Developmental origins of mosaic evolution in the avian cranium

Ryan N. Felice[a,b,1] and Anjali Goswami[a,b,c]

[a]Department of Genetics, Evolution, and Environment, University College London, London WC1E 6BT, United Kingdom; [b]Department of Life Sciences, The Natural History Museum, London SW7 5DB, United Kingdom; and [c]Department of Earth Sciences, University College London, London WC1E 6BT, United Kingdom

Edited by Neil H. Shubin, The University of Chicago, Chicago, IL, and approved December 1, 2017 (received for review September 18, 2017)

Mosaic evolution, which results from multiple influences shaping morphological traits and can lead to the presence of a mixture of ancestral and derived characteristics, has been frequently invoked in describing evolutionary patterns in birds. Mosaicism implies the hierarchical organization of organismal traits into semiautonomous subsets, or modules, which reflect differential genetic and developmental origins. Here, we analyze mosaic evolution in the avian skull using high-dimensional 3D surface morphometric data across a broad phylogenetic sample encompassing nearly all extant families. We find that the avian cranium is highly modular, consisting of seven independently evolving anatomical regions. The face and cranial vault evolve faster than other regions, showing several bursts of rapid evolution. Other modules evolve more slowly following an early burst. Both the evolutionary rate and disparity of skull modules are associated with their developmental origin, with regions derived from the anterior mandibular-stream cranial neural crest or from multiple embryonic cell populations evolving most quickly and into a greater variety of forms. Strong integration of traits is also associated with low evolutionary rate and low disparity. Individual clades are characterized by disparate evolutionary rates among cranial regions. For example, Psittaciformes (parrots) exhibit high evolutionary rates throughout the skull, but their close relatives, Falconiformes, exhibit rapid evolution in only the rostrum. Our dense sampling of cranial shape variation demonstrates that the bird skull has evolved in a mosaic fashion reflecting the developmental origins of cranial regions, with a semi-independent tempo and mode of evolution across phenotypic modules facilitating this hyperdiverse evolutionary radiation.

evolutionary rates | evolutionary development | modularity | morphological diversity | morphometrics

The term "mosaic evolution" was coined following the discovery of body fossils of *Archaeopteryx*, the iconic early bird, which fascinated researchers with its combination of ancestral "reptilian" and derived avian features (1, 2). Mosaic evolution has since become a central part of understanding avian origins and diversification (3–6), but it is rarely quantified (7). Mosaic evolution is the result of traits evolving at different rates or with different modes. For example, the diversification of locomotor behaviors in birds is thought to be related to the evolutionary independence of the forelimb and hind limb (3, 4, 8). Strongly correlated traits are expected to have a coordinated response to selection, whereas dissociated traits can evolve independently. These relationships among traits are governed by genetic, developmental, and functional associations that can form semiautonomous modules.

The tetrapod skull exhibits developmental modularity: the face is primarily derived from cranial neural crest (CNC) cells, whereas the braincase has a primarily mesodermal origin. This dichotomy is the basis of many investigations of phenotypic modularity in the skull: skeletal elements sharing a common embryonic origin can be predicted to covary with each other more than with components of different origins (9, 10). Finer-scale studies of craniofacial development have uncovered examples of phenotypic modularity in the skull regulated by the expression of a small number of genes. For example, manipulating the expression of *Fgf8* generates correlated responses in the growth of the premaxilla and palatine in archosaurs (11). Similarly, variation in avian beak shape and size is regulated by two separate developmental modules (7). Despite the evidence for developmental modularity in the avian skull, some studies have concluded that the cranium is highly integrated (i.e., not subdivided into semiautonomous modules) (9, 10, 12). In light of recent evidence that diversity in beak morphology may not be shaped by dietary factors (12), it is especially critical to investigate other factors that shape the evolution of cranial variation, such as developmental interactions/constraint. Here, we evaluate hypotheses of cranial modularity using a high-dimensional geometric morphometric dataset of unprecedented resolution (757 3D landmarks) that comprehensively describes cranial shape and broad sampling across Neornithes (352 species).

A key question in the study of phenotypic modularity is how trait interactions influence macroevolutionary change (13–15). For example, high integration is frequently hypothesized to constrain disparity and evolutionary rate by limiting axes of variation upon which selection can act and thus limiting the direction or magnitude of response to selection (13, 14). This has been supported with empirical data from cranial modules in primates and carnivorans in which there is an overall positive relationship between integration and constraint (14). In contrast, in the felid axial skeleton high integration has the opposite effect,

Significance

Studies reconstructing morphological evolution have long relied on simple representations of organismal form or on limited sampling of species, hindering a comprehensive understanding of the factors shaping biological diversity. Here, we combine high-resolution 3D quantification of skull shape with dense taxonomic sampling across a major vertebrate clade, birds, to demonstrate that the avian skull is formed of multiple semi-independent regions that epitomize mosaic evolution, with cranial regions and major lineages evolving with distinct rates and modes. We further show that the evolvability of different cranial regions reflects their disparate embryonic origins. Finally, we present a hypothetical reconstruction of the ancestral bird skull using this high-resolution shape data to generate a detailed estimate of extinct forms in the absence of well-preserved three-dimensional fossils.

Author contributions: R.N.F. and A.G. designed research, performed research, analyzed data, and wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

Data deposition: Scan data used in this study will be available for download from Phenome10K.org, subject to copyright rules of the respective repositories.

See Commentary on page 448.

[1]To whom correspondence should be addressed. Email: ryanfelice@gmail.com.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1716437115/-/DCSupplemental.

promoting greater response to selection (16). Analysis of simulated data has suggested that integration does not necessarily influence evolutionary rate or disparity but rather the direction of response to selection (13). We reconstruct macroevolutionary dynamics across modules in the present dataset to further disentangle the complicated relationships between integration, development, and diversification.

