UC Berkeley
UC Berkeley Previously Published Works

Title

A switch in jaw form–function coupling during the evolution of mammals

Permalink

Journal

ISSN

Authors

Tseng, Z Jack
Garcia-Lara, Sergio
Flynn, John J
et al.

Publication Date

DOI

Peer reviewed

Author for correspondence:
Z. Jack Tseng
e-mail: zjt@berkeley.edu

A switch in jaw form–function coupling during the evolution of mammals

Z. Jack Tseng[1,2,3], Sergio Garcia-Lara[1,2], John J. Flynn[3,4], Emily Holmes[1], Timothy B. Rowe[5] and Blake V. Dickson[6]

[1]Department of Integrative Biology, and [2]Museum of Paleontology, University of California, Berkeley, CA 94720, USA
[3]Division of Paleontology, and [4]Richard Gilder Graduate School, American Museum of Natural History, New York, NY 10024, USA
[5]Jackson School of Geological Sciences, University of Texas, Austin, TX 78712, USA
[6]Department of Evolutionary Anthropology, Duke University, Durham, NC 27708, USA

ZJT, 0000-0001-5335-4230

The evolutionary shift from a single-element ear, multi-element jaw to a multi-element ear, single-element jaw during the transition to crown mammals marks one of the most dramatic structural transformations in vertebrates. Research on this transformation has focused on mammalian middle-ear evolution, but a mandible comprising only the dentary is equally emblematic of this evolutionary radiation. Here, we show that the remarkably diverse jaw shapes of crown mammals are coupled with surprisingly stereotyped jaw stiffness. This strength-based morphofunctional regime has a genetic basis and allowed mammalian jaws to effectively resist deformation as they radiated into highly disparate forms with markedly distinct diets. The main functional consequences for the mandible of decoupling hearing and mastication were a trade-off between higher jaw stiffness versus decreased mechanical efficiency and speed compared with non-mammals. This fundamental and consequential shift in jaw form–function underpins the ecological and taxonomic diversification of crown mammals.

This article is part of the theme issue 'The mammalian skull: development, structure and function'.

1. Introduction

The radiation of crown group mammals into adaptive zones vacated by the extinction of non-avian dinosaurs led to the Cenozoic Era (66 Ma to the present) being regarded as the 'Age of Mammals'. Among the most dramatic macroevolutionary transformations leading up to the Cenozoic diversification of the *ca* 6500 living species of crown mammals plus their extinct relatives are two novel osteological features: (i) the addition of two new middle-ear ossicles, elements from postdentary bones homologous to the articular and quadrate in non-mammal vertebrates, to the one element common to all tetrapods (stapes), and (ii) the detachment of postdentary bones from the lower jaw and the related evolution of the temporo-mandibular joint (TMJ, or the squamosal–dentary joint) [1,2]. This second feature is an unreversed mammalian synapomorphy and accomplished the decoupling of hearing acuity on the one hand, and a single-element jaw bone with more precise dental occlusion on the other [3–8]. Study of such morphological transformation and repurposing has fuelled palaeontological research for nearly two centuries, and these shifts are understood to be crucial in enabling mammals to take fuller advantage of a broad range of food sources as active heterotrophs in broad-ranging terrestrial, marine and aerial environments [9,10].

Although past research focused heavily on the implications of this transition for the mammalian ear, this evolutionary transformation also affected jaw mechanics. Historically, we understand vertebrate feeding systems by characterizing the morphological components as lever arms and force transmission linkages (class III
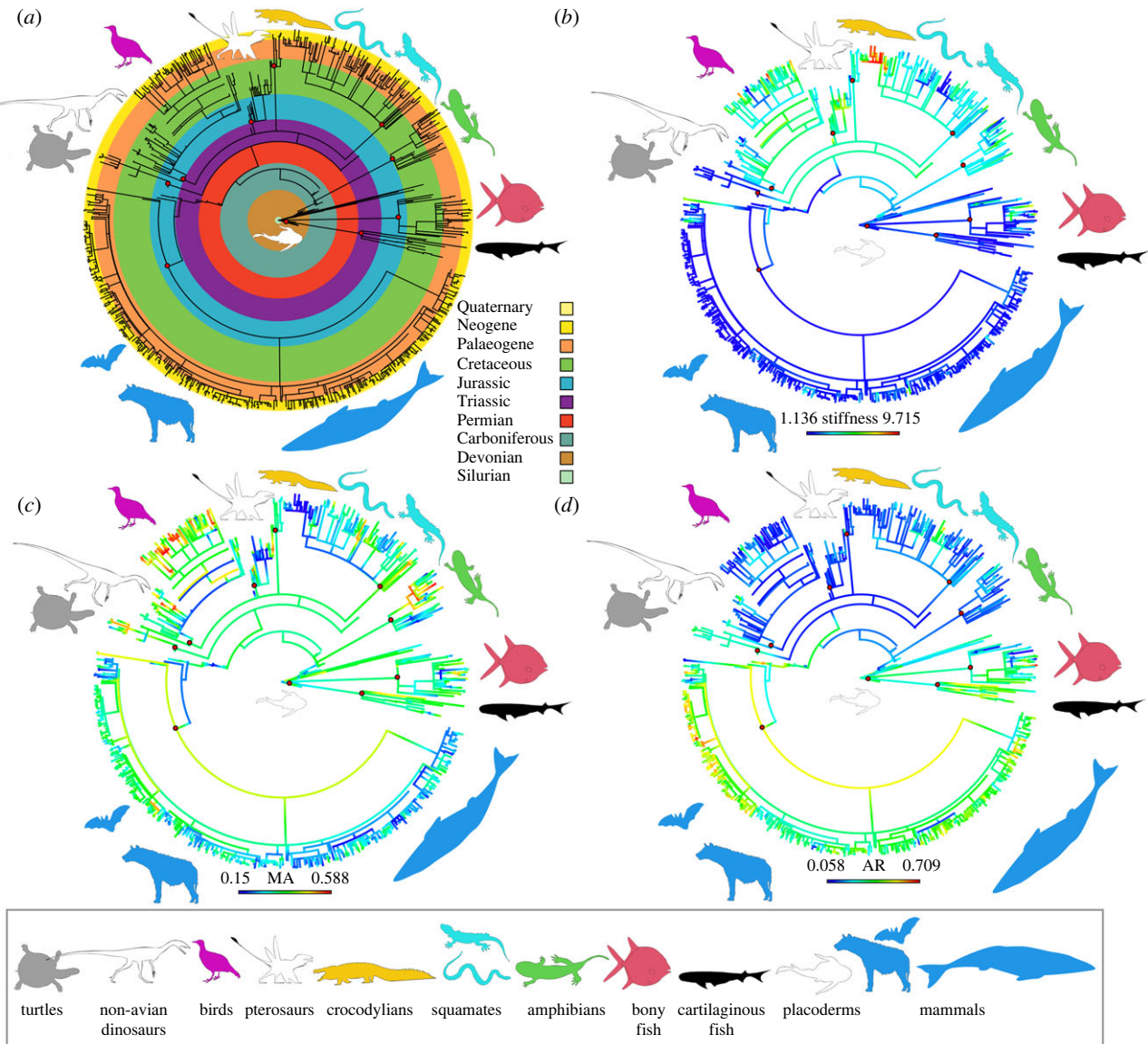
Figure 1. Vertebrate lower jaw shape dataset mapped over a composite phylogeny. (*a*) Temporal and taxonomic distribution of taxa sampled. (*b*) Stiffness (maximum deflection in cantilever bending experiment) with ancestral state reconstruction. Lower deflection values imply higher stiffness. (*c*) Mechanical advantage (MA) heatmap with ancestral state reconstructions. Higher MA values imply more mechanically efficient jaws. (*d*) Aspect ratio (AR) heatmap with ancestral state reconstructions. Lower AR values imply more streamlined and higher velocity transmission jaw shapes. Silhouettes from phylopic.org (see electronic supplementary material, data S21 for attributions).

levers and four-bar linkage systems, for example). In this framework, the fulcrum of the jaw mechanical system (jaw joint) underwent three major transformations during the approximately 440 Myr history of jawed vertebrates. First, the evolution of an upper–lower jaw articulation in the earliest gnathostomes [11] enabled the feeding apparatus to diversify force and speed parameters in a mechanical lever system [12]. The second reorganization occurred in stem Osteichthyes [13], from a quadrate–infragnathal joint to a quadrate–articular joint [14]; 90% of living vertebrates (approx. 58 500 species) still possess this class III lever anatomical configuration, attesting to its evolutionary viability and longevity. The third major transformation, from the quadrate–articular joint to a TMJ in mammals [15], further reduced the number of lower jaw elements and intramandibular joints present in this force transmission system [2]. In recent decades, research on evolutionary transformations of the mammalian jaw has centred on exceptionally preserved fossils documenting the decoupling of hearing and mastication [16–20]; illuminating the biomechanical and kinematic consequences of this important transformation in the loading of the jaw joint and the movements of the now single-boned lower jaw in mammals is a major initiative [4,21,22].

The functional and biomechanical implications of this evolutionary transformation have largely been hypothesized as increased mechanical efficiency and strength with the redistribution of joint loading forces and more advantageous in–out lever configurations for jaw adductor muscles. No broad quantitative test of this scenario surrounding the origin of crown Mammalia has yet been performed, however.

