integration and modularity using two methods: Evaluating Modularity with Maximum Likelihood (EMMLi) (Goswami and Finarelli 2016) and Covariance Ratio (CR) analysis (Adams 2016), both of which have been implemented in previous analyses of modularity using high-dimensional shape data (Felice and Goswami 2018; Marshall et al. 2019; Watanabe et al. 2019; Bardua et al. 2019b). EMMLi is a maximum likelihood approach that allows the testing of multiple hypotheses of modularity, each of which can vary in its number of modules. This is implemented using the "EMMLi" function in the *EMMLi* R package (Goswami and Finarelli 2016). We tested our 27 different model structures with the shape data, as well as with the phylogenetically- and allometry-corrected shape data. To assess the robustness of our results, we subsampled our shape data down to 10% (using random jackknife resampling), and ran EMMLi iteratively 100 times with these subsampled data using the "subSampleEMMLi" function in *EMMLiv2* (https://github.com/hferg/EMMLiv2/). We compared the average

results from these 100 runs to the results from the original analysis. We also ran EMMLi with the landmark-only dataset to compare results using different data types. The landmark-only dataset excludes variably present regions, as these are represented only by surface points. To test whether the creation of "negligible regions" imposed artificially high integration on variably present regions (because surface points occupy ˜ identical positions in negligible regions), we ran phylogenetically-corrected EMMLi analysis for just the specimens that have every region present ($N = 83$). We compared the pattern of trait integration from this analysis to the original analysis, as well as determining whether within-region trait correlations were different.

We observed the pattern of trait correlations between each module for the best-supported model. However, EMMLi is not exhaustive in its comparison of models, and recent analyses suggest that EMMLi tends to favor more highly parameterized models (Adams and Collyer 2019; Felice et al. 2019; Marshall et al. 2019; Watanabe et al. 2019; Bardua et al. 2019b; Fabre et al. 2020). As such, we used the estimated within- and between-module correlations (ρ) for the best-supported model and grouped highly integrated regions into larger modules as previously described (Felice and Goswami 2018; Marshall et al. 2019; Bardua et al. 2019b), to construct an alternative model of modularity. We grouped regions when the between-module correlation for two modules was within 0.2 of the lowest within-module correlation (Bardua et al. 2019b).

Model selection approaches such as EMMLi only select the best fit model, but they do not explicitly test hypotheses of modularity. For this reason, we also used Covariance Ratio analysis. Covariance Ratio analysis compares the overall covariation between hypothesized modules to the covariation within those modules (Adams 2016). We conducted CR analysis (Adams 2016) for the model identified from EMMLi analysis, using the "modularity.test" function in *geomorph*, to observe whether both methods support similar patterns of modularity. Specifically, we ran this analysis six times: for the uncorrected, allometry-corrected, and phylogenetically-corrected complete datasets, as well as for the uncorrected, allometry-corrected, and phylogenetically-corrected landmark-only datasets.

We further ran EMMLi and CR analyses on subsets of specimens, divided based on presence or absence of a feeding larval stage. We split the dataset into species possessing "feeding" and "nonfeeding" larvae prior to Procrustes alignment, so that both datasets were aligned separately. We then performed phylogenetically informed EMMLi and CR analysis for each subset. However, our imbalance in the number of species with ($N = 124$) and without ($N = 39$) a feeding larval stage may affect strength of integration. We therefore also ran EMMLi and CR analyses iteratively 100 times each on subsets of the larger "feeding larvae" group ($N = 124$), taking 39 specimens at random each time. We then calculated the mean estimated correlations/covariations from these 100 runs and compared these results to the results from the original complete "feeding larvae" dataset ($N = 124$), and to the results from the smaller "nonfeeding larvae" group ($N = 39$). We then compared the effect sizes for strength of modularity (*Z*-scores) of the three datasets using the "compare.cr" function in *geomorph*.

Disparity

Disparity (morphological diversity) of each cranial module was defined as Procrustes variance (calculated using the "morphol.disparity" function in *geomorph*), divided by module landmark/semilandmark number, to correct for landmark/semilandmark number (as this affects variance). A regression of disparity on the estimated magnitude of integration (within-module ρ) from the EMMLi analysis was conducted to understand the influence of the latter on the former.

Ossification sequence rank

The median rank position within the frog cranial ossification sequence was taken for each cranial bone from Weisbecker and Mitgutsch (2010). The stapes was excluded as ossification sequence information was absent for this bone, and the "suspensorium" module was split into its individual bones (quadratojugal, squamosal, and pterygoid). The Spearman's rank correlation coefficient was then calculated between the magnitude of integration (within-module ρ) and the median ossification position for each cranial bone to determine the relationship between these two metrics.

Results

PHYLOGENETIC AND ALLOMETRIC SIGNAL

The anuran crania exhibited significant phylogenetic signal ($K_{\text{mult}} = 0.66$, $P = 1 \times 10^{-4}$). Evolutionary allometry was found to be significant, albeit weak ($R^2 = 0.12$, $P = 1 \times 10^{-4}$).

MODULARITY

EMMLi analyses for the uncorrected, phylogenetically-corrected, and allometry-corrected data all recovered the most parametrized model (the 19-region model) as the best supported. In all three analyses, very similar patterns of trait integration were identified between cranial regions (Figs. 2 and S3; Tables S7–S9). For the reasons noted above, comparisons of within- and between-region trait correlations were conducted, resulting in a novel 13-module model recovered from all three analyses. This new model resulted in the construction of the following multi-region modules: (1) a suspensorium module (quadratojugal (jaw joint and lateral process), pterygoid, and squamosal); (2) a maxilla module (both dorsal and ventral regions); (3) a