Results

Using a likelihood-based approach for comparing hypotheses of modularity (17) and covariance ratio analysis (18), we find support for the avian skull as highly modular and consisting of seven anatomical modules (Fig. 1 *A* and *B*), similar to patterns of organization observed in mammals (17, 19). This contrasts with previous analyses that have found the avian cranium to be highly integrated. However, these previous analyses have been restricted to analyzing smaller clades (10, 12) or have excluded large portions of the skull (9).

We computed the rate of evolution in each module directly from the landmark configurations (20) using a recently published dated phylogeny for birds (21). The rostrum, palate, naris, and cranial vault evolve at the fastest rates ($\sigma^2_{mult} = 2.98 \times 10^{-7}$–$3.82 \times 10^{-7}$, *SI Appendix*, Table S1). The occipital and pterygoid-quadrate evolve approximately one-third as fast ($\sigma^2_{mult} = 1.15 \times 10^{-7}$–$1.23 \times 10^{-7}$), and the basisphenoid exhibits the lowest rate ($\sigma^2_{mult} = 0.72 \times 10^{-7}$). Overall, there is an inverse relationship between rate and within-module integration (Fig. 2*A*). Similarly, we observe an inverse relationship between disparity (Procrustes variance) and integration (Fig. 2*B*). These patterns are associated with differences in developmental complexity: the fast-evolving modules develop from anterior mandibular-stream CNC cells only (nares) or with contributions from multiple embryonic cell types including the anterior mandibular-stream CNC cells (rostrum, palate, vault), whereas the more slowly evolving regions arise from either the posterior mandibular CNC only (pterygoid-quadrate) or mesoderm only (occipital and basisphenoid) (22). Together, these findings suggest that strong within-module trait correlations, related to developmental origin, constrain both the rate of evolution and the potential for the evolution of high diversity in some avian cranial modules, as in mammals (23).

For each anatomical module, we conducted evolutionary model fitting on the principal component axes that explain 95% of the within-module variation (*SI Appendix*, Table S1) to detect shifts in evolutionary rates across the Neornithes. Brownian motion with a lambda tree transformation is the most likely model for the whole skull and for each individual module. High lambda values (mean $\lambda = 0.83$–0.94) emphasize that phenotype is strongly influenced by shared ancestry. Multivariate phylogenetic signal (24) is lower for each module ($K_{mult} = 0.48$–0.66, $P = 0.001$, *SI Appendix*, Table S1) but still indicates a significant effect of phylogeny on morphology. Tracing evolutionary rate through time illustrates that each module experienced high rates early in the diversification of birds, leading up to and corresponding with the Cretaceous–Paleogene (K–Pg) boundary (Fig. 1*C*). This pattern is concordant with recent reconstructions of high net species diversification rates in the latest Cretaceous (25) and with the estimated time of the origination of major clades including Neoaves, Strisores, Inopaves, Aequorlitornithes, and Columbaves (21). This peak is likely to be a result of the origin of major phenotypic innovations early in the diversification of these clades. Phenotypic evolutionary rates are very low immediately following the K-Pg mass extinction. Perhaps representing a signal of recovery from the extinction event, there is a small peak centered around 60 Mya, coinciding with the origin of several orders including Sphenisciformes, Coliiformes, Musophagiformes, and Accipitrifomes (21). Fossil evidence indicates that these clades had begun to acquire key phenotypic differences during this interval (26, 27). In the rostrum, vault, palate, and pterygoid-quadrate, rates