Here, we demonstrate that shape evolution of the mammalian mandible is characterized by a concomitant increase in morphological disparity but decrease in functional disparity, representing a form–function decoupling. Furthermore, the mammalian jaw transformation represents a trade-off between strength and efficiency; the radiation of mammalian mandibular shape diversity is coupled with higher jaw stiffness but lower jaw speed and mechanical advantage (MA) compared with non-mammals. We leverage an integrative array of analyses using elliptic Fourier transforms, experimental mechanical testing, finite-element modelling, gene–phenotype–biomechanics association analyses, adaptive landscapes and phylogenetic comparative methods on a jaw shape dataset of 1000+ vertebrates (479 mammaliaform genera and 563 non-mammal vertebrate genera; figure 1, electronic
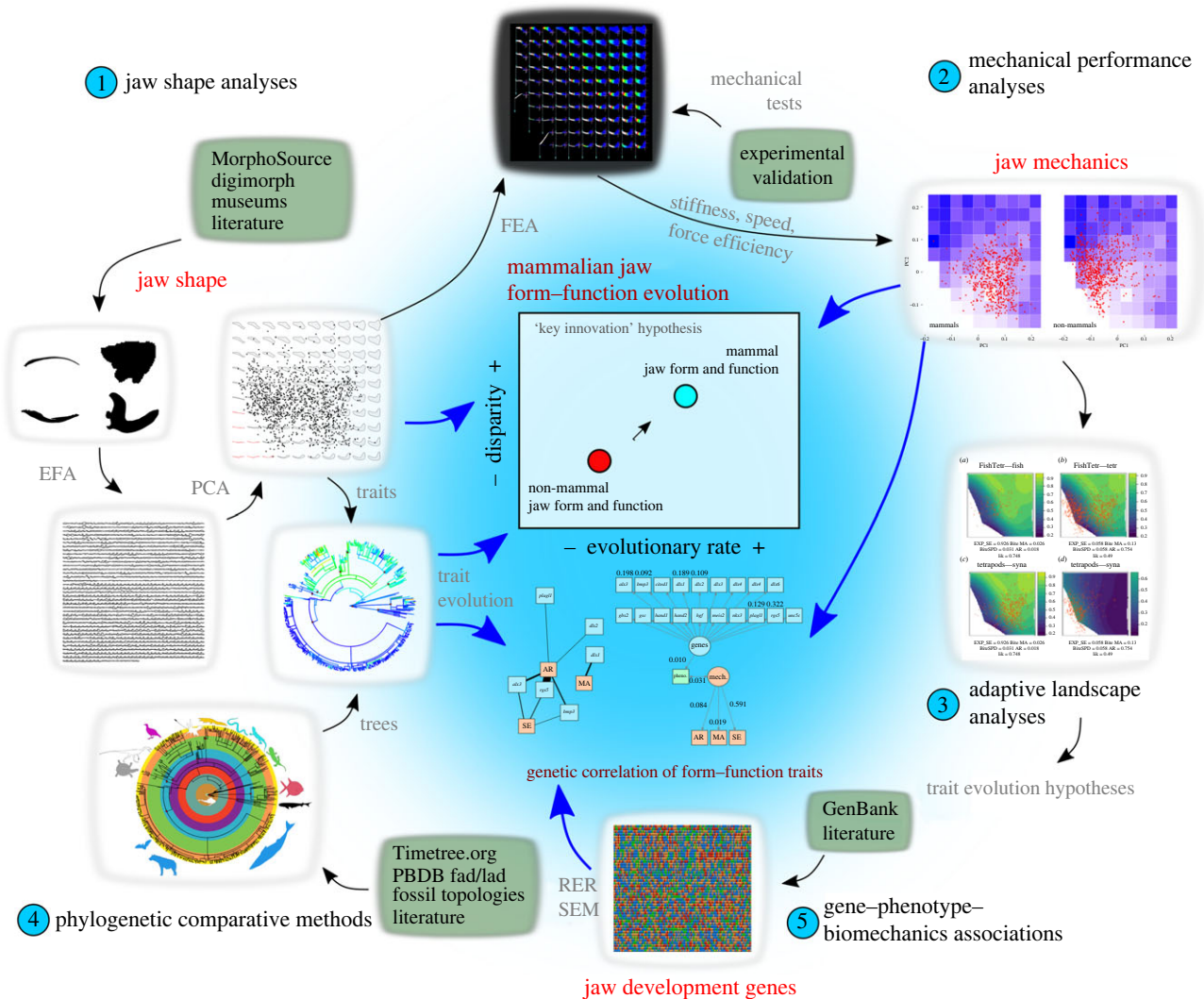
Figure 2. The integrated methodological framework used in this study. The major components of the analyses include (1) jaw shape analyses using elliptical Fourier analysis (EFA), principal components analysis (PCA) and morphospace visualization; (2) mechanical performance analyses using finite-element analysis (FEA) and cantilever bending tests on theoretical warp models, interpolation of biomechanical traits in individual genera; (3) adaptive landscape analyses that optimize the relative fitness coefficient among biomechanical traits; (4) phylogenetic comparative methods that test fit to trait evolution models using interpolated biomechanical traits and a new composite phylogeny for both extant and fossil vertebrate taxa; (5) relative evolutionary rates (RER) and path analyses using structural equation modelling (SEM) of genotype–phenotype–biomechanics correlations in rates of evolution and in overall similarity.

supplementary material, figure S1 and data S1, S2). The mammal dataset includes representatives of all major crown and key stem mammal clades; the non-mammal dataset includes a sampling of fossil and living non-tetrapod vertebrates, amphibians, squamates, archosaurs, turtles and non-mammalian synapsids. We constructed an 'all-mandible' morphospace using the largest 2D profile shape dataset of jawed vertebrates to date and used it to test a 'key innovation' hypothesis of a major mammalian functional morphological evolutionary rate and disparity shift coinciding with a single-element lower jaw. The form–function analyses are grounded in the first principles of jaw lever mechanics [23] and the fundamental property of force–speed trade-off in quasi-static lever systems [24].

2. Material and methods

(a) Mandible shape analysis

We used 2D outline shape analysis to quantify vertebrate jaw shape [12,25–30] (figure 2). To capture a range of jaw shapes across vertebrates, we compiled mandible specimen image files (see electronic supplementary material for detailed sources) and converted them into binary images. The image dataset was imported into the R programming environment for elliptical Fourier analysis using the *Momocs* R package [31]. A series of smoothing and standardization operations ensured the shapes represented outlines of comparable semi-landmark density (see electronic supplementary material for additional details). A full generalized Procrustes alignment was performed. The aligned shape dataset was then subjected to elliptic Fourier transformation with calibrated harmonics search to obtain 99% harmonic power. A principal components analysis (PCA) was then conducted on the transformed data and the first two PC axes used to create a two-dimensional jaw shape morphospace. Within the PC1–PC2 jaw shape morphospace we generated 100 warped theoretical jaw shapes characterizing different regions of the morphospace. Warped jaw shapes were set at regular intervals to cover the X (10 intervals) and Y (10 intervals) axes using the plotting function in R. The resulting warped shapes (hereafter called 'theoretical warp models') were exported as vector graphics files both for linear biomechanical trait measurements and for 3D bending experiments and finite element (FE) simulations (see below).

(b) Biomechanical trait estimation and validation

Two linear biomechanical traits based on beam theory and lever mechanics principles were measured for each of the 100 theoretical warp models from the morphospace: aspect ratio (AR) and MA. Mechanical lever ratios have been shown to overestimate force transmission and underestimate velocity transmission [32]. Therefore, we determined both MA and AR as separate measures of force and velocity, respectively. Jaw AR is defined as the length of the longest line drawn across the dorsal–ventral axis of each jaw shape outline divided by that of the longest line drawn across the anterior–posterior axis of the jaw shape. A lower AR is interpreted as representative of higher jaw displacement or compliance, and higher AR is an indication of low jaw displacement or compliance. Jaw velocity is estimated as 1/AR; jaws that are elongate will have higher velocity than jaws that are less elongate. MA is defined here as the length of the line between the edge of the outline at the jaw joint to the dorsal-most point of the coronoid/adductor process divided by the length of the line between the jaw joint and the anterior-most point of the jaw shape outline. This ratio measures the relative in-lever arm length of the generalized adductor muscle insertion site at the coronoid process to the out-lever arm length of a hypothetical bite point at the front of the jaw (equivalent to an incisor tooth position in mammals). Both AR and MA values are unit-less and size-free. Measurements used in calculating these ratios were taken on the exported vector graphics file of the 100 theoretical warp models using the linear measurement tool in Fiji [33].

We estimated additional biomechanical traits using FE simulation and validated our simulation dataset using experimental mechanical tests. The 100 theoretical warp models were converted from vector graphics files into extruded (with a thickness of 10 mm) models that were then 3D printed for experimental bending test and digitally solid-meshed for FE simulations, respectively. For experimental bending tests, hypothetical warp models were printed and mounted onto a custom-fitted vice grip attachment to a Mark-10 ESM1500 electromechanical testing frame (Mark-10, New York, USA). Cantilever bending tests were performed on each plastic jaw model by fixing the jaw joint/grip cube and applying a ventrally directed force at the anterior tip of the model. Two mechanical property values were collected from each 3D printed model experiment: total deflection (up to 10 mm) and strain energy (SE). Deflection is a measure of the stiffness of a particular jaw shape and is measured as the vertical distance travelled by the testing frame during a given experiment. SE is another measurement of stiffness or resistance to deformation, defined as the work done in deforming a material under a particular load, and was estimated as the area under the force–deflection curve by calculating the integral of the curve (in N mm).

For simulation-based bending tests using the FE method, the extruded jaw shape models (in STL format) were imported into Strand7 finite-element software (v.2.4.6, Strand7, Sydney, Australia) and set to have identical dimensions to the physical models tested using bending experiments. All jaw models were assigned a homogeneous set of material properties suitable for linear static analysis; the elastic (Young's) modulus was set to 20 GPa, and Poisson's ratio to 0.3, as in previous FE studies of jaw structures [12,34]. Nodal constraints preventing free body movement were applied in each jaw model at both the posterior–dorsal and the posterior–ventral extremes of the jaw shape. A nodal force of 100 N was applied to the anterior-most node on the edge of each jaw shape model to simulate a ventrally directed load as in a cantilever bending experiment. The FE models were solved with a linear static solver using the direct sparse scheme option. After solution, total deflection (in mm) was extracted from each jaw model at the point of the nodal force. Total stored SE (in joules) was obtained from the solution file using the SE visualization option.