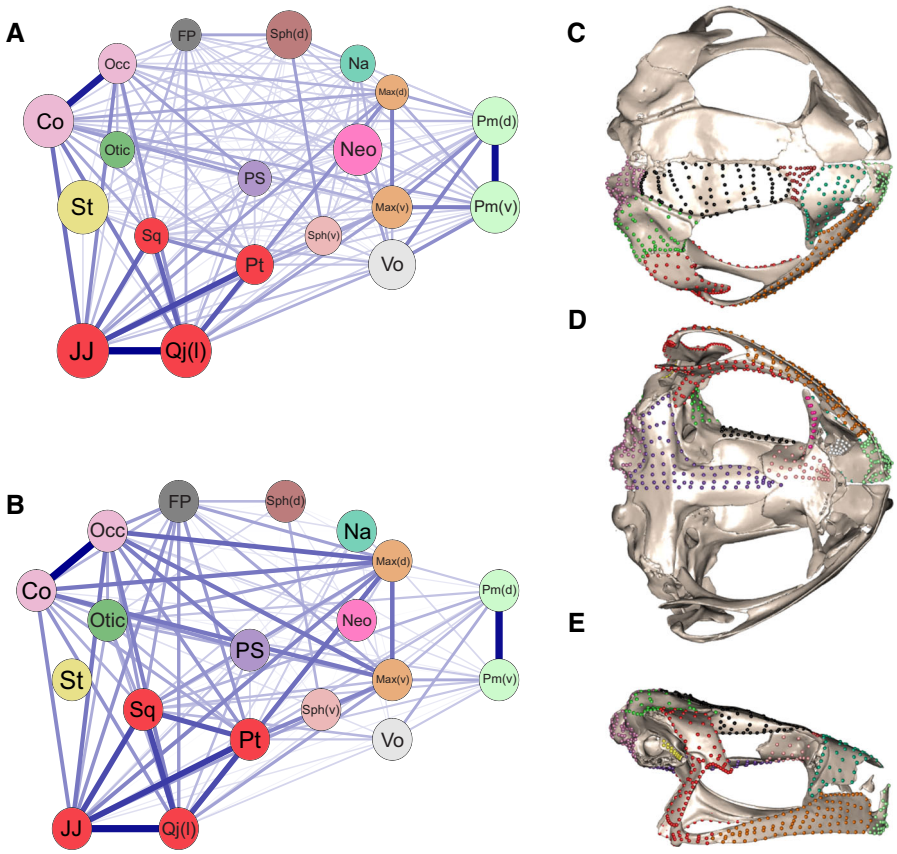

Figure 2. **The 13-module model identified from the 19 cranial regions. Network graphs of the results from (A) phylogenetically-corrected EMMLi analysis and (B) phylogenetically-corrected Covariance Ratio analysis, showing the 19 cranial regions defined in this study, color coded by the 13 identified modules. Regions were grouped into modules when the between-region trait correlations identified from EMMLi analysis (represented by line thickness in [A]) was within 0.2 of the lowest internal trait correlation (represented by circle size in [A]). Line thickness in (B) relates to the covariance ratio for each region pairing, and circle size is held constant. The resulting 13 modules are visualized on *Adenomus kelaartii* (FMNH 1580) in (C) dorsal, (D) ventral, and (E) lateral aspect. Modules are as follows: FP (black): frontoparietal; Max (orange): maxilla (dorsal and ventral surfaces); Na (green): nasal; Neo (hot pink): neopalatine; Occ (light purple): occipital region (Occ) and occipital condyle (Co); Otic (light green): otic region; PS (purple): parasphenoid; Pm (pale green): premaxilla (dorsal and ventral surfaces); Sph (d) (brown): sphenethmoid (dorsal surface); Sph (v) (light pink): sphenethmoid (ventral surface); St (yellow): stapes; Susp (red): suspensorium (squamosal, pterygoid, and both quadratojugal regions); Vo (white): vomer. For region definitions, see Table S2.**

premaxilla module (both dorsal and ventral regions); and (4) an occipital module (occipital and occipital condyle). All remaining regions formed single-region modules: parasphenoid, frontoparietal, nasal, otic, neopalatine, stapes, sphenethmoid (dorsal), sphenethmoid (ventral), and vomer.

Resampling our data down to 10% and taking the average of 100 runs returned very similar results to the full run. The pattern of trait integration was the same, and the identical 13-module model was recovered following the steps outlined above (Fig. S3; Table S10).

The landmark-only EMMLi analyses (uncorrected, phylogenetically-corrected, and allometry-corrected) showed very similar results to one another with only subtle differences (Fig. S4; Tables S11–S13). Two highly integrated modules were recovered for all three landmark-only EMMLi analyses: (1) an occipital module (the occipital and occipital condyle) and (2) a "facial module," consisting of both regions of both the maxilla and premaxilla. Besides these two modules, the within-region trait correlations observed for all landmark-only analyses were considerably lower than for the complete shape data analyses, and extremely low for the nasal, frontoparietal, and maxilla (dorsal) (0.14, 0.23, and 0.27, respectively). Between-region trait correlations were also lower, but only slightly, so that in some cases the between-region correlations were actually higher than the within-region correlations (e.g., occipital within: 0.57, frontoparietal within: 0.23, between: 0.32). Many within-region and between-region trait correlations were below 0.3, and the effect of this was that most regions did not stand out as independent modules, nor could they be grouped with other regions as being strongly integrated. Most of the cranium therefore

appeared unintegrated, with the exception of the occipital and facial modules. The suspensorium module identified in the complete dataset analyses could not be investigated here, as two of these regions (the quadratojugal regions) were not present in the landmark-only data.

Rerunning EMMLi analysis excluding all specimens with any "negligible regions" revealed a near-identical pattern of trait integration to that recovered from the entire dataset for phylogenetically-corrected data (Fig. S5). For this analysis, the within-region trait correlations for the variably present regions were either unchanged, or only marginally different (Table S14) to analysis using the complete dataset. The quadratojugal regions exhibited no change in within-region trait correlations, and the largest difference was a 0.05 decrease in within-region trait integration (stapes and sphenethmoid (ventral) regions). We recovered a modular structure from this analysis that was identical to the model recovered using the complete dataset. The use of the "negligible region" method therefore did not artificially exaggerate within-region trait integration for regions that are variably present.

All three CR analyses of the complete shape data (using uncorrected, phylogenetically-, and allometry-corrected datasets) recovered significant modular signal ($CR = 0.52$, 0.48, and 0.49, respectively, $P = 0.001$ for all). When pairwise CR values were investigated, the pattern of trait relationships was extremely similar to the results from EMMLi for all three analyses of the complete dataset (Fig. S3). For all three analyses, the strongest covariances between regions were those within the identified occipital module ($CR = 0.97$–0.99), premaxilla module ($CR = 0.95$–0.97), and suspensorium module ($CR = 0.78$–0.96) (Tables S15–S17). The maxilla module also showed strong covariation ($CR = 0.73$–0.8), although the maxilla (dorsal) covaried more strongly with the occipital than the maxilla (ventral) using phylogenetically-corrected data and also more strongly with the occipital condyle using uncorrected data.

Conducting CR for the landmark-only datasets (uncorrected, phylogenetically-, and allometry-corrected) revealed weaker, but still significant, modular signal compared with the full datasets ($CR = 0.82$, 0.77, and 0.78, respectively, $P = 0.001$ for all). The pattern of integration observed for these datasets was similar to the landmark-only EMMLi analyses; there was strong covariation within the occipital module ($CR = 1.14$, *phylogenetically-corrected* $CR = 1.12$, *allometry-corrected* $CR = 1.17$, respectively) and strong covariations between the four regions comprising the "facial" module ($CR = 1.04$–1.34, *phylogenetically-corrected* $CR = 0.95$–1.31, *allometry-corrected* $CR = 1.03$–1.17), although the two maxilla regions also covaried strongly with some other regions (Fig. S4; Tables S18–S20).

Table 2. Strength of evolutionary integration and disparity for each cranial module. Results for the 13 identified cranial modules. Integration (within-module correlation) is taken from the phylogenetically-corrected EMMLi analysis.