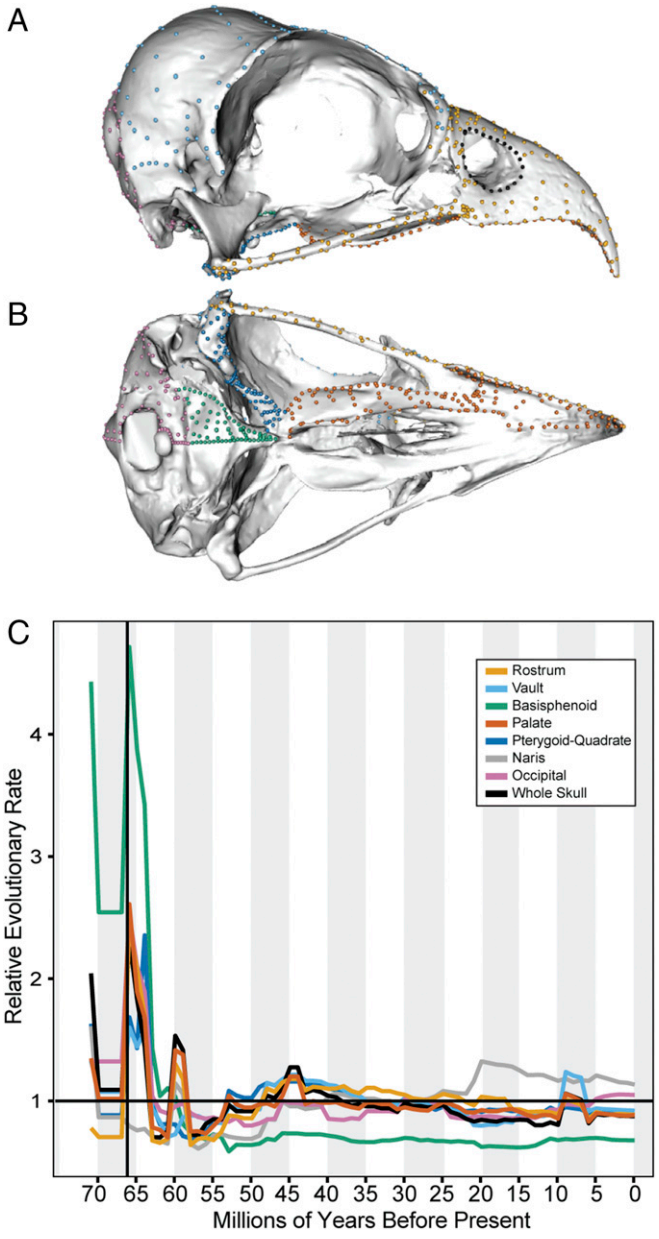

Fig. 1. Cranial modularity in the avian skull. Cranial morphology was quantified using 757 3D landmarks, here illustrated on *Pandion haliaetus* (USNM 623422) in lateral (*A*) and ventral (*B*) perspective. Landmark colors reflect the reconstructed pattern of seven cranial modules. (*C*) Rate through time plots for each of the seven modules and the whole skull.

accelerate gradually through the Eocene, peaking at 45 Mya. This was a major period of cladogenesis: Charadriiformes, Anseriformes, Galliformes, and Strigiformes all originated in this interval. Finally, the rostrum, vault, and palate modules exhibit a peak in rates between 5 and 10 Mya. This is likely to represent recent intrafamilial and intrageneric divergences.

We further examined the patterns of rate shifts throughout the tree to understand the tempo and mode of evolution in each module across different lineages. For each module, rapid evolutionary change tends to occur (*i*) at the origin of major clades, (*ii*) throughout the evolution of diverse clades, and/or (*iii*) in association with the acquisition of novel phenotypes relative to the broader sample. These patterns are typified by the rostrum module. Falconiformes, Anatidae, Strisores, and Pelecaniformes

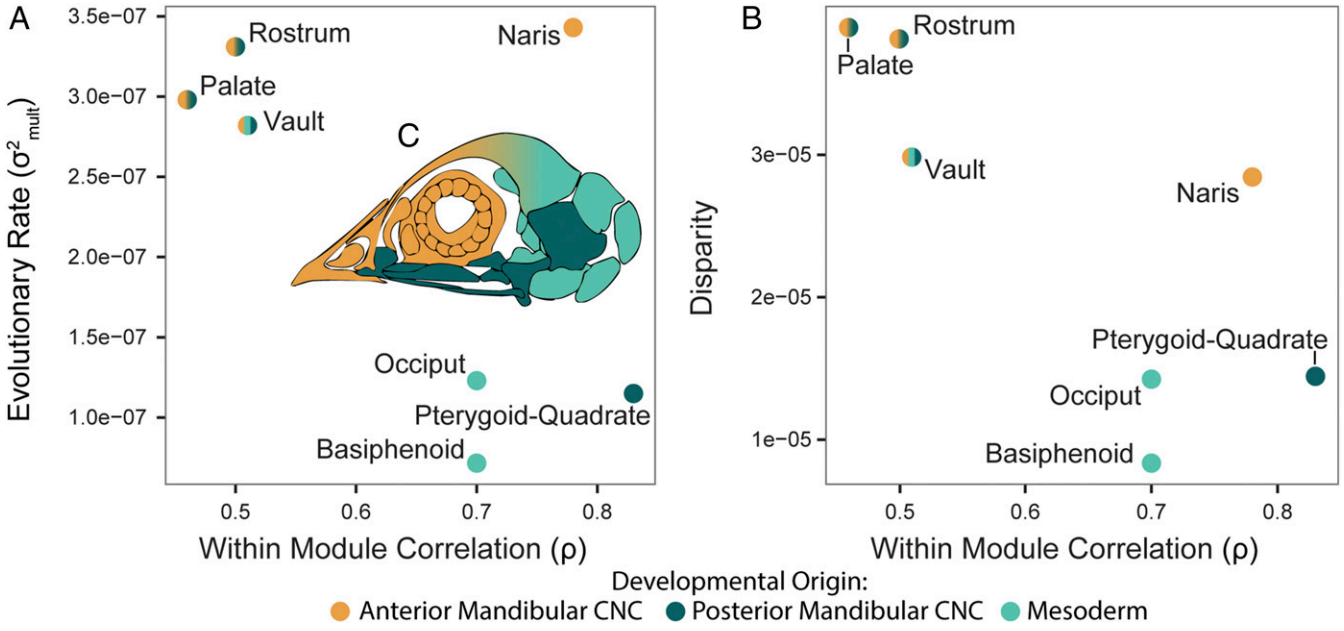

Fig. 2. Module integration and evolution reflects developmental origin. Evolutionary rate vs. within-module correlation (*A*) and disparity vs. within-module correlation (*B*). Disparity is quantified for each module using Procrustes variance divided by the number of landmarks. Disparity and rate are highest in modules that are composed of anterior mandibular CNC or multiple embryonic cell types. Embryonic origin of cranial elements is shown in *C*, modified from ref. 22.

exhibit bursts of rapid rostrum evolution at their origins (Fig. 3 and *SI Appendix*, Fig. S1). Core Passeroidea (28) exhibits an early burst with sustained elevated rates relative to other passeriforms. Within Coraciimorphae, the highest estimated rates of rostrum evolution are exhibited by hornbills (Bucerotidae) and toucans (Ramphastidae), which are frequently cited as classic examples of Old World and New World phenotypic convergence for their long, broad bills (29). Lineages with unique beak phenotypes (relative to their parent clade in the present sample) and rapid rostrum evolution include Pelecanidae, Recurvirostridae, Phoenicopteridae, *Campylorhamphus*, *Psarocolius*, *Rostratula*, and *Nyctibius* (*SI Appendix*, Fig. S1). Overall patterns of rostrum-shape evolution are comparable to those reconstructed in a recent study of bill evolution (30). As in the rostrum, the bill is characterized by bursts of "quantum evolution" at the origin of major clades, as well as by rapid evolution on branches leading to the first occurrence of a unique phenotype within the given sample (30).