(c) Biomechanical trait interpolation

We estimated taxon-specific values of AR, MA, total deflection (stiffness) and SE (another measure of stiffness) using double linear interpolation, implemented in the *interp* R package. Double linear (or bilinear) interpolation sequentially applies linear interpolation to two variables (x, y) in predicting a third (i.e. z). The biomechanical values obtained from linear measurement, experimental and FE analyses were set as the z values for the theoretical warp models and used to interpolate z values for per-species biomechanical trait in the taxon-specific shape dataset. The resulting mean interpolated values for each genus were used in all subsequent analyses.

(d) Jaw shape and biomechanical trait distribution

To evaluate the assumption that the simplified lateral/2D profile jaw outlines used in this study retain significant functional morphological signal relative to a null model of randomly distributed shape and biomechanical traits, we conducted a series of non-parametric trait distribution analyses using a combination of bootstrap resampling and Wilcoxon rank sum tests. We first generated 1000 datasets of 1000 combinations of PC1 and PC2 values within the jaw shape morphospace as simulated taxa. The biomechanical traits (MA, AR, total deflection, SE) were then estimated for simulated taxa using the same interpolation formulae as those used to estimate actual taxa values. Lastly, Wilcoxon rank sum tests of similarity between the distributions of both shape and biomechanical trait values in the simulated shape dataset versus values interpolated for actual taxon shapes were conducted. The hypothesis tested here is that both the PC1–PC2 and biomechanical trait values in the taxonomic dataset collected were significantly different from randomly sampled distributions of those values in the same morphospace, and thus provide support for the 2D approach.

(e) Trait disparity analyses

To examine potential differences in jaw shape and biomechanics trait disparity between mammals and non-mammals predicted by a 'key innovation' hypothesis (figure 2), we calculated trait disparity using the *dispRity* R package [35]. In addition, to account for potential biases in disparity metrics introduced by uneven taxonomic sampling and sample sizes, each of the disparity comparisons was both bootstrapped and rarefied. Disparity values were compared between mammals and non-mammals for five morphological and biomechanical traits: (i) PC1 and PC2 values only (reflecting the morphological data used to build the jaw shape morphospace), (ii) all PC values from the PCA, (iii) taxon-specific interpolated MA values, (iv) taxon-specific interpolated AR values, and (v) taxon-specific interpolated, experimentally derived maximum deflection values.

(f) Body mass allometry of biomechanical and morphological traits

To assess the potential correlation between the morphological and biomechanical traits analysed and body mass, we collected both body mass and jaw adductor muscle mass data from a range of mammals and other vertebrates. Average adult body masses for 411 mammal genera were extracted from the PanTHERIA database [36] (electronic supplementary material, data S19). In addition, total adductor muscle masses (wet weight) were compiled for selected vertebrates for which dissection and comparative muscle anatomical data were available in the literature [37–43] (electronic supplementary material, data

S20). All mass variables were $\log_{10}$-transformed prior to further analysis. The relationships between each of five functional morphological traits (MA, AR, maximum deflection, PC1 shape score, PC2 shape score) and $\log_{10}$(body mass in grams) were examined using scatterplots. The relationship between $\log_{10}$(body mass) and total jaw adductor was plotted and then linearly regressed using the 'power trend line' function in Microsoft Excel.

(g) Adaptive landscape analysis

We contextualized the evolutionary and biomechanical optimality of vertebrate mandibles using adaptive landscape analysis [44,45]. Briefly, we generated performance surfaces over the main jaw shape morphospace for each biomechanical trait by fitting a second-order polynomial surface using least squares. A combined adaptive landscape was then computed as the summation of individual performance surfaces, each weighted by their relative importance or contribution to overall fitness. In equation form:

$$W = w_1 F_1 + w_2 F_2 + w_3 F_3 + \cdots + w_n F_n,$$

where W is the combined adaptive landscape, w_n is the per-performance surface relative fitness and F_n is the individual performance surface [46]. For each examined clade (see below), the value W is optimized as the likelihood of combinations of performance surface and relative fitness, under the definition that the total fitness sums to 1 and the variance of all surfaces is set to be equal. Using experimental SE, relative bite force (MA), bite speed (1/AR) and jaw displacement (AR) traits as performance variables, adaptive landscapes were optimized for all gnathostomes at five taxonomic levels: non-tetrapods versus tetrapods; between synapsids and sauropsids; and within synapsids, sauropsids and archosaurs, respectively. The optimized adaptive landscapes were then used as hypotheses for testing models of trait evolution.

(h) Testing models of biomechanical trait evolution

To examine evolutionary rates and modes of biomechanical traits (MA, AR, stiffness as measured by SE and maximum deflection), we constructed a composite phylogeny including extant and fossil taxa using existing literature (see electronic supplementary material for additional details). We also estimated branch lengths and their uncertainty over the phylogeny by generating 100 time-tree samples using the Paleobiology Database and the *paleotree* R package [47]. Using these 100 time-tree samples, we evaluated the fit of the trees to different models of trait evolution for MA, AR and stiffness. We evaluated models of trait evolution using the *motmot* R package to test the results from the adaptive landscape analyses: these models are (i) Brownian motion (BM; a random walk model in which trait variance increases with time since divergence at a constant rate), (ii) Ornstein–Uhlenbeck (OU; a BM model with presence of a single optimal trait value or central tendency), and (iii) ACDC ('early burst'; model of exponential increase or decrease in evolutionary rate through time).

Furthermore, to assess whether there is a significant change in the rate of biomechanical evolution from non-mammal vertebrates to mammaliaforms during the transition to the mammalian mandible, we modelled biomechanical evolutionary rate heterogeneity using the *traitMEDUSA* method implemented in *motmot*. This rate shift detection method first fits a BM model to the data, then systematically fits two-rate models to each node in the phylogeny, and then identifies the best-fitting two-rate model using Akaike information criterion (AIC). We used this approach to test the hypothesis of a rate shift occurring near the non-mammal to mammal transition but without specifying an exact node, by allowing the fit to be searched over all possible nodes. Because of model misspecification and other biases associated with using individual multivariate shape trait variables such as individual PC axis scores for univariate comparative analysis [48,49], we conducted a separate evolutionary rate analysis for jaw shape using a distance-based method previously proposed for high-dimensional shape and other biological data [50]. Pairwise between-species differences are used in this approach to estimate multivariate evolutionary rates using a BM model [50]. We used permutation to obtain statistical significance values in the comparison between mammal and non-mammal shape evolution rates.

To select the best-performing biomechanical trait evolution model, we used weights from AICc (AIC corrected for small sample sizes) scores of the three different trait models to identify the preferred model for each of the three biomechanical traits analysed. To identify locations of potential evolutionary rate change in biomechanical traits on the composite phylogeny, we summarized the results of the two-rate model-fitting analyses as a percentage of the 100 tree samples in which a rate increase or decrease is found at a particular node. The same 100 tree samples were used to estimate magnitude and significance of multivariate evolutionary rate differences between mammal and non-mammal jaw shapes.

(i) Relative evolutionary rates of genes implicated in mandible development

We conducted a relative evolutionary rate analysis using GenBank sequences of genes that have been shown to be involved in normal mandible development in vertebrates. Nineteen genes were identified from the literature as definitely expressed in the lower jaw. Reference sequences of these genes were obtained using *Homo sapiens* as the query sequence in the BLAST search engine. A summary of the gene families and groups involved in jaw development is included in the electronic supplementary material.

We first assessed the extent to which the candidate genes to be used in relative evolutionary rate analyses display high degrees of covariation with each other. We used a phenetic approach (neighbour-joining) to build individual gene trees from FASTA sequences. All possible pairwise comparisons of the 19 jaw development candidate genes were then conducted by pruning each pair of gene trees to the same ingroup and outgroup taxa. The trees then were compared using the entanglement algorithm implemented in the *dendextend* R package [51]. We then used the algorithms implemented in the *RERconverge* R package for relative evolutionary rate analysis [52]. Three continuous trait variables from the functional morphological analyses (deflection/SE, MA and AR) for matching taxa in the morphometric dataset were imported, and correlated evolutionary rates between the jaw development genes and each of the continuous trait variables were calculated, followed by statistical testing to identify significant gene–morphology correlations.

(j) Structure equation modelling of gene–phenotype– biomechanics correlations

To examine the correlation between genetic and jaw shape distances among jawed vertebrates, we used structure equation modelling (also known as path analysis or latent variable analysis) to quantify gene sequence covariation with multivariate jaw shape covariation and biomechanical traits. We treated the combined effects of the 19 identified jaw development genes as a genetic latent variable on which the jaw shape, represented by the pairwise Euclidean distance of all principal component shape scores among all taxa for which genetic information is available, depends. We created a second latent variable representing jaw biomechanical performance as measured by AR, MA and SE; this latent variable is correlated to jaw shape in our model, representing structure–function linkage. We fitted

6

royalsocietypublishing.org/journal/rstb   Phil. Trans. R. Soc. B **378**: 20220091

this structural equation model to the genetic, shape and biomechanical data using maximum likelihood in R (*lavaan* package) to estimate path coefficients, which provide measures of the amount of variance explained by individual traits contributing to genetic and biomechanical latent variables, respectively.

3. Results

(a) Mandible shape analysis results

The bivariate plot defined by the first two PC axes (out of a total of 40 PC axes) together represents 59.1% of the total variance in the jaw shape dataset (figure 3*a*). The *x*-axis (PC1) represents a gradient of shallow (dorsoventrally narrow) jaw shapes with low adductor processes on the negative axis to dorsoventrally deep jaws with high adductor (coronoid) processes on the positive axis. The *y*-axis (PC2) mainly represents a gradient of jaw shapes with shallow dorsoventral depth at the adductor process on the negative axis to those with deep dorsoventral depth at the adductor process on the positive axis. The lower left-hand corner of the morphospace (negative PC1 and negative PC2 values) contains a region represented by biologically impossible shapes (those with dorsal and ventral outlines contacting or crossing each other); several taxa possess jaw shapes that border those impossible regions, but none falls within it. Such intersecting outlines are a common occurrence in elliptic Fourier analyses and are treated herein as in other studies as geometrically and biologically impossible configurations and removed from further analysis [12,53]. Data points in the morphospace are clustered around the origin, with a larger spread in the *x*-axis compared with the *y*-axis. Generally, the major gnathostome clades overlap each other, and most clades sampled have representative taxa near the origin. Clades such as pterosaurs and squamates contain clusters of taxa that fall near the edge, representing jaw shapes with high velocity (electronic supplementary material, figure S2).