Module	Integration (phylogenetically-corrected)	Disparity (Procrustes variance) ($\times 10^{-5}$)
Frontoparietal	0.58	1.04
Maxilla	0.57	1.35
Nasal	0.67	1.07
Neopalatine	0.93	1.91
Occipital	0.75	0.72
Otic	0.66	1.13
Parasphenoid	0.64	0.60
Premaxilla	0.87	0.89
Sphenethmoid (d)	0.87	1.29
Sphenethmoid (v)	0.7	1.96
Stapes	0.97	1.86
Suspensorium	0.58	1.76
Vomer	0.9	0.99

Species with and without a feeding larval stage exhibited very similar patterns of trait integration, although specimens without a feeding larval stage had very slightly more strongly integrated crania, as recovered from both phylogenetically informed EMMLi and CR analyses (Fig. S6; Tables S21–S26). However, subsampling the "feeding larvae" group down to 39 specimens and finding the average of running phylogenetically informed EMMLi and CR analyses iteratively 100 times eliminated the very slight difference between the two groups. *Z*-scores for the three datasets (taxa with nonfeeding larvae, taxa with feeding larvae [$N = 124$] and subsampled taxa with feeding larvae) were not significantly different (pairwise *P*-values = 0.72–1.00).

DISPARITY AND EVOLUTIONARY INTEGRATION

Cranial modules displayed a wide range of values for morphological disparity. Morphological disparity was lowest in the parasphenoid (6.00×10^{-6}) and occipital (7.20×10^{-6}) modules, and highest in the sphenethmoid(v) (1.96×10^{-5}), neopalatine (1.91×10^{-5}), stapes (1.86×10^{-5}), and suspensorium (1.76×10^{-5}) modules (Table 2). Magnitude of integration was not significantly correlated with disparity (Multiple $R^2 = 0.05$, $P = 0.47$) (Fig. 3).

OSSIFICATION SEQUENCE

There was a significant relationship observed between ossification sequence rank and magnitude of integration for individual cranial bones (Spearman's rho = 0.59, $P = 0.027$) (Fig. 4).

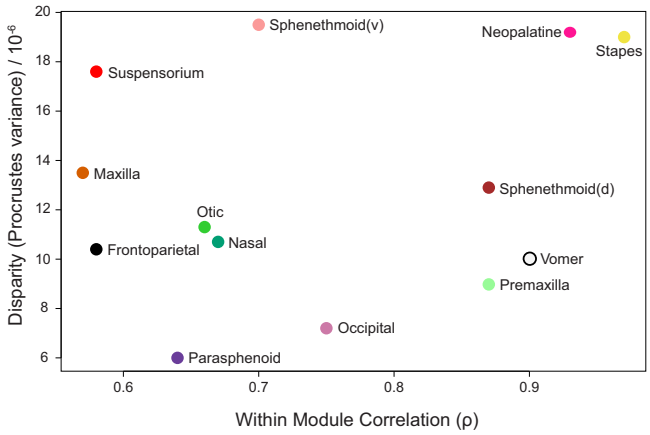

Figure 3. Magnitude of integration versus disparity. Regression of strength of integration (estimated within-module correlation) against disparity (Procrustes variance) for each cranial module.

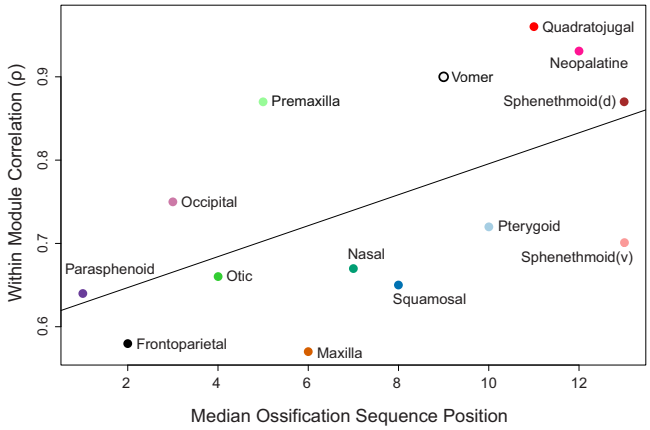

Figure 4. Influence of ossification sequence timing on integration, using the median rank position within the frog cranial ossification sequence for each cranial bone (Weisbecker and Mitgutsch 2010). The "suspensorium" module was split into its constituent bones (quadratojugal, squamosal, and pterygoid) and the stapes was excluded. Spearman's rank correlation between ossification sequence timing and integration (within-module correlation) revealed the relationship was significant (Spearman's rank correlation rho = 0.59, *P* = 0.027).

Discussion

Trait integration is of fundamental interest in reconstructing the evolution of phenotype and of phenotypic diversity. When integrated traits experience divergent selection pressures, their covariation can hinder the ability of each trait to evolve toward its respective optimum. This constraint may drive the fragmentation of sets of integrated traits into smaller, functionally linked units comprising traits with aligned selection pressures and with reduced linkages to other sets of traits, thus freeing modules from the constraints of integration with functionally unrelated traits. This modular organization may thus be expected to facilitate evolutionary diversification. Although it is the variational, rather than the evolutionary, relationships among traits that shape evolvability and influence response to selection, evolutionary modularity informs on the outcome of those processes.

Variational and evolutionary modularity often correspond, for example, in caecilians (Marshall et al. 2019; Bardua et al. 2019b) and in salamanders (Bon et al. 2020; Fabre et al. 2020), suggesting that the genetic, developmental, and functional interactions of traits that shape phenotypic integration are replicated in patterns of trait evolution. As discussed above, many studies have suggested that evolutionary modularity promotes diversification of form. Studies have also found that high levels of phenotypic and evolutionary integration within modules are associated with limited morphological diversity (disparity) and/or evolutionary rate of those modules (Goswami and Polly 2010; Felice and Goswami 2018), or conversely, that strong integration of traits is associated with higher disparity (Claverie and Patek 2013; Parr et al. 2016; Randau and Goswami 2017; Navalón et al. 2020). Many other studies suggest there may be no simple relationship between integration and either morphological disparity or rate at either level (Drake and Klingenberg 2010; Goswami et al. 2014; Felice et al. 2018; Marshall et al. 2019; Watanabe et al. 2019; Bardua et al. 2019b). Improving our understanding of the relationship between modularity in trait evolution and morphological diversification ultimately requires more (and more diverse) empirical data on modularity and morphological evolution within species and across clades. Here, we have expanded the study of evolutionary modularity to a taxonomically, developmentally, ecologically, and morphologically diverse tetrapod clade: frogs.

HIGHLY MODULAR ANURAN CRANIA

We find anuran crania evolve in a highly modular manner, exhibiting 13 distinct evolutionary modules. Our results contrast with previous analyses recovering only weak support for phenotypic or evolutionary modularity across frog crania (Simon and Marroig 2017; Vidal-García and Keogh 2017), and we find no support for the functional models investigated in these studies. However, those studies investigated modularity at a finer taxonomic scale, sampling at subgenus or family level, respectively, with likely greater similarity (and less variation) in function and ecology than is sampled here. Furthermore, differences in data type and study design hinder direct comparisons, with our study the first to investigate highly modular hypotheses with dense morphometric data. Our analyses suggest a complex model, where osteological units were mainly recovered as either distinct modules (premaxilla and maxilla) or split into multiple modules (e.g., dorsal and ventral surfaces of the sphenethmoid). One multi-region module, the "suspensorium," comprises three bones (quadratojugal, squamosal, and pterygoid) and corresponds to a similar multi-bone module in salamanders and caecilians

(Bardua et al. 2019b; Fabre et al. 2020). Function (adult feeding) thus appears to be a more proximal driver of evolutionary integration than development (i.e., tissue origin) is for this region.