This pattern of early bursts and quantum evolution alongside acceleration in some unusual lineages is also observed in other modules. Psittaciformes, Bucerotiformes, and Passeroidea show very similar patterns in the palate and rostrum (Fig. 4*A* and *SI Appendix*, Fig. S2). The high ancestral rate at the origin of Psittaciformes is presumably associated with the evolution of the characteristic vertically oriented palatine (31). Rates of evolution in the cranial vault are less variable, with bursts of rapid evolution at the origin of Strigiformes, Strisores, and Trochilidae (Fig. 4*B*). The highest rates of cranial vault evolution are seen in genera with cranial ornaments (e.g., *Casuarius*, *Numida*, *Balearica*, *Bucorvus*), suggesting that display structures may evolve particularly quickly (*SI Appendix*, Fig. S3) (32).

Each of the other cranial modules also exhibit distinct and heterogenous patterns of evolutionary rate (*SI Appendix*, Figs. S4–S8). The occipital region is characterized by sustained evolutionary rates throughout major clades, including high rates in Passeri and Phasianidae (Fig. 4*C*). The most notable rate shift in the pterygoid-quadrate module includes early bursts at the origin of Strisores, Gallonserae, and the most recent common ancestor of Strigiformes and Piciformes and sustained high rates in Psittaciformes (*SI Appendix*, Fig. S5). In the naris module, patterns are largely defined by clades with divergent naris morphology or position such as Bucerotidae and Ramphastidae, which both have extremely posteriorly positioned nares (*SI Appendix*, Fig. S6). Finally, the basisphenoid module has the slowest overall rate of evolution observed across the skull, but quantum evolution is observed at the origin of several groups including Strisores, Aequorlitornithes, and Trochilidae (*SI Appendix*, Fig. S7) This module most typifies an early burst, with the highest rates in the Late Cretaceous and lowest rates in the Cenozoic (Fig. 1*C*).

The earliest crown bird fossils include representatives of Galloanserae (33), total-clade Sphenisciformes (26), and total-clade Coliiformes (27). However, these specimens are known from primarily postcranial remains or highly deformed specimens, hindering estimation of the cranial phenotype for the ancestral neornithine. An advantage of high-dimensional geometric morphometrics is that it allows for the visualization of hypothetical phenotypes (34). We reconstructed the ancestral state for each module to generate a composite hypothesis of the earliest neornithine using a likelihood-based approach (35). Ancestral values were calculated from the Procrustes-aligned right-hand landmarks. This ancestral landmark configuration was projected into principal component morphospace along with the empirically derived specimens to find the species that it most closely resembles: *Vanga curvirostris*. The 3D mesh of *V. curvirostris* (*SI Appendix*, Fig. S9) was warped (36) to match the ancestral character state, generating the reconstruction shown in Fig. 3 (*Inset*). The reconstructed skull (Fig. 3) has a gently curved beak that is approximately equal in length to the rest of the skull and elongate choanae [schizognathous palate (31)]. In addition to *V. curvirostris*, overall skull shape is similar to that of *Oriolus oriolus*. Both are omnivorous passeriforms that are aerial and canopy foragers and range in mass from 64 to 79 g (37). This reconstruction serves as a first attempt at visualizing a hypothesis of avian origins using high-dimensional data and as a testable model of ancestral phenotype and potential feeding ecology. The accuracy of this reconstruction will be greatly improved by incorporation of data from extinct species (38), whether through discoveries of fossils that are well preserved in 3D or retrodeformation of existing fossils of both early crown and stem birds.