(b) Biomechanical trait estimation and validation results

MA and AR values obtained from the linear measurements are included in electronic supplementary material, data S6. Experimentally measured total deflection and SE values were collected for 93 out of 100 warped theoretical jaw shape models; seven of the theoretical models were biologically impossible and were not tested. Both deflection and SE values show a similar pattern across the morphospace: the jaw shapes with the lowest deflection (i.e. stiffest) and SE values (i.e. most work-efficient) are in the upper right corner of the morphospace (positive PC1 and PC2 values). Displacement and SE values both increased with decreasing PC1 and PC2 values. Models bordering biologically impossible forms were flexible and did not fall within the parameters of the cantilever bending experiments, and thus did not have displacement or SE values. Given the strong correlation between deflection and SE values, we used the experimentally calculated SE values and maximum deflection values interchangeably to represent overall structural stiffness in adaptive landscape, trait evolution, relative evolutionary rate and latent variable path analyses.

FE simulations of cantilever bending across the 93 theoretical warp models that represent shapes at regular intervals across the all-mandible morphospace returned highly correlated maximum vertical deflection values

between models with intramandibular joints and those without. Outputs of these computational models also correlated strongly to experimentally derived measurements (electronic supplementary material, movie S1). The lower left-hand edge of the morphospace revealed large discrepancies between simulated and experimental data (electronic supplementary material, figure S5). Such discrepancies stemmed from jaw shapes with very low AR (=high jaw closing velocity). Jaw deflection values in experimental data increased exponentially relative to change in AR towards thinner jaws (AR = $0.5987 \times 10^{-0.284 \times \text{Displacement}}$, $R^2 = 0.4728$), reflecting the rapid drop in stiffness of the jaw structure as AR shifts towards higher jaw closing velocity (electronic supplementary material, figure S4).

(c) Jaw shape and biomechanical trait distribution results

Wilcoxon ranked sum tests between bootstrapped null distributions and actual measured (PC1, PC2) or interpolated (MA, AR, stiffness) taxon-specific trait values showed significant differences. All of the taxon-specific attributes are significantly different from a model of random distribution. There are more occurrences of lower MA and lower AR values in sampled taxa than expected from a random distribution of jaw shapes in the morphospace. There are more higher SE and deflection (both indicating lower stiffness) values in sampled taxa compared with randomized jaw shape samples (electronic supplementary material, figure S6). PC1 scores (jaw dorsoventral depth and degree of adductor process development) show a distinct bimodal distribution on either side of the origin (zero value) compared with a random distribution model. On the other hand, PC2 scores exhibit a unimodal distribution around the origin rather than a random distribution. Pairwise Wilcoxon tests for between-clade differences in biomechanical trait values were conducted (electronic supplementary material, data S10); however, we caution against overinterpreting comparisons between clades that have very small sample sizes (e.g. turtles, ornithischian dinosaurs). Instead, we focus our discussion on testing mammal versus non-mammal differences.

(d) Trait disparity results

Mammals exhibit significantly lower disparity in combined PC values compared with non-mammals but have significantly higher disparity specifically in PC1 and PC2 values. In addition, mammals have significantly lower disparity in all three biomechanical traits (MA, AR, total deflection/stiffness) from PC1 and PC2 theoretical warp models compared with non-mammals (electronic supplementary material, figures S14 and S15). Mammals exhibit significantly higher ARs (deeper jaws) and higher stiffness (lower maximum deflection values) than non-mammals, but lower MA compared with non-mammals (table 1).

(e) Body mass allometry results

None of the scatterplots indicates a correlation between body mass and biomechanical and morphological traits; in all analyses the R^2 was smaller than 0.01. Furthermore, linear regression analysis between body mass and total adductor mass suggests that there is no differential allometry for adductor mass to body mass scaling in the mammals
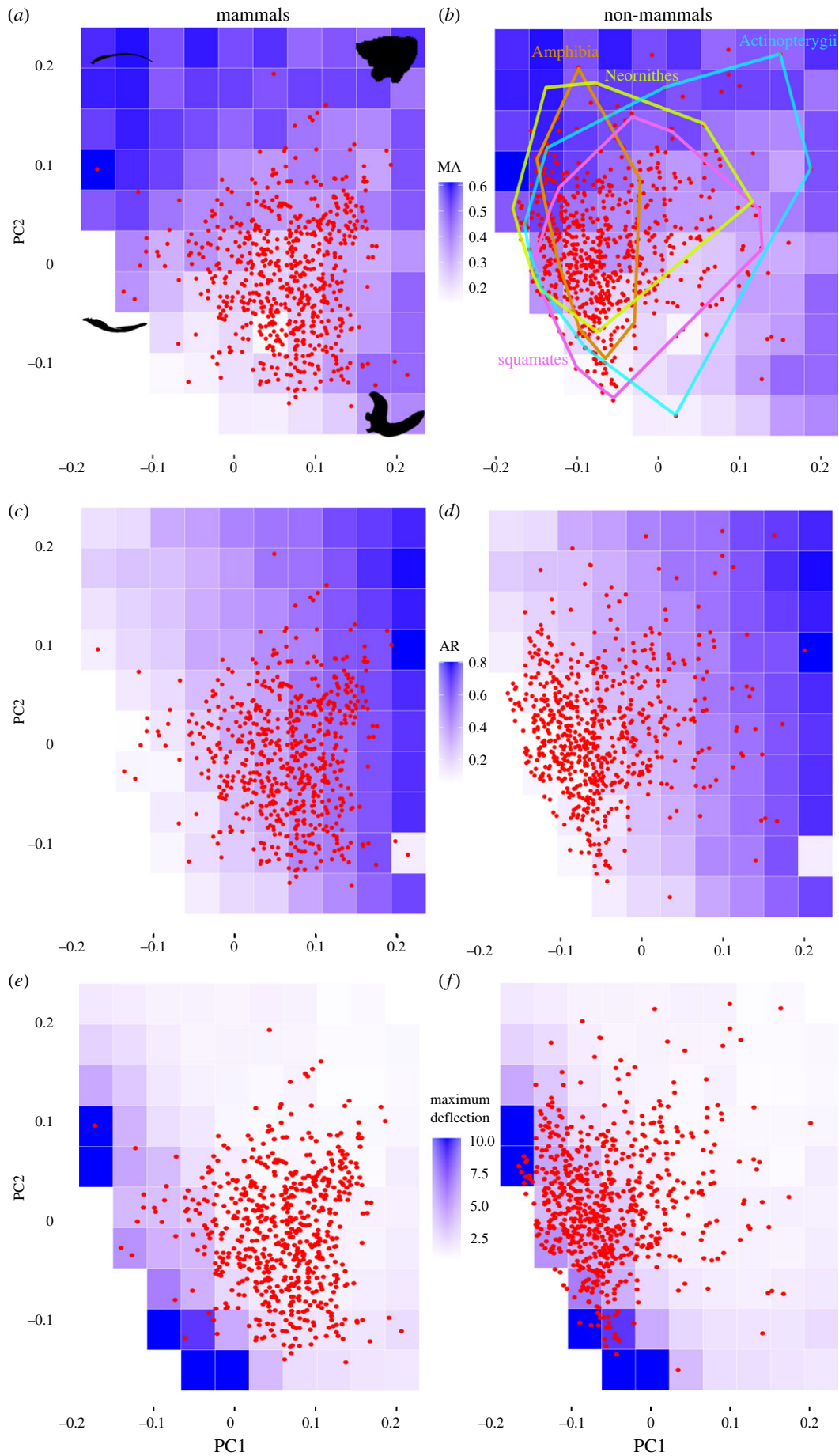
Figure 3. (*Caption overleaf.*)

(57 spp.), birds (48 spp.) or squamates (6 spp.) sampled. The squamate slope is lower (i.e. jaw adductor mass increases slower with body mass) than in either birds or mammals, but the small sample size from the squamate dataset precludes a proper statistical comparison (electronic supplementary material, data S19 and S20).

Figure 3. (*Overleaf*.) All-mandible morphospace and biomechanical traits. (*a*) Distribution of mammals in the morphospace (PC1 38.9%, PC2 20.2% variation) of mechanical advantage (MA) represented by background heatmaps of values calculated from theoretical warp models. Some representative jaw shapes are shown at the corners of the morphospace. (*b*) Distribution of non-mammals in the MA morphospace; higher MA implies more efficient jaw force generation. Representative extant clades are indicated by convex hulls to show the extensive overlap of clade jaw shapes in the morphospace. (*c*) Mammal distribution in jaw aspect ratio morphospace; lower aspect ratio (AR) implies more streamlined and higher velocity jaw shapes. (*d*) Non-mammal distribution in jaw AR morphospace. (*e*) Mammal distribution in jaw stiffness morphospace measured by maximum deflection in cantilever bending experiments; stiffer jaws have lower deflection values. (*f*) Non-mammal distribution in jaw stiffness morphospace. The theoretical model at (0.2, −0.1) returned unrealistic values in some simulations, so all interpolated taxonomic values were calculated after excluding that particular theoretical model from consideration. (Online version in colour.)

Table 1. Summary of biomechanical trait value differences between mammals and non-mammals. MA, mechanical advantage; AR, aspect ratio (higher AR implies lower jaw velocity); ANOVA, one-way analysis of variance. Lower maximum deflection values indicate higher stiffness.