The suspensory apparatus braces and suspends the jaws against the neurocranium for autosystylic jaw suspension and has been long recognized as an important functional unit in frogs (Trueb 1973). The squamosal and pterygoid act to protect and strengthen the cartilaginous system (including the quadrate), and the quadrate (which ossifies in some species) is often invaded by the ossified quadratojugal (Trueb 1973). The suspensorium likely acts as a functional unit of selection with studies noting coordinated change in this unit. For example, the suspensorium has been found to be more vertically oriented in miniaturized anuran species, in which a more rostrally positioned jaw joint articulation leads to a smaller gape size (Yeh 2002). Also, direct-developing species have been shown to have accelerated appearance of the jaw and suspensorium bones through their cranial bone formation (Hanken et al. 1992; Kerney et al. 2007; Vassilieva 2017), demonstrating coordinated changes of the suspensorium module due to heterochronic repatterning. The range of feeding mechanisms across frogs associated with different environments, for example, aquatic suction feeding (Carreño and Nishikawa 2010; Fernandez et al. 2017) and hydrostatic elongation in many fossorial (and some terrestrial) species (e.g., Trueb and Gans 1983; Nishikawa et al. 1999), may have been facilitated by an integrated jaw suspensorium module, allowing these mechanisms to arise convergently multiple times (Nishikawa 2000).

Life history strategies are extremely varied across frogs, and differences in strategy (in particular, the presence or absence of a larval stage that interacts with its environment) may be expected to exert an influence on the structure of evolutionary modules. Frogs undergo major morphological restructuring of the cranium through metamorphosis (Rose and Reiss 1993), either inside (direct-developing species; Callery and Elinson 2000; Ziermann and Diogo 2013) or outside (biphasic species) the egg. Unlike frogs with nonfeeding larvae, frogs with a feeding larval stage need to adapt to two environments, each with unique functional pressures (Reiss 2002). Species with feeding larvae thus may have greater autonomy across stages and consequently fewer developmental constraints, which may be expected to manifest in greater variation in patterns of phenotypic integration and thus greater evolutionary modularity. In contrast, taxa with nonfeeding larvae may be more developmentally constrained, resulting in greater similarity in patterns of phenotypic integration across species and thus stronger evolutionary integration. Surprisingly, despite these expectations, we do not find evolutionary integration to differ substantially between taxa with and without a feeding larval stage. A recent study on salamander crania (Fabre et al. 2020) found that taxa undergoing complete metamorphosis (biphasic and direct-developing) displayed more modular crania than did pedomorphic taxa. Thus, metamorphosis in salamanders may represent an adaptive mechanism through which developmental processes shaping morphological variation are decoupled across different life history stages (Moran 1994; Fabre et al. 2020). It may therefore be the process of metamorphosis, and not the interaction with a larval environment, that results in a decoupling of life history stages. Furthermore, life history may be a bigger influence on cranial morphological evolution of salamanders than frogs. Ecology (habitat) has been found to be correlated with skull shape in frogs (Paluh et al. 2020) and may be a primary factor shaping cranial morphological evolution, in contrast to salamanders, where ecology is a relatively minor factor compared to life history (Fabre et al. 2020). Thus, the influence of life history on evolutionary modularity may be confounded by ecological influences, and the process of metamorphosis may be a greater driver of trait fragmentation than the possession of life stages with distinct ecological pressures.

Developmental associations of traits are often hypothesized to drive phenotypic and evolutionary integration (e.g., Simon and Marroig 2017). We found no evidence of strong relationships among regions sharing contributions from CNC streams or with similar ossification sequence ranks. However, frogs (evidenced by *Xenopus*) exhibit a considerably different pattern of CNC contributions to cranial elements compared with salamanders and other vertebrates (Hanken and Gross 2005; Gross and Hanken 2008; Piekarski et al. 2014). The pattern of CNC contributions therefore appears evolutionarily labile (Hanken and Gross 2005) and may have been changing throughout anuran evolution, which is not captured in the CNC model of modularity.

An additional developmental influence on the strength of cranial integration is the timing of bone ossification. Cranial ossification sequences are highly evolvable across Anura (Weisbecker and Mitgutsch 2010), with large variation in the relative timing of cranial bone ossification, although early and late ossifying bones are generally the most conservative. Contrary to our expectation, ossification sequence rank correlated positively with the strength of evolutionary integration within cranial bones. The latest-ossifying bones showed the highest evolutionary integration and are those that are lost in many taxa, possibly due to pedomorphosis (Weisbecker and Mitgutsch 2010). The intermediate-ossifying bones, which are generally the least conserved in ossification sequence rank (Weisbecker and Mitgutsch 2010; Vassilieva 2017), displayed the broadest range of magnitudes of integration. Direct and biphasic developers differ most in the relative timing of intermediate-ossifying bones (Vassilieva 2017), so the wide range of timings of intermediate-ossifying bones across taxa may therefore have weakened the relationship between ossification and integration across these bones. Regardless, the significant relationship between ossification sequence rank and cranial integration across frogs, compared with the lack of influence

of life history on integration, suggests at least some aspects of development do significantly influence the coordinated evolution of traits in the anuran cranium.

A COMMON PATTERN OF LISSAMPHIBIAN EVOLUTIONARY MODULARITY

A large, multi-region module associated with feeding and prey capture appears to be common across lissamphibian lineages (Bardua et al. 2019b; Fabre et al. 2020), comprising the lateral and articular surfaces of the quadrate/quadratojugal, squamosal, and, in frogs and biphasic salamanders, pterygoid. This tight functional feeding unit is thus replicated across lissamphibians and selection on this module is likely to be a primary and persistent factor shaping morphological evolution across lissamphibians. The model of evolutionary modules in the anuran cranium supported here is remarkably similar in number and pattern to that recovered recently from analyses of both phenotypic and evolutionary modularity in caecilian (Marshall et al. 2019; Bardua et al. 2019b) and salamander (Bon et al. 2020; Fabre et al. 2020) crania, despite substantial morphological differences among these clades. Replication of patterns of cranial modularity from static to evolutionary level across caecilians and salamanders has already been demonstrated (Marshall et al. 2019; Bardua et al. 2019b; Bon et al. 2020; Fabre et al. 2020), and now our results suggest that these evolutionary modules are shared across Lissamphibia. Lissamphibian crania also display more modular evolution than amniote crania, which may be partly attributed to their more frequent losses and regains of cranial elements as well as their greater number of cranial elements. Studies of amniote crania using similar morphometric data have identified numbers of evolutionary modules ranging from seven (Aves; Felice et al. 2018), to nine (snakes; Watanabe et al. 2019) or 10 (dingos and dogs; Parr et al. 2016; lizards; Watanabe et al. 2019). Comparison of patterns of trait evolutionary integration recovered from these studies reveals some striking similarities. For example, the occipital region and the jaw joint region are both strongly integrated across caecilians (Bardua et al. 2019b), salamanders (Fabre et al. 2020), birds (Felice and Goswami 2018), squamates (Watanabe et al. 2019), crocodylomorphs and nonavian dinosaurs (Felice et al. 2019), and, as demonstrated here, frogs. However, caecilians, salamanders, and frogs exhibit the strongest concordance in their patterns of cranial integration, suggesting a possible amniote-amphibian divergence in the structure of evolutionary modules of the cranium.