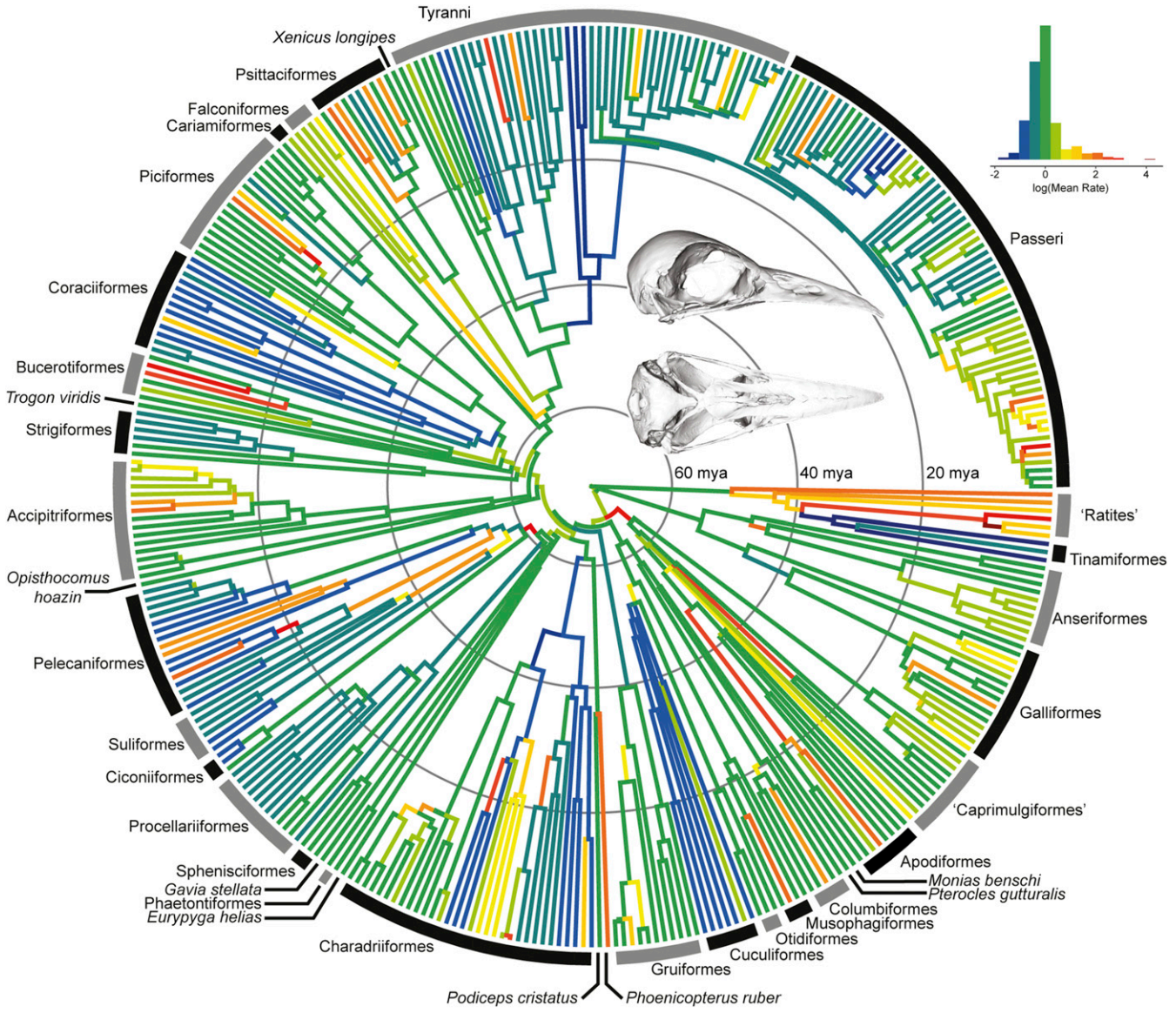

Fig. 3. Evolutionary rates in the avian rostrum. Estimated using BayesTraitsV3 using a variable-rates model and lambda tree transformation. (*Inset*) Reconstruction of the ancestral neornithine skull (*Materials and Methods*).

Discussion

Using high-dimensional morphometrics, we find that the avian skull is more modular than previously known (9, 10, 12). Weak correlations among anatomical modules allowed each to evolve relatively independently of the others, generating a pattern of mosaicism that characterizes several aspects of avian evolution (3–6). For each module, evolutionary rates vary greatly across clades. Morphologically diverse clades, including Strisores, Aequorlitornithes (waterbirds and shorebirds), and Passeroidea, have elevated evolutionary rates at their origin in multiple modules. In contrast, Columbaves and Coraciiformes tend to evolve slowly in all modules. As in the avian bill, high rates of evolution are associated with the origin of divergent or unique phenotypes among the sampled taxa (30) including elongate or curved premaxillae (e.g., Phoenicopteridae), cranial ornaments (e.g., *Bucorvus*), or distinctive palates (e.g., Psittaciformes).

The finding that high integration is associated with low evolutionary rates and low disparity supports the long-standing hypothesis that strong correlations among traits constrain evolutionary change. A similar pattern has been demonstrated in the mammal cranium (14). However, this pattern may not represent a general property of phenotypic integration as the strength of within-module correlation is positively correlated with disparity in other systems (16). Rather, the macroevolutionary consequences of integration are likely to be highly dependent on the direction and magnitude of selection on each module (13). Complementary analyses of other clades and anatomical regions will aid in expanding our understanding of the link between integration and constraint. Interestingly, we find a major difference in developmental complexity between modules with high and low integration: weakly integrated modules arise from multiple embryonic cell populations whereas strongly integrated modules are composed of just one. This mirrors the finding that, in mammals, anatomical modules with high developmental complexity have low within-module integration (23). Additional study of the functional and developmental constraints on cranial evolution, incorporating data from the brain and soft tissues, will allow for further evaluation of the mechanistic underpinnings of these patterns. For example, the morphogenic primacy of the brain in