	MA	AR	stiffness (max. deflection)
mammal median	0.28	0.41	1.52
non-mammal median	0.34	0.18	3.40
% (mammal : non-mammal)	83.61%	228.44%	44.71%
p-value, Wilcoxon test	<0.0001	<0.0001	<0.0001
mammal mean	0.29	0.41	1.77
non-mammal mean	0.34	0.22	3.41
% (mammal : non-mammal)	83.38%	188.34%	51.91%
p-value, ANOVA	<0.0001	<0.0001	<0.0001

(f) Adaptive landscape analysis results

Although each taxonomic group returned a unique landscape, pairwise comparisons found significant differences only between synapsids (mammals and their immediate predecessors) and sauropsids (all amniotes more closely related to extant reptiles than to mammals; electronic supplementary material, figure S27 and table S7), and between turtles and all other sauropsids. Our analyses indicated no differences between non-tetrapods and tetrapods, among archosaurs, among non-testudine sauropsids or along the grade of synapsids leading to mammals.

Two primary adaptive regimes were responsible for these differences: (i) a strength regime that is primarily adapted for reducing SE (i.e. increasing stiffness) and only small contributions from bite force, bite speed and jaw compliance (low AR); and (ii) a displacement regime primarily adapted for jaw compliance with moderate contribution from bite force, and small contributions from bite speed and SE. The strength regime provides generally high adaptive performance for a majority of morphospace, to which morphologically diverse groups are well adapted. Non-tetrapod vertebrates, stem tetrapodomorphs, turtles and synapsids are generally broadly distributed through morphospace and are best adapted to the strength regime, while lepidosaurs and archosaurs are generally more restricted in morphospace and are best adapted to the jaw compliance regime (electronic supplementary material, figure S27). Based on these findings, we

hypothesize that an evolutionary rate shift in biomechanical traits occurred in lepidosaurs and archosaurs relative to other vertebrates.

The landscape of stem tetrapodomorphs does not provide any clarity owing to limited sample size; although they are 'best' adapted for the strength regime, they perform poorly on it ($z = 0.51$), and nearly as well on a compliance regime. All synapsids were best adapted to a strength regime, and although sample sizes for non-mammalian synapsids were limited, we did find that pelycosaurs performed poorly, with a large increase in cynodonts and then again in mammals. Given the majority of the vertebrates exhibit a strength regime, we hypothesize that SE/stiffness evolved via an OU-like process of optimization, whereas MA and AR evolution is better explained by a BM process as they were apparently not optimized in most vertebrate groups examined.

(g) Trait evolution mode and rate model fitting results

By contrast to the hypothesized evolutionary modes based on adaptive landscape analyses, the best-fitted trait evolution models for all biomechanical traits (MA, AR, stiffness) across vertebrates are OU models. An OU model is also a better fit than either BM or ACDC when mammal and non-mammal datasets are fitted separately. There is no consistent evidence for evolutionary rate change in biomechanical traits in mammals versus non-mammals. In 68% of trees, laurasiatheres (hedgehogs, bats, hoofed mammals, carnivores) or all crown Mammalia are reconstructed with substantially lower rates of evolution in MA compared with other jawed vertebrates. No clade exhibits substantial increase in MA under a 50% majority rule criterion. In 74% of trees New World monkeys/primates exhibit a two-order increase in AR rate evolution. In 69% of trees Toxicofera (snakes, anguimorph lizards, iguanas) or Serpentes exhibit greater than 80% reduction in AR evolutionary rate. In jaw stiffness, 93% of trees show a substantial slowdown (less than 10% of average rate) in Euarchontoglires (lagomorphs, rodents, tree shrews, primates) and a substantial speed-up (5–10 times average rate) for ornithodirans (dinosaurs and pterosaurs).

(h) Relative evolutionary rate and path analysis results

The calculated entanglement values suggest that the genes *cited1*, *gbx2*, *dlx4*, *meis2* and *rgs5* exhibit the largest differences from other jaw development genes sampled (electronic supplementary material, data S18). By contrast, *dlx2* shared similar relative taxonomic differences to most other genes sampled. The default set of relative evolutionary rate analyses used no transformations, weighting or scaling. The results show that evolutionary rates in four genes (*alx3*, *plagl1*, *bmp3*, *dlx5*) are significantly ($p_{\mathrm{adjusted}} < 0.05$) correlated to jaw AR evolution. Only *dlx1* is significantly correlated to MA evolutionary
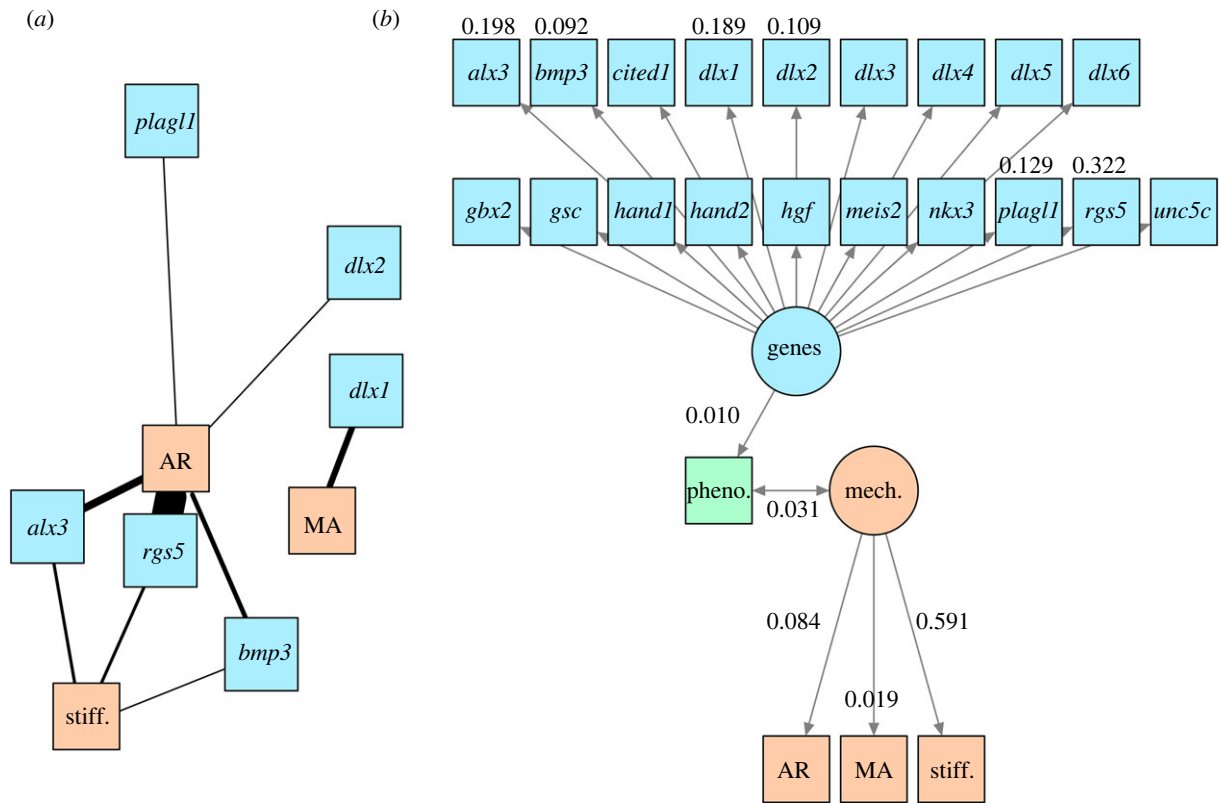
Figure 4. Gene–phenotype–biomechanics association networks. (a) Correlated evolutionary rates between biomechanical (AR, MA, stiffness (measured by maximum deflection in cantilever experiments)) and genes involved in mammalian mandible development portrayed as a network graph; thicker lines indicate higher quantity of significant correlations across sensitivity analyses. (b) Path diagram showing all jaw development genes analysed as components of a latent variable representing genetic effects on phenotype, and the interaction between phenotype and a latent variable representing biomechanics as measured by AR, MA and stiffness. Phenotype is measured by multivariate jaw shape data represented by all 40 principal component axes from the elliptic Fourier analysis. Only path coefficients for the genes found to be significantly correlated with biomechanical evolutionary traits are shown. Gene–phenotype–biomechanics associations are weak but statistically significant. Stiffness explains the largest component of variation in phenotype–biomechanics covariation relative to AR and MA. (Online version in colour.)

rate. Evolutionary rates of three of the analysed genes (*alx3*, *rgs5* and *bmp3*) are significantly correlated to jaw stiffness evolutionary rate (figure 4*a*). A series of sensitivity analyses were conducted using other combinations of transforms, weights and scales; these analyses provided differential support for the significant correlations identified above (see electronic supplementary material, table S6 and figures S17–S21).

Relative evolutionary rates of the analysed genes were plotted over the composite phylogeny (electronic supplementary material, figure S16), and laurasiatheres are reconstructed as having ancestrally accelerated evolutionary rate for *alx3*, whereas euarchontoglires exhibit an ancestrally accelerated *plagl1* evolutionary rate. Both clades are also reconstructed to have relatively low rates of *bmp3* evolutionary rate and no ancestral change in *rgs5* evolutionary rate. No other clade exhibited statistically significant deviations from the overall evolutionary rates. The genes found to be significantly correlated with biomechanical trait evolutionary rates show the same relative contribution to genotype–phenotype correlation in the path analysis (figure 4*b*). There is a significant but weak correlation between phenotype and biomechanics, with stiffness (maximum deflection) contributing the most to this association (path coefficient = 0.59), followed by AR (0.08) and MA (0.02).

4. Discussion

Functional morphological analyses of the vertebrate lower jaw indicate the presence of biomechanical decoupling and

trade-off associated with the morphological diversification of the mammalian single-element mandible. Mammalian and non-mammalian jaws are distinguished by a switch in morphological and biomechanical trait disparity relationships. Bootstrapped and rarefied disparity metrics show statistically significant increases in mandible shape disparity in mammals (electronic supplementary material, figures S14, S15 and data S10). On the other hand, mammalian jaw MA, AR and stiffness were significantly lower in disparity, even after accounting for uneven taxon sampling. Furthermore, mammals exhibit significantly higher jaw stiffness, but lower jaw efficiency (MA) and speed (high AR/low jaw velocity), compared with non-mammals (table 1). This combination of high morphological shape disparity, low biomechanical trait disparity and biomechanical trade-off characterizes many-to-one structure–function mapping, one of the principal manifestations of a redundant mapping of morphology to performance in biological systems. Other prime examples of such complex mapping have been extensively documented in actinopterygian jaw systems [54]. Presence of this many-to-one structure–function relationship in mammals indicates that they independently capitalized on this form–function complex to diversify in jaw shape while maintaining stereotyped, stiffness-optimized jaw biomechanical characteristics.