Other studies of tetrapod cranial modularity have recovered weaker or less complex modular structures, including six-module models (mammals, Cheverud 1982; Goswami 2006; Goswami and Polly 2010; Goswami and Finarelli 2016), two-module models (Alpine newt, Ivanović and Kalezić 2010; Anolis lizards Sanger et al. 2012; lacertids Urošević et al. 2018), or no modules at all (i.e., fully integrated) (birds, Klingenberg and Marugán-Lobón 2013). However, many of these studies used many fewer two-dimensional or three-dimensional landmarks and thus necessarily tested less-complex patterns of modularity, hindering comparison with studies incorporating surface semilandmarks.

EVOLUTIONARY MODULARITY AND MORPHOLOGICAL EVOLUTION

With these new data, we return to the question of whether evolutionary modularity has promoted the diversification of the frog cranium. Our results demonstrate that anuran cranial modules vary widely in terms of morphological disparity, suggesting that modular variation has indeed promoted the mosaic evolution of the anuran cranium. However, the magnitude of evolutionary integration within each cranial module does not show a strong or consistent relationship with disparity, similar to findings from caecilian (Bardua et al. 2019b), salamander (Fabre et al. 2020), and squamate (Watanabe et al. 2019) crania. Other studies have found high evolutionary integration to correspond with either higher or lower disparity or rates of evolution in different clades (Goswami and Polly 2010; Claverie and Patek 2013; Parr et al. 2016; Randau and Goswami 2017; Felice and Goswami 2018; Navalón et al. 2020). The observed relationship between evolutionary integration and morphological evolution depends on many factors, including the alignment of the direction of selection with the path that phenotypic integration facilitates in morphospace and the evolutionary stability of both of those attributes, as well as biases in fully capturing morphological diversity due to extinction. The relationship evolutionary integration has with morphological disparity is therefore likely complex, with patterns for each module reflecting their diverse selection histories, rather than following a single simple relationship across a larger structure.

Finally, on a more practical note, comparing landmark-only analyses to our analyses of more complete descriptions of cranial shape highlights the benefits of implementing a surface-based approach for investigating modularity across frog crania. The potential for landmarks to amplify between-region trait correlations for adjacent regions and understate within-region correlations (Goswami et al. 2019; Marshall et al. 2019; Bardua et al. 2019b) may result in the cranium appearing less modular, and more integrated, than when incorporating curve and surface information. In addition, crucially, landmarks fail to capture seven cranial regions in anurans, because these are variably present across the clade. This exclusion of nearly one third of the regions from modularity analyses would prevent patterns of integration from being comprehensively explored across the whole cranium. We assigned "negligible regions" in taxa with absent regions (Bardua et al. 2019a,b), and our comparison of analyses including and

excluding these taxa demonstrates that these "negligible regions" do not affect the observed pattern of integration, as results are near-identical with and without taxa lacking those regions. Thus, our high-dimensional surface-based approach allows a more representative quantification of trait integration and cranial morphological evolution across frogs.

Conclusion

We implemented a high-density approach to quantify anuran crania across the entire clade, with extensive sampling of taxonomy and morphology. We identified 13 evolutionary modules, which replicate patterns recovered across caecilians and salamanders, but are less similar to those observed in amniotes, suggesting a possible lissamphibian-amniote divergence in cranial modularity. Surprisingly, contrasting life histories have no impact on evolutionary integration of the anuran skull, in contradiction to the hypothesis that a feeding larval stage may have arisen as a mechanism to fragment evolutionary associations among regions (although it may well apply to different stages). Timing of ossification shows a significant correlation with evolutionary integration, with late-ossifying bones displaying higher magnitudes of integration than earlier-ossifying bones. Cranial modules display a wide range of morphological disparities, but magnitude of within-module integration did not correspond significantly with module disparity, supporting recent studies suggesting a complex relationship between evolutionary integration and morphological diversification. Finally, we have illustrated the utility of this high-dimensional approach, for quantifying cranial morphology across every frog family, allowing for incorporation of regions that were absent in some specimens. Applying a similar high-dimensional approach to other extremely diverse clades, many of which are currently undersampled, will facilitate the study of phenotypic and evolutionary integration in these clades and allow broader comparisons across different scales of analysis and across the tree of life.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

AUTHOR CONTRIBUTIONS

CB and AG designed the study. CB, ELS, KD, MB, and DCB generated and reconstructed the scans. CB undertook data collection, performed analyses, and primarily wrote the manuscript. All authors provided ideas and discussion, contributed to the writing of the manuscript, and read and approved the final version.

ACKNOWLEDGMENTS

We would like to thank R.N. Felice, J. Clavel, and A. Watanabe for help with analyses, and members of the Goswami Lab for helpful discussions. We would like to thank A.-M. Ohler and A. Herrel for help with CT scanning at the MNHN collection, Paris, and K. Mahlow, N. Fröbisch, F. Tillack, J. Müller, and M.-O. Rödel for facilitating the CT scanning at the MfN, Berlin. The CT scan data for specimens scanned at the MfN are available at https://doi.org/10.7479/cb53-6d14. We would also like to thank two anonymous reviewers and the Associate Editor Dr. M. Zelditch for their comments, which greatly helped to improve this manuscript. This work was supported by ERC grant STG-2014–637171 to AG. This research also received support from the SYNTHESYS Project http://www.SYNTHESYS.info/ (FR-TAF-5583), which is financed by European Community Research Infrastructure Action under the FP7 Integrating Activities Program. Data collection performed at the University of Florida was supported in part by National Science Foundation (DBI-1701714) to DCB and ELS. MB was supported by Erasmus+. Funding for scanning *Chiromantis rufescens* MCZ A 136738, *Rhacophorus reinwardtii* FMNH 252384, and *Theloderma stellatum* FMNH 253619 was provided by a National Science Foundation Grant (DEB-0345885) to Dr. Hanken, and funding for image processing was provided by an National Science Foundation Digital Libraries Initiative grant to Dr. T. Rowe of The University of Texas at Austin. Funding for scanning *Incilius (Bufo) periglenes* TNHC 39447 was provided by Dr. Cossel and funding for image processing was provided by the High-Resolution X-ray CT Facility.

DATA ARCHIVING

Surface meshes of all specimens are available on Phenome10k.org, Morphosource.org, and DigiMorph.org. Data and code are provided on https://github.com/anjgoswami/frogs_modularity.