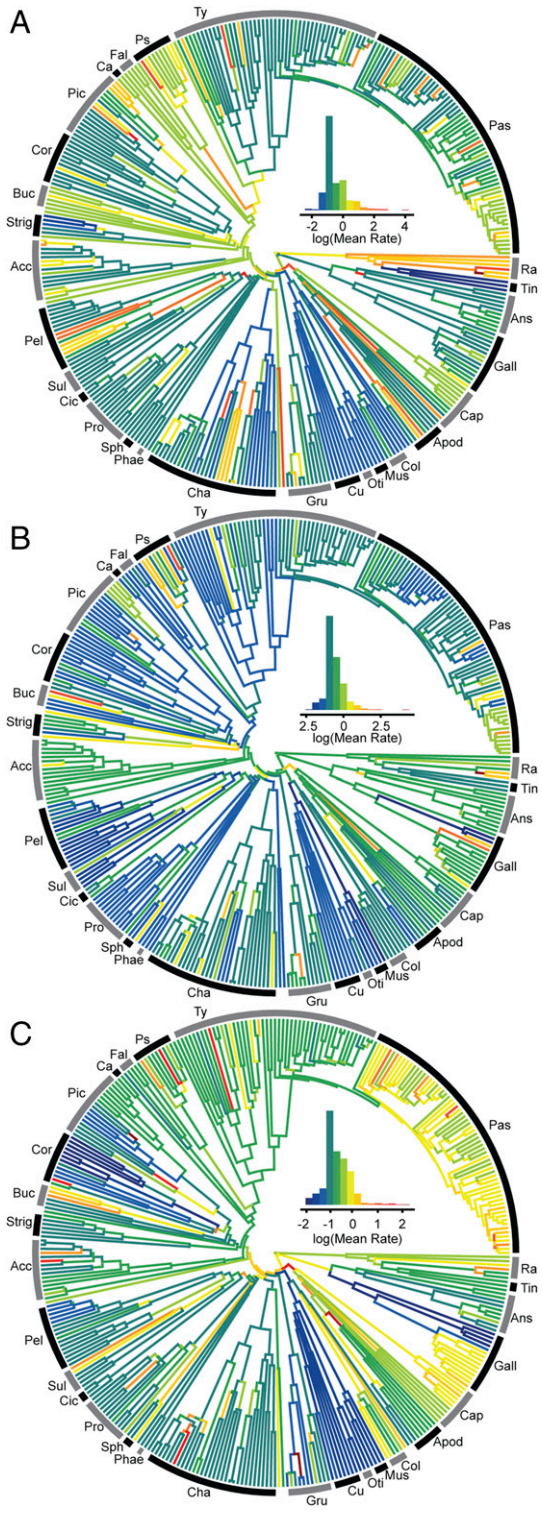

Fig. 4. Evolutionary rates across cranial modules. Palate (*A*), cranial vault (*B*), and occiput (*C*). Clade abbreviations: Acc, Accipitriformes; Ans, Anseriformes; Apod, Apodiformes; Buc, Bucerotiformes; Ca, Cariamiformes; Cap, Caprimulgiformes; Cha, Charadriiformes; Cic, Ciconiiformes; Col, Columbiformes; Cor, Coraciiformes; Cu, Cuculifomes; Fal, Falconiformes; Gall, Galliformes; Gru, Gruiformes; Mus, Musophagiformes; Oti, Otidiformes; Pas, Passeri; Pel, Pelecaniformes; Phae, Phaethontiformes; Pic, Piciformes; Pro, Procellariiformes; Ps, Psittaciformes; Ra, Ratites; Sph, Sphenisciformes; Strig, Strigiformes; Sul, Suliformes; Tin, Tinamiformes; Ty, Tyranni.

the development of the cranium (39) may contribute to the relatively limited variation in evolutionary rates observed in the vault in this study. Taken together, these results support the fundamental prediction that macroevolutionary patterns are shaped by phenotypic modularity and integration, which are in turn determined by developmental processes (7, 11, 13, 16, 23).

From *Archaeopteryx* to crown birds, mosaic evolution has been a vital part of avian diversification. Mosaic evolution is possible because the cranium, like the brain (6) and postcranium (2–4), is composed of modular subunits that evolve at different rates in different lineages. Our high-density sampling of morphology and phylogeny allowed for identification of complex patterns of evolution that gave rise to the extraordinary diversity of living birds. Our results show that individual regions of the skull experienced independent rapid bursts of evolution at different times during the early adaptive radiation of Neornithes and later at the origin of some clades. We further demonstrate that structures derived from multiple embryonic cell types or from anterior mandibular CNC evolve at faster rates than those originating from only posterior mandibular CNC or mesoderm, highlighting the importance of intrinsic factors in shaping the evolution of biodiversity.

Materials and Methods

Data Collection. Skull morphology was characterized in 352 species of neornithine birds representing 320 genera and 159 families of the 238 families recognized by the International Ornithological Congress (*SI Appendix*, Data File S1). One specimen was digitized per species using laser surface scanning (FARO EDGE Scan Arm HD) or microCT scanning (SkyScan1172). For sexually dimorphic taxa, the sex with smaller or absent ornaments was selected for scanning. Because this study concerns skeletal evolution, crania were scanned without the rhamphotheca. Digital models were landmarked with 36 anatomical landmarks (26 bilateral, 10 midline) and 23 sliding semilandmark curves composed of a total of 335 landmarks (*SI Appendix*, Data File S2). Landmark placement was carried out using IDAV Landmark (34), and landmarking was conducted by a single investigator to avoid multiuser bias in placement.

Because some museum specimens are partially damaged or incomplete, semilandmark curves were digitized on the right side of the specimen and midline only. To generate an extremely detailed characterization of the entire surface of the skull, we then used a semiautomated procedure to distribute surface sliding semilandmarks across the skull. First, landmarks and semilandmark curves were placed on a simple hemispherical template (*SI Appendix*, Fig. S10*B*). Surface semilandmark points were placed on this template on the surface of the regions corresponding to the rostrum, jugal bar, cranial vault, occipital bones, basisphenoid, palate, and the ventral surfaces of the quadrate and pterygoid. We then used the R package Morpho (40) to apply the surface semilandmarks from the template to each specimen, generating a dataset of 770 landmarks per specimen (*SI Appendix*, Fig. S10*C*). Because the structure of the face, anterior orbit, and naris of ratites is substantially different from other birds, a separate rostrum template was used to ensure proper placement of the rostrum-surface semilandmarks to the five ratite specimens. After application of surface semilandmarks, all surface landmarks were slid to reduce bending energy (41). Analysis of unilateral landmarks on bilaterally symmetrical structures can introduce undesirable error during superimposition (42). To mitigate these effects, all right-side semilandmarks were mirrored to the left side, and then specimens were subjected to a Procrustes alignment (43). The mirrored left-side landmarks and semilandmarks were then deleted to reduce dimensionality of the data, resulting in a final dataset of 757 aligned landmarks. See *SI Appendix*, *SI Methods*, for additional considerations when dealing with high-dimensional data such as these.

Phylogenetic Trees. We utilized a recent hypothesis of the phylogenetic relationships of birds based on next-generation sequencing data (21) for all comparative analyses. To generate a phylogeny containing all measured taxa for this study, we generated a composite topology following the procedure in ref. 30. This incorporates the backbone of relationships among major clades from ref. 21 along with the fine-scale species relationships from a maximum clade credibility tree generated from www.birdtree.org (44). The tree of 9,993 species was then pruned to match the 352 taxa in our dataset. Finally, our dataset contains the recently extinct *Xenicus longipes*, which is not present in the published molecular phylogenies. We substituted *X. longipes* at the position of *Xenicus gilviventri*.