Appearance of the single-element mandible in mammals did not correspond to a shift in the optimized adaptive landscape regime for jaw biomechanics. Jawed vertebrates collectively are optimized by a strength regime in an adaptive

landscape framework; jaw stiffness is high in both non-tetrapods and tetrapods (electronic supplementary material, figures S13 and S24). Mammals and non-mammalian synapsids maintain this strength regime (electronic supplementary material, table S1), but sauropsids (all tetrapods more closely related to extant reptiles than to mammals) have an optimized adaptive landscape characterized by an AR regime focused more on compliant and faster-closing jaws (electronic supplementary material, figure S24). During their evolution, crown mammals, with their single-element jaw bones, optimized on the existing strength regime within the adaptive landscape by further elevating jaw stiffness (figure 3e,f). However, the lower biomechanical trait disparity associated with the mammalian mandible within this strength regime reinforces the presence of functional constraints despite high shape disparity and stiffness; many differently shaped mammalian jaws exhibit similar mechanical traits. Furthermore, the significant increase in jaw stiffness comes at the expense of lowered jaw efficiency and speed (table 1). Taken together, these findings indicate that the morphological outcome of crown mammalian radiation was an evolutionary exploration of jaw shape space within prescribed bounds of the tetrapod adaptive landscape, characterized by further augmentation and stereotyping of high jaw strength. This mammalian pattern is partially convergent with the earliest gnathostomes, which evolved under a similar regime of form–function decoupling [12]; however, the diverse dental morphologies and environments in which mammals have radiated represent a unique context for interpreting the consequences of such decoupling.

Given the findings of the adaptive landscape analyses, we hypothesized that speed (AR) and efficiency (MA) functions evolved without evolutionary rate shifts during the transition from mammals to non-mammals; we also predicted that jaw stiffness experienced an evolutionary rate shift between mammals and non-mammals. We tested these hypotheses by modelling AR, MA and stiffness trait evolution using a variable rate shift model and found no evidence for evolutionary rate shifts in AR or MA at the transition to mammals. We also did not observe a rate shift in stiffness in the transition to mammals, although the fitted model suggests Euarchontoglires (the crown mammal clades including lagomorphs, rodents, primates, tree shrews) experienced a substantial slowdown in stiffness evolutionary rate, in part reflecting the decreased disparity in stiffness values. These findings indicate that the OU-type jaw stiffness evolution suggested by the adaptive landscape analysis and supported by trait evolution modelling did not couple with modified evolutionary rates of change in jaw stiffness in all mammals; rather, the increase in jaw shape disparity in mammals occurred without modifying the jaw stiffness evolutionary tempo or mode already observed in their non-mammalian relatives. Furthermore, although adaptive landscape analyses optimized stiffness/strength as the primary regime and stiffness data best fit an OU process of evolution, both AR and MA trait evolution fit an OU model as well; this is interpreted as evidence of a biomechanical trade-off where all three traits may be evolving toward their respective functional optima, but strength/stiffness appears to be the principal trait that mammals and non-sauropsid vertebrates in general optimized at the expense of efficiency and speed. This observation suggests stiffness could be a universally optimized functional trait across vertebrate clades and is evidenced by its importance in mechanically demanding feeding tasks as disparate as suction feeding [55] and bone-cracking [56].

Relative gene–biomechanics evolutionary rate analyses demonstrate that the narrower range of biomechanical trait values that distinguish mammalian jaws from non-mammalian jaws is genetically correlated (figure 4; electronic supplementary material, figures S16–S21). In comparison with evolutionary rates of 19 candidate genes known to play key roles in the development of the mandible within the first pharyngeal/branchial arch (BA1), biomechanical trait evolutionary rates are statistically correlated to genetic variation across living mammal species (figure 4a). Stiffness evolutionary rates are significantly correlated with alx3, rgs5 and bmp3. MA (correlated to dlx1) and jaw velocity (correlated to bmp3, plagl1, alx3, rgs5, dlx2) trait values are also found to be genetically linked. Nested expression of the dlx gene family by cranial neural crest cells is associated with normal mandible development and jaw adductor formation, and is also known to regulate the downstream expression of plagl1, alx3, rgs5 and other genes during normal jaw development [57]. The association of these genes to biting-specific biomechanical traits provides evidence of heritability in jaw force, velocity transmission and bending stiffness variation [58,59]. Given this finding, we speculate that these and potentially other candidate genes involved in craniofacial development were targets of selection during the Mesozoic origin of the single-element mandible in mammals.

Path analyses show that the genes with the highest degree of correlated evolutionary rates to biomechanical variables also contribute the highest amount of variance towards the gene–jaw shape covariance, which is statistically significant but weak (figure 4b; path coefficient = 0.01). Among the biomechanical traits, SE explains the most variance in the jaw shape–biomechanics correlation among the three measured traits (SE path coefficient = 0.591 compared with 0.084 for AR and 0.019 for MA). These findings suggest that the low disparity of jaw stiffness in mammals is in part related to correlated rates of biomechanical evolution with jaw development genes, but such correlation is only weakly associated with overall jaw shape. We speculate that genetically encoded structure–function linkages for jaw stiffness are localized in certain regions of the jaw rather than to the overall shape. Future comparative evolutionary developmental studies would provide insights into the expression patterns of these biomechanically relevant candidate genes.

The presence of many-to-one form–function mapping in the mammalian mandible suggests that the single-element lower jaw may be an example of a 'fly in a tube' model of evolution, whereby potential for evolutionary disparity and/or rate shifts within a trait complex (in this case, jaw biomechanics) is constrained in certain regions of a morphospace and shaped by genotypic and/or phenotypic integration [60]. The absence of a significant rate shift or trait evolution model shift in biomechanical traits along the mammalian transition to a dentary–squamosal joint is consistent with expectations of this model; the single-element mammal jaw represents integrated biomechanical trait complexes that traverse a smaller part of the functional space compared with non-mammals ([60]; figure 4b). Such a pattern of high shape disparity coupled with lower biomechanical disparity explains previous findings of redundant and complex form–function mapping of mammalian feeding systems at lower taxonomic rank levels [61] and suggests a revised conceptual framework for understanding crown mammalian jaw evolution (figure 5). These results do not support an interpretation whereby simplified,
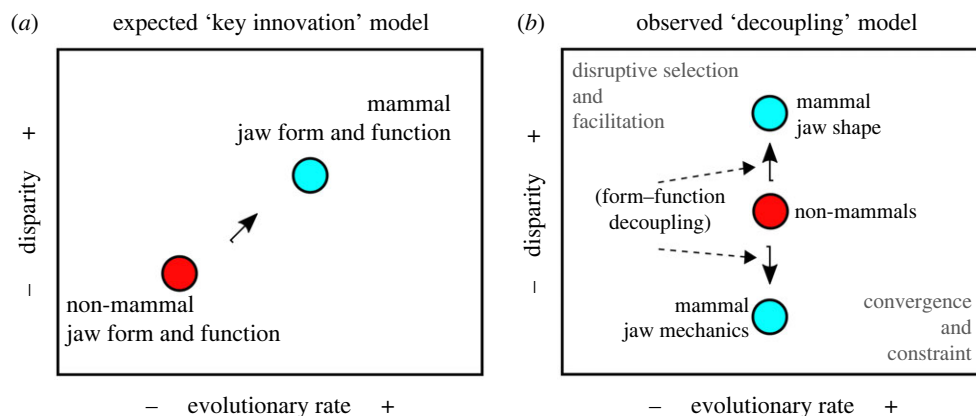
Figure 5. Evolutionary rate–disparity framework for mammalian jaw form–function evolution. (a) Graphical representation of a 'key innovation' type of model for non-mammal to mammal jaw transformation. In this model, the single-element jaw of mammals is associated with increase in both jaw shape and function disparity and evolutionary rate as a result of an adaptive radiation at the origin of mammals. (b) The observed rate-disparity shift from non-mammals to mammals is one of form–function decoupling, whereby mammal jaw shapes increased in disparity and mammal jaw biomechanical traits decreased in disparity relative to non-mammals. Modified from [60]. (Online version in colour.)

single-element jaws substantially improved the force and velocity transmission lever system in terms of MA and AR during the radiation of crown mammals, but do support the interpretation that stiffness is the principal axis of mammalian jaw adaptation. The differences between a single-element mammalian jaw and multi-element non-mammalian jaw are thus best characterized by a switch in form–function coupling rather than a wholesale 'key innovation'. Multi-element systems, such as in squamate skulls, can be highly integrated morphologically but are nevertheless able to respond to dietary selective pressure when examined at lower taxonomic levels [62]; this phenomenon is consistent with the diversity of adaptive landscape regimes across sauropsid clades (electronic supplementary material, figure S24). Specific adaptive advantages for a single-element jaw may lie instead in enabling diverse jaw shapes optimized for stiffness to support the extremely diverse dentitions that are characteristic of mammals. Stiffness-optimized mandibular shapes may also facilitate evolutionary transformations in other, potentially less stereotyped components of the feeding system such as muscle physiology, bone microstructure or other morphological traits with higher functional or mechanical sensitivity [63].

Body size is an important organismal trait underlying many biological scaling relationships [64]; in the context of mammalian jaws, body mass is poorly correlated with morphological and biomechanical traits. This size independence applies both to shape (PC1 and PC2) and biomechanical traits (electronic supplementary material, figures S22, S23 and data S19, S20). This suggests that during their evolutionary radiation, crown mammals of different sizes comparably explored the jaw morphospace region within which they are found in terms of mandible biomechanics. This finding corroborates the recent observation that ecomorphological signals overtake those of body size in mammalian jaws [65] and provides new evidence that despite the extraordinarily large size range represented by living mammals, from 2 g bats to almost 200 000 kg whales, the same genetically correlated, shape-induced mechanical constraints exist.