REFERENCES

Adams, D. C. 2014. A generalized K statistic for estimating phylogenetic signal from shape and other high-dimensional multivariate data. Syst. Biol. 63:685–697.

———. 2016. Evaluating modularity in morphometric data: challenges with the RV coefficient and a new test measure. Methods Ecol. Evol. 7:565–572.

Adams, D. C., and M. L. Collyer. 2019. Comparing the strength of modular signal, and evaluating alternative modular hypotheses, using covariance ratio effect sizes with morphometric data. Evolution 73:2352–2367.

Adams, D. C., and E. Otárola-Castillo. 2013. geomorph: an R package for the collection and analysis of geometric morphometric shape data. Methods Ecol. Evol. 4:393–399.

Adams, D. C., M. L. Collyer, A. Kaliontzopoulou, and E. Sherratt. 2017. Geomorph: software for geometric morphometric analyses. R package version 3.1.3.

AmphibiaWeb. 2020. AmphibiaWeb: information on amphibian biology and conservation. Available via https://amphibiaweb.org.

Bardua, C., R. N. Felice, A. Watanabe, A.-C. Fabre, and A. Goswami. 2019a. A practical guide to sliding and surface semilandmarks in morphometric analyses. Integr. Org. Biol. 1:1–34.

Bardua, C., M. Wilkinson, D. J. Gower, E. Sherratt, and A. Goswami. 2019b. Morphological evolution and modularity of the caecilian skull. BMC Evol. Biol. 19:1–23.

Bon, M., C. Bardua, A. Goswami, and A.-C. Fabre. 2020. Cranial integration in the fire salamander, *Salamandra salamandra* (Caudata: Salamandridae). Biol. J. Linn. Soc. 130:178–194. https://doi.org/10.1093/biolinnean/blaa020

Bookstein, F. L. 1991. Morphometric tools for landmark data: geometry and biology. Cambridge Univ. Press, Cambridge, U.K.

Botton-Divet, L., R. Cornette, A.-C. Fabre, A. Herrel, and A. Houssaye. 2016. Morphological analysis of long bones in semi-aquatic mustelids and their terrestrial relatives. Integr. Comp. Biol. 56:1298–1309.

Bragg, A. N., and W. N. Bragg. 1958. Variations in the mouth parts in tadpoles of *Scaphiopus (Spea) bombifrons* Cope (Amphibia: Salientia). Southwest. Nat. 3:55–69.

Callery, E. M., and R. P. Elinson. 2000. Thyroid hormone-dependent metamorphosis in a direct developing frog. Proc. Natl. Acad. Sci. USA 97:2615–2620.

Campos, L. A., H. R. Da Silva, and A. Sebben. 2010. Morphology and development of additional bony elements in the genus *Brachycephalus* (Anura: Brachycephalidae). Biol. J. Linn. Soc. 99:752–767.

Cardini, A. 2016. Left, right or both? Estimating and improving accuracy of one-side-only geometric morphometric analyses of cranial variation. J. Zool. Syst. Evol. Res. 55:1–10.

Carreño, C. A., and K. C. Nishikawa. 2010. Aquatic feeding in pipid frogs: the use of suction for prey capture. J. Exp. Biol. 213:2001–2008.

Cheverud, J. M. 1982. Phenotypic, genetic, and environmental morphological integration in the cranium. Evolution 36:499–516.

———. Quantitative genetics and developmental constraints on evolution by selection. J. Theor. Biol. 110:155–171.

Claverie, T., and S. N. Patek. 2013. Modularity and rates of evolutionary change in a power-amplified prey capture system. Evolution 67:3191–3207.

De Beer, G. 1954. Archaeopteryx lithographica: a study based upon the British Museum specimen. No. 224. order Trust. Br. Museum.

Dellinger, A. S., S. Artuso, S. Pamperl, F. A. Michelangeli, D. S. Penneys, D. M. Fernández-Fernández, M. Alvear, F. Almeda, W. Scott Armbruster, Y. Staeder, et al. 2019. Modularity increases rate of floral evolution and adaptive success for functionally specialized pollination systems. Commun. Biol. 2:1–11.

Drake, A. G., and C. P. Klingenberg. 2010. Large-scale diversification of skull shape in domestic dogs: disparity and modularity. Am. Nat. 175:289–301.

Duellman, W. E., and L. Trueb. 1986. Biology of the amphibians. McGraw-Hill Publishing Company, New York.

Dumont, M., C. E. Wall, L. Botton-Divet, A. Goswami, S. Peigné, and A.-C. Fabre. 2015. Do functional demands associated with locomotor habitat, diet, and activity pattern drive skull shape evolution in musteloid carnivorans? Biol. J. Linn. Soc. 117:858–878.

Ebenman, B. 1992. Evolution in organisms that change their niches during the life cycle. Am. Nat. 139:990–1021.

Fabre, A., C. Bardua, M. Bon, J. Clavel, R. N. Felice, J. W. Streicher, J. Bonnel, E. L. Stanley, D. C. Blackburn, and A. Goswami. 2020. Metamorphosis shapes cranial diversity and rate of evolution in salamanders. Nat. Ecol. Evol. https://doi.org/10.1038/s41559-020-1225-3.

Felice, R. N., and A. Goswami. 2018. Developmental origins of mosaic evolution in the avian cranium. Proc. Natl. Acad. Sci. USA 115:555–560.

Felice, R. N., M. Randau, and A. Goswami. 2018. A fly in a tube: macroevolutionary expectations for integrated phenotypes. Evolution 72:2580–2594.

Felice, R. N., A. Watanabe, A. R. Cuff, E. Noirault, D. Poll, L. M. Witmer, M. A. Norell, P. M. O'Connor, and A. Goswami. 2019. Evolutionary integration and modularity in the archosaur cranium. Integr. Comp. Biol. 59:371–382.

Felsenstein, J. 1985. Phylogenies and the comparative method. Am. Nat. 125:1–15.

Fernandez, E., F. Irish, and D. Cundall. 2017. How a frog, *Pipa pipa*, succeeds or fails in catching fish. Copeia 105:108–119.

Goswami, A. 2006. Cranial modularity shifts during mammalian evolution. Am. Nat. 168:270–280.

Goswami, A., and J. A. Finarelli. 2016. EMMLi: a maximum likelihood approach to the analysis of modularity. Evolution 70:1622–1637.

Goswami, A., and P. D. Polly. 2010. The influence of modularity on cranial morphological disparity in Carnivora and primates (Mammalia). PLoS ONE 5:e9517.

Goswami, A., J. B. Smaers, C. Soligo, and P. D. Polly. 2014. The macroevolutionary consequences of phenotypic integration: from development to deep time. Philos. Trans. R. Soc. Lond. B Biol. Sci. 369:20130254.

Goswami, A., A. Watanabe, R. N. Felice, C. Bardua, A.-C. Fabre, and P. D. Polly. 2019. High-density morphometric analysis of shape and integration: the good, the bad, and the not-really-a-problem. Integr. Comp. Biol. 59:669–683.