Modularity Analysis. We used a likelihood-based approach (17) to evaluate the degree to which the avian skull is structured as a set of interrelated

anatomical modules. EMMLi allows between- and within-module correlations to be calculated based on user-defined models of modular organization and then evaluates the likelihood of each model. We evaluated 16 different hypotheses of the structure of modularity in the landmark configurations, ranging from 2 to 13 modules. The model with 13 modules had the highest likelihood. However, in examining the observed correlations between modules, a clear pattern becomes evident (*SI Appendix*, Table S2*A*). The correlation between the dorsomedial and ventrolateral margins of the naris are extremely high ($\rho = 0.73$), justifying the binning of these regions into a single module. Similarly, the pairwise correlations between all regions of both the pterygoid and quadrate are very high ($\rho = 0.54$–0.95). We therefore combined the landmarks and semilandmarks on each of these two elements into a single pterygoid-quadrate module. This resulted in a seven-module hypothesis of the organization of the skull, composed of the rostrum (dorsal surface of the premaxilla, nasal, jugal bar), cranial vault (frontal, parietal, and squamosal; anatomical terminology from ref. 39), occipital (supraoccipital, paraoccipital, basioccipital), basisphenoid, pterygoid-quadrate (ventral surface or pterygoid and articular surface of quadrate), palate (ventral surface of premaxilla and palatine), and naris (the perimeter of the external naris). Each subsequent analysis was carried out on each of the seven modules individually and also on the whole skull configuration. To test the effects of shared ancestry on trait correlations, we also carried out EMMLi analysis on the phylogenetic independent contrasts (45) of the landmark configurations. This analysis supported the same most likely model and relative strength of correlations within and among modules as the analysis of the raw data did.

We also evaluated the seven-module hypothesis (Fig. 1) by calculating covariance ratios (CR) between all pairs of modules (18). There is significant modularity between all pairs of modules ($P = 0.001$, *SI Appendix*, Table S3). CR is highest between rostrum and palate modules (CR = 0.99); however, this value represents significant modularity compared with the distribution of CR calculated from 1,000 simulations.

Phylogenetic Signal. Phylogenetic signal was calculated for each module using the K_{mult} statistic, a method specifically designed for the challenges of working with high-dimensional landmark configurations (24).

Evolutionary Rates. The rate of evolution in each module was analyzed with BayesTraitsV3 (www.evolution.rdg.ac.uk/), using principal-component scores as the input data (Figs. 3 and 4 and *SI Appendix*, Figs. S1–S8). We also calculated the multivariate rate of evolution (σ^2_{mult}) directly from the landmark data (46). See *SI Appendix, SI Methods*, for additional detail.

ACKNOWLEDGMENTS. We thank Judith White, Christine Lefevre, Chris Milensky, Steve Rogers, Ben Marks, Janet Hinshaw, Paul Sweet, Lydia Garetano, Kristof Zyskowski, and Greg Watkins-Colwell for facilitating data collection; Anne-Claire Fabre, Henry Ferguson-Gow, and Léo Botton-Divet for methodological assistance; Catherine Early for CT scanning; and Aki Watanabe, Carla Bardua, Thomas Halliday, Marcela Randau, Eve Noirault, and three anonymous reviewers for providing helpful feedback and comments. This research was funded by European Research Council Grants STG-2014-637171 (to A.G.) and SYNTHESYS FR-TAF-5635 (to R.N.F.).