Although it is likely that functional morphological variation in 3D bone material properties and shape are not fully captured by the 2D mandibular profiles examined in this study, the significantly non-random distributions of both

shape and biomechanical traits in our dataset suggest critical functional signals are still present in 2D data. It is likely that MA, AR and stiffness measures taken from 2D jaw shapes only partially capture the biomechanical characteristics of 3D jaw function, but any systematic errors in estimating those traits using 2D data should not bias the interpretations made herein, unless there are significant differences in how well 2D represent 3D structures from clade to clade. On the other hand, the extent to which jaw shapes of different vertebrate clades overlap each other in morphospace may change if 3D shapes are used; this may be particularly important for taxa in which there are substantial mediolateral morphological and biomechanical differences. Despite these potential shortcomings, the important role of dorsoventral kinetics and kinematics of mammalian and all gnathostome vertebrate jaws (the jaw adductor muscle vectors are principally oriented in the parasagittal plane [66,67] and create rotation mainly around the mediolateral axis and translation mainly in that parasagittal plane) indicates that 2D projections are both informative and time-efficient sources of relevant functional morphology variation.

Other factors to be considered in future research and which may interact with the stiffness-governed form–function regime identified in this study include those associated with mammalian jaw movements about the dorsoventral axis (yaw), which increased in MA relative to the movement about a mediolateral axis in crown mammals [5,68]. Research on the origin of the vertebrate dentition and its evolutionary interactions with the jaw bones housing it also suggests that mechanical environment at the tooth–bone junction is key in jaw system evolution [69]; the importance of tooth–bone interactions became greatly amplified in mammals as a complex neuroanatomical system evolved along with an occlusal dentition and mastication [70]. Furthermore, a single-bone mandible that is stiff regardless of shape may permit larger jaw closing forces via effective resistance to deformation caused by higher absolute muscle contraction magnitudes as mammals evolved increasingly large body sizes through the Cenozoic, starting during the immediate aftermath of the end-Cretaceous extinctions [71–73]. Evolutionary encephalization and epigenetic effects from growing a larger brain likely exerted a cascading influence on mammalian

skull configuration, especially in the auditory, olfactory and masticatory systems [74–76]. Within crown mammals, marsupial mammals appear to exhibit lower jaw disparity relative to placentals given the former's developmental constraints [77]; future investigations into within-mammal variations should clarify how developmental differences among mammalian clades further modulate this broad, stereotyped form–function configuration.

Although mandibular evolution was once thought principally to have enabled mammals to innovate in feeding mechanical versatility and efficiency over those of non-mammal vertebrates, the new data presented here across a wide sample of vertebrates clearly document that stereotyped stiffness in the mandible was a pre-existing condition underpinning the diverse jaw shapes expressed in crown mammals. A further optimized stiffness regime across diverse mammal jaw shapes evolved at the expense of efficiency and speed, but such a regime may have served as the hotbed for the radiation of dental morphologies and masticatory strategies that exemplify mammalian biodiversity. These new findings should encourage further investigations into how crown mammals radiated into new adaptive zones and became ecologically dominant in various ecosystems during the Cenozoic by capitalizing on a fundamentally modified and stereotyped jaw form–function configuration.

royalsocietypublishing.org/journal/rstb Phil. Trans. R. Soc. B 378: 20220091

12

References

1. Wang J, Wible JR, Guo B, Shelley SL, Hu H, Bi S. 2021 A monotreme-like auditory apparatus in a Middle Jurassic haramiyidan. Nature 590, 279–283. (doi:10.1038/s41586-020-03137-z)

2. Barghusen HR, Hopson JA. 1970 Dentary-squamosal joint and the origin of mammals. Science 168, 573–575. (doi:10.1126/science.168.3931.573)

3. Mao F, Hu Y, Li C, Wang Y, Hill Chase M, Smith A, Meng J. 2020 Integrated hearing and chewing modules decoupled in a Cretaceous stem therian mammal. Science 367, 305–308. (doi:10.1126/science.aay9220)

4. Lautenschlager S, Gill PG, Luo Z-X, Fagan MJ, Rayfield EJ. 2018 The role of miniaturization in the evolution of the mammalian jaw and middle ear. Nature 561, 533–537. (doi:10.1038/s41586-018-0521-4)

5. Grossnickle DM. 2017 The evolutionary origin of jaw yaw in mammals. Scient. Rep. 7, 45094. (doi:10.1038/srep45094)

6. Reed DA, Iriarte-Diaz J, Diekwisch TGH. 2016 A three dimensional free body analysis describing variation in the musculoskeletal configuration of the cynodont lower jaw. Evol. Dev. 18, 41–53. (doi:10.1111/ede.12171)

7. Crompton AW, Jenkins FA. 1973 Mammals from reptiles: a review of mammalian origins. Annu. Rev. Earth Planet. Sci. 1, 131–155. (doi:10.1146/annurev.ea.01.050173.001023)

8. Kermack KA, Mussett F, Rigney H. 1973 The lower jaw of Morganucodon. Zool. J. Linn. Soc. 53, 87–175. (doi:10.1111/j.1096-3642.1973.tb00786.x)

9. Crompton AW. 1963 The evolution of the mammalian jaw. Evolution 17, 431–439.

10. Bramble DM. 1978 Origin of the mammalian feeding complex: models and mechanisms. Paleobiology 4, 271–301. (doi:10.1017/S009483730000600X)

11. Cerny R, Cattell M, Sauka-Spengler T, Bronner-Fraser M, Yu F, Medeiros DM. 2010 Evidence for the prepattern/cooption model of vertebrate jaw evolution. Proc. Natl Acad. Sci. USA 107, 17 262–17 267. (doi:10.1073/pnas.1009304107)

12. Deakin WJ, Anderson PS, den Boer W, Smith TJ, Hill JJ, Rücklin M, Donoghue PC, Rayfield EJ. 2022 Increasing morphological disparity and decreasing optimality for jaw speed and strength during the radiation of jawed vertebrates. Sci. Adv. 8, eabl3644. (doi:10.1126/sciadv.abl3644)

13. Olson EC. 1961 Jaw mechanisms: rhipidistians, amphibians, reptiles. Am. Zool. 1, 205–215. (doi:10.1093/icb/1.2.205)

14. Zhu M, Ahlberg P, Pan Z, Zhu Y, Qiao T, Zhao W, Jia L, Lu J. 2016 A Silurian maxillate placoderm illuminates jaw evolution. Science 354, 334–336. (doi:10.1126/science.aah3764)

15. Romer AS. 1969 Cynodont reptile with incipient mammalian jaw articulation. Science 166, 881–882. (doi:10.1126/science.166.3907.881)

16. Zhou C-F, Bhullar B-A, Neander A, Martin T, Luo Z-X. 2019 New Jurassic mammaliaform sheds light on early evolution of mammal-like hyoid bones. Science 365, 276–279. (doi:10.1126/science.aau9345)

17. Luo Z-X, Schultz JA, Ekdale EG. 2016 Evolution of the middle and inner ears of mammaliaforms: the approach to mammals. In Evolution of the vertebrate ear: evidence from the fossil record (eds JA Clack, RR Fay, AN Popper), pp. 139–174. Cham, Switzerland: Springer International Publishing.

18. Martin T, Luo ZX. 2005 Homoplasy in the mammalian ear. Science 307, 861–862. (doi:10.1126/science.1107202)

19. Luo Z-X, Meng Q-J, Grossnickle DM, Liu D, Neander AI, Zhang Y-G, Ji Q. 2017 New evidence for mammaliaform ear evolution and feeding adaptation in a Jurassic ecosystem. Nature 548, 326–329. (doi:10.1038/nature23483)

20. Luo Z-X, Gatesy S, Shubin N. 2015 Mandibular and dental characteristics of Late Triassic mammaliaform Haramiyavia and their ramifications for basal mammal evolution. Proc. Natl Acad. Sci. USA 112, E7101–E7109. (doi:10.1073/pnas.1519387112)

21. Benevento GL, Benson RBJ, Friedman M. 2019 Patterns of mammalian jaw ecomorphological disparity during the Mesozoic/Cenozoic transition.

Proc. R. Soc. B **286**, 20190347. (doi:10.1098/rspb.2019.0347)

22. Morales-García NM, Gill PG, Janis CM, Rayfield EJ. 2021 Jaw shape and mechanical advantage are indicative of diet in Mesozoic mammals. *Commun. Biol.* **4**, 242. (doi:10.1038/s42003-021-01757-3)

23. Wainwright PC. 1996 Ecological explanation through functional morphology: the feeding biology of sunfishes. *Ecology* **77**, 1336–1343. (doi:10.2307/2265531)

24. McHenry MJ. 2011 There is no trade-off between speed and force in a dynamic lever system. *Biol. Lett.* **7**, 384–386. (doi:10.1098/rsbl.2010.1029)

25. Crampton JS. 1995 Elliptic Fourier shape analysis of fossil bivalves: some practical considerations. *Lethaia* **28**, 179–186. (doi:10.1111/j.1502-3931.1995.tb01611.x)

26. Hill JJ, Puttick MN, Stubbs TL, Rayfield EJ, Donoghue PCJ. 2018 Evolution of jaw disparity in fishes. *Palaeontology* **61**, 847–854. (doi:10.1111/pala.12371)

27. Navarro CA, Martin-Silverstone E, Stubbs TL. 2022 Morphometric assessment of pterosaur jaw disparity. *R. Soc. Open Sci.* **5**, 172130. (doi:10.1098/rsos.172130)

28. Schaeffer J, Benton MJ, Rayfield EJ, Stubbs TL. 2020 Morphological disparity in theropod jaws: comparing discrete characters and geometric morphometrics. *Palaeontology* **63**, 283–299. (doi:10.1111/pala.12455)