Gross, J. B., and J. Hanken. 2008. Segmentation of the vertebrate skull: neural-crest derivation of adult cartilages in the clawed frog, *Xenopus laevis*. Integr. Comp. Biol. 48:681–696.

Gunz, P., P. Mitteroecker, and F. L. Bookstein. 2005. Chapter three: semilandmarks in three dimensions. Pp. 73–98 *in* D. E. Slice, ed. Modern morphometrics in physical anthropology. Kluwer Academic/Plenum, New York.

Haddad, C. F. B., and C. P. A. Prado. 2005. Reproductive modes in frogs and their unexpected diversity in the Atlantic forest of Brazil. Bioscience 55:207–217.

Hall, J. A., and J. H. Larsen. 1998. Postembryonic ontogeny of the Spadefoot Toad. Scaphiopus intermont*anus* (Anura: Pelobatidae): skeletal morphology. J. Morphol. 238:179–244.

Hanken, J., and J. B. Gross. 2005. Evolution of cranial development and the role of neural crest: insights from amphibians. J. Anat. 207:437–446.

Hanken, J., M. W. Klymkowsky, C. H. Summers, D. W. Seufert, and N. Ingebrigtsen. 1992. Cranial ontogeny in the direct-developing frog, *Eleutherodactylus coqui* (Anura: Leptodactylidae), analyzed using whole-mount lmmunohistochemistry. J. Morphol. 211:95–118.

Harrington, S. M., L. B. Harrison, and C. A. Sheil. 2013. Ossification sequence heterochrony among amphibians. Evol. Dev. 15:344–364.

Heatwole, H., and M. Davies, eds. 2003. Amphibian biology (volume 5), Osteology. Surrey Beatty & Sons, Chipping Norton, U.K.

Ivanović, A., and J. W. Arntzen. 2014. Evolution of skull and body shape in *Triturus* newts reconstructed from three-dimensional morphometric data and phylogeny. Biol. J. Linn. Soc. 113:243–255.

Ivanović, A., and M. L. Kalezić. 2010. Testing the hypothesis of morphological integration on a skull of a vertebrate with a biphasic life cycle: a case study of the alpine newt. J. Exp. Zool. Part B Mol. Dev. Evol. 314B:527–538.

Jetz, W., and R. A. Pyron. 2018. The interplay of past diversification and evolutionary isolation with present imperilment across the amphibian tree of life. Nat. Ecol. Evol. 2:850–858.

Kerney, R., M. Meegaskumbura, K. Manamendra-Arachchi, and J. Hanken. 2007. Cranial ontogeny in *Philautus silus* (Anura: Ranidae: Rhacophorinae) reveals few similarities with other direct-developing anurans. J. Morphol. 268:715–725.

Klingenberg, C. P. 2008. Novelty and "homology-free" morphometrics: what's in a name? Evol. Biol. 35:186–190.

Klingenberg, C. P., and J. Marugán-Lobón. 2013. Evolutionary covariation in geometric morphometric data: analyzing integration, modularity, and allometry in a phylogenetic context. Syst. Biol. 62:591–610.

Lambert, S. M., C. R. Hutter, and M. D. Scherz. 2017. Diamond in the rough: a new species of fossorial diamond frog (*Rhombophryne*) from Ranomafana National Park, southeastern Madagascar. Zoosyst. Evol. 93:143–155.

Larouche, O., M. L. Zelditch, and R. Cloutier. 2018. Modularity promotes morphological divergence in ray-finned fishes. Sci. Rep. 8:7278.

Lucas, T., and A. Goswami. 2017. paleomorph: geometric morphometric tools for paleobiology. R package version 0.1.4.

Marshall, A. F., C. Bardua, D. J. Gower, M. Wilkinson, E. Sherratt, and A. Goswami. 2019. High-density three-dimensional morphometric analyses support conserved static (intraspecific) modularity in caecilian (Amphibia: Gymnophiona) crania. Biol. J. Linn. Soc. 126:721–742.

Martín-Serra, A., O. Nanova, C. Varón-González, G. Ortega, and B. Figueirido. 2019. Phenotypic integration and modularity drives skull shape divergence in the Arctic fox (*Vulpes lagopus*) from the Commander Islands. Biol. Lett. 15:20190406.

McDiarmid, R. W., and R. Altig, eds. 1999. Tadpoles: the biology of anuran larvae. Univ. of Chicago Press, Chicago, IL.

Moore, M. K., and V. R. Townsend Jr. 2003. Intraspecific variation in cranial ossification in the tailed frog, *Ascaphus truei*. J. Herpetol. 37:714–717.

Moran, N. A. 1994. Adaptation and constraint in the complex life cycles of animals. Annu. Rev. Ecol. Syst. 25:573–600.

Navalón, G., J. Marugán-lobón, J. A. Bright, C. R. Cooney, and E. J. Rayfield. 2020. The consequences of craniofacial integration for the adaptive radiations of Darwin's finches and Hawaiian honeycreepers. Nat. Ecol. Evol. 4:270–278.

Nishikawa, K. C. 2000. Feeding in frogs. Pp. 117–148 *in* K. Schwenk, ed. Feeding: form, function and evolution in tetrapod vertebrates. Academic Press, San Diego, CA.

Nishikawa, K. C., W. M. Kier, and K. K. Smith. 1999. Morphology and mechanics of tongue movement in the African pig-nosed frog *Hemisus marmoratum*: a muscular hydrostatic model. J. Exp. Biol. 202: 771–780.

Paluh, D. J., E. L. Stanley, and D. C. Blackburn. 2020. Evolution of hyperossification expands skull diversity in frogs. Proc. Natl. Acad. Sci. USA 117:8554–8562.

Paradis, E., J. Claude, and K. Strimmer. 2004. APE: analyses of phylogenetics and evolution in R language. Bioinformatics 20:289–290.

Parr, W. C. H., L. A. B. Wilson, S. Wroe, N. J. Colman, M. S. Crowther, and M. Letnic. 2016. Cranial shape and the modularity of hybridization in dingoes and dogs; hybridization does not spell the end for native morphology. Evol. Biol. 43:171–187.

Pereyra, M. O., M. C. Womack, J. S. Barrionuevo, B. L. Blotto, D. Baldo, M. Targino, J. J. Ospina-Sarria, J. M. Guayasamin, L. A. Coloma, K. L. Hoke, et al. 2016. The complex evolutionary history of the tympanic middle ear in frogs and toads (Anura). Sci. Rep. 6:1–9.

Pfennig, D. 1990. The adaptive significance of an environmentally-cued developmental switch in an anuran tadpole. Oecologia 85:101–107.

Piekarski, N., J. B. Gross, and J. Hanken. 2014. Evolutionary innovation and conservation in the embryonic derivation of the vertebrate skull. Nat. Commun. 5:1–9.

Randau, M., and A. Goswami. 2017. Unravelling intravertebral integration, modularity and disparity in Felidae (Mammalia). Evol. Dev. 19:85–95.

Reiss, J. O. 2002. The phylogeny of amphibian metamorphosis. Zoology 105:85–96.

Riedl, R. 1978. Order in living organisms. John Wiley & Sons, New York.