1. Owen R (1863) On the *Archaeopteryx* of Von Meyer, with a description of the fossil remains of a long-tailed species from the lithographic stone of Solnhofen. *Philos Trans R Soc Lond* 153:33–47.
2. de Beer GR (1954) Archaeopteryx and evolution. *Adv Sci* 11:160–170.
3. Clarke JA, Middleton KM (2008) Mosaicism, modules, and the evolution of birds: Results from a Bayesian approach to the study of morphological evolution using discrete character data. *Syst Biol* 57:185–201.
4. Gatesy SM, Dial KP (1996) Locomotor modules and the evolution of avian flight. *Evolution* 50:331–340.
5. Cracraft J (1970) Mandible of *Archaeopteryx* provides an example of mosaic evolution. *Nature* 226:1268.
6. Balanoff AM, Smaers JB, Turner AH (2016) Brain modularity across the theropod-bird transition: Testing the influence of flight on neuroanatomical variation. *J Anat* 229: 204–214.
7. Mallarino R, et al. (2011) Two developmental modules establish 3D beak-shape variation in Darwin's finches. *Proc Natl Acad Sci USA* 108:4057–4062.
8. Abourachid A, Höfling E (2012) The legs: A key to bird evolutionary success. *J Ornithol* 153:193–198.
9. Klingenberg CP, Marugán-Lobón J (2013) Evolutionary covariation in geometric morphometric data: Analyzing integration, modularity, and allometry in a phylogenetic context. *Syst Biol* 62:591–610.
10. Kulemeyer C, Asbahr K, Gunz P, Frahnert S, Bairlein F (2009) Functional morphology and integration of corvid skulls: A 3D geometric morphometric approach. *Front Zool* 6:2.
11. Bhullar B-AS, et al. (2015) A molecular mechanism for the origin of a key evolutionary innovation, the bird beak and palate, revealed by an integrative approach to major transitions in vertebrate history. *Evolution* 69:1665–1677.
12. Bright JA, Marugán-Lobón J, Cobb SN, Rayfield EJ (2016) The shapes of bird beaks are highly controlled by nondietary factors. *Proc Natl Acad Sci USA* 113:5352–5357.
13. Goswami A, Smaers JB, Soligo C, Polly PD (2014) The macroevolutionary consequences of phenotypic integration: From development to deep time. *Philos Trans R Soc Lond B Biol Sci* 369:20130254.
14. Goswami A, Polly PD (2010) The influence of modularity on cranial morphological disparity in carnivora and primates (Mammalia). *PLoS One* 5:e9517–e9518.
15. Marroig G, Shirai LT, Porto A, Oliveira FB, Conto V (2009) The evolution of modularity in the mammalian skull II: Evolutionary consequences. *Evol Biol* 36:136–148.
16. Randau M, Goswami A (2017) Unravelling intravertebral integration, modularity and disparity in Felidae (Mammalia). *Evol Dev* 19:85–95.
17. Goswami A, Finarelli JA (2016) EMMLi: A maximum likelihood approach to the analysis of modularity. *Evolution* 70:1622–1637.
18. Adams DC (2016) Evaluating modularity in morphometric data: Challenges with the RV coefficient and a new test measure. *Methods Ecol Evol* 7:565–572.
19. Parr WCH, et al. (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.
20. Adams DC (2014) Quantifying and comparing phylogenetic evolutionary rates for shape and other high-dimensional phenotypic data. *Syst Biol* 63:166–177.
21. Prum RO, et al. (2015) A comprehensive phylogeny of birds (Aves) using targeted next-generation DNA sequencing. *Nature* 526:569–573.
22. Maddin HC, Piekarski N, Sefton EM, Hanken J (2016) Homology of the cranial vault in birds: New insights based on embryonic fate-mapping and character analysis. *R Soc Open Sci* 3:160356.
23. Goswami A (2006) Cranial modularity shifts during mammalian evolution. *Am Nat* 168:270–280.
24. Adams DC (2014) A generalized K statistic for estimating phylogenetic signal from shape and other high-dimensional multivariate data. *Syst Biol* 63:685–697.
25. Claramunt S, Cracraft J (2015) A new time tree reveals Earth history's imprint on the evolution of modern birds. *Sci Adv* 1:e1501005.
26. Slack KE, et al. (2006) Early penguin fossils, plus mitochondrial genomes, calibrate avian evolution. *Mol Biol Evol* 23:1144–1155.
27. Ksepka DT, Stidham TA, Williamson TE (2017) Early Paleocene landbird supports rapid phylogenetic and morphological diversification of crown birds after the K-Pg mass extinction. *Proc Natl Acad Sci USA* 114:8047–8052.
28. Barker FK, Barrowclough GF, Groth JG (2002) A phylogenetic hypothesis for passerine birds: Taxonomic and biogeographic implications of an analysis of nuclear DNA sequence data. *Proc Biol Sci* 269:295–308.
29. Seki Y, Bodde SG, Meyers MA (2010) Toucan and hornbill beaks: A comparative study. *Acta Biomater* 6:331–343.
30. Cooney CR, et al. (2017) Mega-evolutionary dynamics of the adaptive radiation of birds. *Nature* 542:344–347.
31. Zusi RL, Livezey BC (2006) Variation in the os palatinum and its structural relation to the palatum osseum of birds (Aves). *Ann Carnegie Mus* 75:137–180.
32. Cuervo JJ, Moller AP (1999) Evolutionary rates of secondary sexual and non-sexual characters among birds. *Evol Ecol* 13:283–303.
33. Clarke JA, Tambussi CP, Noriega JI, Erickson GM, Ketcham RA (2005) Definitive fossil evidence for the extant avian radiation in the Cretaceous. *Nature* 433:305–308.
34. Wiley DF, et al. (2005) Evolutionary morphing. *Visualization 2005* (IEEE, Washington, DC), pp 431–438.
35. Goolsby EW (2017) Rapid maximum likelihood ancestral state reconstruction of continuous characters: A rerooting-free algorithm. *Ecol Evol* 7:2791–2797.
36. Adams DC, Otárola-Castillo E (2013) geomorph: An R package for the collection and analysis of geometric morphometric shape data. *Methods Ecol Evol* 4:393–399.
37. Wilman H, et al. (2014) EltonTraits 1.0: Species-level foraging attributes of the world's birds and mammals. *Ecology* 95:2027.
38. Finarelli JA, Flynn JJ (2006) Ancestral state reconstruction of body size in the Caniformia (Carnivora, Mammalia): The effects of incorporating data from the fossil record. *Syst Biol* 55:301–313.
39. Fabbri M, et al. (2017) The skull roof tracks the brain during the evolution and development of reptiles including birds. *Nat Ecol Evol* 1:1543–1550.
40. Schlager S (2017) Morpho and Rvcg–Shape analysis in R. *Statistical Shape and Deformation Analysis*, eds Zheng G, Li S, Székely GJ (Academic, London), pp 217–256.
41. Gunz P, Mitteroecker P, Bookstein F, Slice D (2005) Semilandmarks in three dimensions. *Modern Morphometrics in Physical Anthropology*, ed Slice DE (Springer, New York), pp 73–98.
42. Cardini A (2017) 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.
43. Rohlf FJ, Slice D (1990) Extensions of the procrustes method for the optimal superimposition of landmarks. *Syst Zool* 39:40–59.
44. Jetz W, Thomas GH, Joy JB, Hartmann K, Mooers AO (2012) The global diversity of birds in space and time. *Nature* 491:444–448.
45. Felsenstein J (1985) Phylogenies and the comparative method. *Am Nat* 125:1–15.
46. Adams DC (2013) Comparing evolutionary rates for different phenotypic traits on a phylogeny using likelihood. *Syst Biol* 62:181–192.