Rose, C. S., and J. O. Reiss. 1993. Metamorphosis and the vertebrate skull: ontogenetic patterns and developmental mechanisms. Pp. 289–346 in J. Hanken and B. K. Hall, eds. The skull volume 1: development. Univ. of Chicago Press, Chicago, IL.

Sanger, T. J., D. L. Mahler, A. Abzhanov, and J. B. Losos. 2012. Roles for modularity and constraint in the evolution of cranial diversity among anolis lizards. Evolution 66:1525–1542.

Schlager, S. 2017. Morpho and rvcg - shape analysis in R. Pp. 217–256 *in* G. Zheng, S. Li, and G. Szekely, eds. Statistical shape and deformation analysis. Academic Press, Cambridge, MA.

Schoch, R. R. 2014. Amphibian skull evolution: the developmental and functional context of simplification, bone loss and heterotopy. J. Exp. Zool. Part B Mol. Dev. Evol. 322B:619–630.

Sherratt, E. 2011. Evolution of the caecilian skull. Ph.D. thesis, The University of Manchester, Manchester, U.K.

Sherratt, E., D. J. Gower, C. P. Klingenberg, and M. Wilkinson. 2014. Evolution of cranial shape in caecilians (Amphibia: Gymnophiona). Evol. Biol. 41:528–545.

Sherratt, E., M. Vidal-García, M. Anstis, and J. S. Keogh. 2017. Adult frogs and tadpoles have different macroevolutionary patterns across the Australian continent. Nat. Ecol. Evol. 1:1385–1391.

Simon, M. N., and G. Marroig. 2017. Evolution of a complex phenotype with biphasic ontogeny: contribution of development versus function and climatic variation to skull modularity in toads. Ecol. Evol. 7:10752–10769.

Trueb, L. 1970. Evolutionary relationships of headed tree frogs with coossified skulls (Family Hylidae). Univ. Kansas Publ. Museum Nat. Hist. 18:547–716.

———. 1973. Bones, frogs, and evolution. Pp. 65–132 *in* J. L. Vial, ed. Evolutionary biology of the anurans: contemporary research on major problems. Univ. of Missouri Press, Columbia, MO.

———. 1993. Patterns of cranial diversity among the Lissamphibia. Pp. 255–338 in J. Hanken and B. K. Hall, eds. The skull: patterns of structural and systematic diversity. The Univ. of Chicago Press, Chicago, IL.

Trueb, L., and C. Gans. 1983. Feeding specializations of the Mexican burrowing toad, *Rhinophrynus dorsalis* (Anura: Rhinophrynidae). J. Zool. 199:189–208.

Urošević, A., K. Ljubisavljević, and A. Ivanović. 2018. Multilevel assessment of the lacertid lizard cranial modularity. J. Zool. Syst. Evol. Res. 57:1–14.

Vassilieva, A. B. 2017. Heterochronies in the cranial development of Asian tree frogs (Amphibia: Anura: Rhacophoridae) with different life histories. Dokl. Biol. Sci. 473:110–113.

Vidal-García, M., and J. S. Keogh. 2017. Phylogenetic conservatism in skulls and evolutionary lability in limbs – morphological evolution across an ancient frog radiation is shaped by diet, locomotion and burrowing. BMC Evol. Biol. 17:1–15.

Volume Graphics. 2001. VGStudio MAX version 2.0. Volume Graphics GmbH, Heidelberg, Germany.

Wagner, G. P., and L. Altenberg. 1996. Perspective: complex adaptations and the evolution of evolvability. Evolution 50:967–976.

Watanabe, A., A.-C. Fabre, R. N. Felice, J. Maisano, J. Müller, A. Herrel, and A. Goswami. 2019. Ecomorphological diversification in squamates from conserved pattern of cranial integration. Proc. Natl. Acad. Sci. USA 116:14688–14697.

Weisbecker, V., and C. Mitgutsch. 2010. A large-scale survey of heterochrony in anuran cranial ossification patterns. J. Zool. Syst. Evol. Res. 48:332–347.

Wells, K. D. 2010. The ecology and behavior of amphibians. Univ. of Chicago Press, Chicago.

Wiley, D. F., N. Amenta, G. Alcantara, D. A. Deboshmita, Y. J. Kil, E. Delson, W. Harcourt-Smith, F. J. Rohlf, K. St.John, and B. Hamann. 2005. Evolutionary morphing. Proc. IEEE Vis. 2005:431–438.

Yeh, J. 2002. The effect of miniaturized body size on skeletal morphology in frogs. Evolution 56:628–641.

Ziermann, J. M., and R. Diogo. 2013. Cranial muscle development in frogs with different developmental modes: direct development versus biphasic development. J. Morphol. 275:398–413.

Associate Editor: M. L. Zelditch

Supporting Information

Additional supporting information may be found online in the Supporting Information section at the end of the article.

Figure S1. Landmark and semilandmark data displayed on specimen.
Figure S2. Morphospace of the first two phylogenetic principal components for anuran crania.
Figure S3. Network graphs displaying results from EMMLi and CR analyses.
Figure S4. Network graphs displaying results from landmark-only EMMLi and CR analyses.
Figure S5. Network graph for EMMLi analysis excluding specimens with absent regions.
Figure S6. Effect of life history on trait integration.
Table S1. Specimen information table legend- see additional document for table.
Table S2. Region definitions.
Table S3. Landmark definitions.
Table S4. Curve definitions.
Table S5. Number of surface points within each cranial region.
Table S6. Centroid size.
Table S7. Results of EMMLi analysis using uncorrected data.
Table S8. Results of EMMLi analysis using phylogenetically-corrected data.
Table S9. Results of EMMLi analysis using allometry-corrected data.
Table S10. Results of EMMLi analysis using subsampled data.
Table S11. Results of EMMLi analysis using uncorrected landmark-only data.
Table S12. Results of EMMLi analysis using phylogenetically-corrected landmark-only data.
Table S13. Results of EMMLi analysis using allometry-corrected landmark-only data.
Table S14. Results of phylogenetic EMMLi analysis excluding specimens with absent regions.
Table S15. Covariance Ratio results using uncorrected data.
Table S16. Covariance Ratio results using phylogenetically-corrected data.
Table S17. Covariance Ratio results using allometry-corrected data.
Table S18. Covariance Ratio results using uncorrected landmark-only data.
Table S19. Covariance Ratio results using phylogenetically-corrected landmark-only data.
Table S20. Covariance Ratio results using allometry-corrected landmark-only data.
Table S21. Results of phylogenetic EMMLi analysis for species possessing a feeding larval stage.
Table S22. Results of subsampled phylogenetic EMMLi analysis for species possessing a feeding larval stage.
Table S23. Results of phylogenetic EMMLi analysis for species lacking a feeding larval stage.
Table S24. Results of phylogenetic CR analysis for species possessing a feeding larval stage.
Table S25. Results of subsampled phylogenetic CR analysis for species possessing a feeding larval stage.
Table S26. Results of phylogenetic CR analysis for species lacking a feeding larval stage.