29. Ferson S, Rohlf FJ, Koehn RK. 1985 Measuring shape variation of two-dimensional outlines. *Syst. Zool.* **34**, 59–68. (doi:10.2307/2413345)

30. Lestrel PE. 1989 Method for analyzing complex two-dimensional forms: elliptical Fourier functions. *Am. J. Hum. Biol.* **1**, 149–164. (doi:10.1002/ajhb.1310010204)

31. Bonhomme V, Picq S, Gaucherel C, Claude J. 2014 Momocs: outline analysis using R. *J. Stat. Softw.* **56**, 1–24. (doi:10.18637/jss.v056.i13)

32. Westneat MW. 2003 A biomechanical model for analysis of muscle force, power output and lower jaw motion in fishes. *J. Theor. Biol.* **223**, 269–281. (doi:10.1016/S0022-5193(03)00058-4)

33. Schindelin J et al. 2012 Fiji: an open-source platform for biological-image analysis. *Nat. Methods* **9**, 676–682. (doi:10.1038/nmeth.2019)

34. Tseng ZJ, Su DF, Wang X, White SC, Ji X. 2017 Feeding capability in the extinct giant *Siamogale melilutra* and comparative mandibular biomechanics of living Lutrinae. *Scient. Rep.* **7**, 15225. (doi:10.1038/s41598-017-15391-9)

35. Guillerme T. 2018 dispRity: a modular R package for measuring disparity. *Methods Ecol. Evol.* **9**, 1755–1763. (doi:10.1111/2041-210X.13022)

36. Jones KE et al. 2009 PanTHERIA: a species-level database of life history, ecology, and geography of extant and recently extinct mammals: ecological archives E090-184. *Ecology* **90**, 2648. (doi:10.1890/08-1494.1)

37. Schaerlaeken V, Herrel A, Aerts P, Ross CF. 2008 The functional significance of the lower temporal bar in *Sphenodon punctatus*. *J. Exp. Biol.* **211**, 3908–3914. (doi:10.1242/jeb.021345)

38. Herrel A, Adriaens D, Verraes W, Aerts P. 2002 Bite performance in clariid fishes with hypertrophied jaw adductors as deduced by bite modeling. *J. Morphol.* **253**, 196–205. (doi:10.1002/jmor.1121)

39. van der Meij M, Bout R. 2004 Scaling of jaw muscle size and maximal bite force in finches. *J. Exp. Biol.* **207**, 2745–2753. (doi:10.1242/jeb.01091)

40. Herrel A, De Smet A, Aguirre LF, Aerts P. 2008 Morphological and mechanical determinants of bite force in bats: do muscles matter? *J. Exp. Biol.* **211**, 86–91. (doi:10.1242/jeb.012211)

41. Perry JMG, Hartstone-Rose A, Wall CE. 2011 The jaw adductors of strepsirrhines in relation to body size, diet, and ingested food size. *Anat. Rec.* **294**, 712–728. (doi:10.1002/ar.21354)

42. Hartstone-Rose A, Hertzig I, Dickinson E. 2019 Bite force and masticatory muscle architecture adaptations in the dietarily diverse Musteloidea (Carnivora). *Anat. Rec.* **302**, 2287–2299. (doi:10.1002/ar.24233)

43. Hartstone-Rose A, Dickinson E, Deutsch AR, Worden N, Hirschkorn GA. 2022 Masticatory muscle architectural correlates of dietary diversity in Canidae, Ursidae, and across the order Carnivora. *Anat. Rec.* **305**, 477–497. (doi:10.1002/ar.24748)

44. Dickson BV, Clack JA, Smithson TR, Pierce SE. 2021 Functional adaptive landscapes predict terrestrial capacity at the origin of limbs. *Nature* **589**, 242–245. (doi:10.1038/s41586-020-2974-5)

45. Dickson BV, Pierce SE. 2019 Functional performance of turtle humerus shape across an ecological adaptive landscape. *Evolution* **73**, 1265–1277.

46. Polly PD, Stayton CT, Dumont ER, Pierce SE, Rayfield EJ, Angielczyk KD. 2016 Combining geometric morphometrics and finite element analysis with evolutionary modeling: towards a synthesis. *J. Vertebr. Paleontol.* **36**, e1111225. (doi:10.1080/02724634.2016.1111225)

47. W D. 2012 Bapst, paleotree: an R package for paleontological and phylogenetic analyses of evolution. *Methods Ecol. Evol.* **3**, 803–807. (doi:10.1111/j.2041-210X.2012.00223.x)

48. Uyeda JC, Caetano DS, Pennell MW. 2015 Comparative analysis of principal components can be misleading. *Syst. Biol.* **64**, 677–689. (doi:10.1093/sysbio/syv019)

49. Adams DC, Collyer ML. 2018 Multivariate phylogenetic comparative methods: evaluations, comparisons, and recommendations. *Syst. Biol.* **67**, 14–31. (doi:10.1093/sysbio/syx055)

50. Adams DC. 2014 Quantifying and comparing phylogenetic evolutionary rates for shape and other high-dimensional phenotypic data. *Syst. Biol.* **63**, 166–177. (doi:10.1093/sysbio/syt105)

51. Galili T. 2015. dendextend: an R package for visualizing, adjusting and comparing trees of hierarchical clustering. *Bioinformatics* **31**, 3718–3720. (doi:10.1093/bioinformatics/btv428)

52. Kowalczyk A, Meyer WK, Partha R, Mao W, Clark NL, Chikina M. 2019 RERconverge: an R package for associating evolutionary rates with convergent traits. *Bioinformatics* **35**, 4815–4817. (doi:10.1093/bioinformatics/btz468)

53. Diaz G. 2017 Contour recognition of complex leaf shapes. *PLoS ONE* **12**, e0189427. (doi:10.1371/journal.pone.0189427)

54. Wainwright PC, Alfaro ME, Bolnick DI, Hulsey CD. 2005 Many-to-one mapping of form to function: a general principle in organismal design? *Integr. Comp. Biol.* **45**, 256–262. (doi:10.1093/icb/45.2.256)

55. Wainwright PC, McGee MD, Longo SJ, Hernandez LP. 2015, Origins, innovations, and diversification of suction feeding in vertebrates. *Integr. Comp. Biol.* **55**, 134–145. (doi:10.1093/icb/icv026)

56. Tseng ZJ. 2013 Testing adaptive hypotheses of convergence with functional landscapes: a case study of bone-cracking hypercarnivores. *PLoS ONE* **8**, e65305. (doi:10.1371/journal.pone.0065305)

57. Jeong J, Li X, McEvilly RJ, Rosenfeld MG, Lufkin T, Rubenstein JLR. 2008 Dlx genes pattern mammalian jaw primordium by regulating both lower jaw-specific and upper jaw-specific genetic programs. *Development* **135**, 2905–2916. (doi:10.1242/dev.019778)

58. Depew MJ, Lufkin T, Rubenstein JLR. 2002 Specification of jaw subdivisions by *Dlx* genes. *Science* **298**, 381–385. (doi:10.1126/science.1075703)

59. Heude É, Bouhali K, Kurihara Y, Kurihara H, Couly G, Janvier P, Levi G. 2010 Jaw muscularization requires *Dlx* expression by cranial neural crest cells. *Proc. Natl Acad. Sci. USA* **107**, 11 441–11 446. (doi:10.1073/pnas.1001582107)

60. Felice RN, Randau M, Goswami A. 2018 A fly in a tube: macroevolutionary expectations for integrated phenotypes. *Evolution* **72**, 2580–2594.

61. Tseng ZJ, Flynn JJ. 2018 Structure-function covariation with nonfeeding ecological variables influences evolution of feeding specialization in Carnivora. *Sci. Adv.* **4**, eaao5441. (doi:10.1126/sciadv.aao5441)

62. Watanabe A, Fabre A-C, Felice RN, Maisano JA, Müller J, Herrel A, Goswami A. 2019 Ecomorphological diversification in squamates from conserved pattern of cranial integration. *Proc. Natl Acad. Sci. USA* **116**, 14 688–14 697. (doi:10.1073/pnas.1820967116)

63. Muñoz MM, Anderson PSL, Patek SN. 2017 Mechanical sensitivity and the dynamics of evolutionary rate shifts in biomechanical systems. *Proc. R. Soc. B* **284**, 20162325. (doi:10.1098/rspb.2016.2325)

64. LaBarbera M. 1989 Analyzing body size as a factor in ecology and evolution. *Annu. Rev. Ecol. Syst.* **20**, 97–117. (doi:10.1146/annurev.es.20.110189.000525)

65. Grossnickle DM. 2020 Feeding ecology has a stronger evolutionary influence on functional morphology than on body mass in mammals. *Evolution* **74**, 610–628. (doi:10.1111/evo.13929)

66. Datovo A, Vari RP. 2013 The jaw adductor muscle complex in teleostean fishes: evolution, homologies and revised nomenclature (Osteichthyes:

Actinopterygii). *PLoS ONE* **8**, e60846. (doi:10.1371/journal.pone.0060846)

67. Ziermann JM, Diaz RE, Diogo R. 2019 *Heads, jaws, and muscles: anatomical, functional, and developmental diversity in chordate evolution*. Cham, Switzerland: Springer.

68. Bhullar B-AS, Manafzadeh AR, Miyamae JA, Hoffman EA, Brainerd EL, Musinsky C, Crompton AW. 2019 Rolling of the jaw is essential for mammalian chewing and tribosphenic molar function. *Nature* **566**, 528–532. (doi:10.1038/s41586-019-0940-x)

69. Valéria V, Donglei C, Paul T, Zerina J, Boris E, Henning B, Erik AP. 2020 Marginal dentition and multiple dermal jawbones as the ancestral condition of jawed vertebrates. *Science* **369**, 211–216. (doi:10.1126/science.aaz9431)

70. Rowe TB. 2020 The emergence of mammals. In *Evolutionary neuroscience*, 2nd edn (ed. JH Kaas), pp. 263–319. London, UK: Academic Press.

71. Alroy J. 1998 Cope's rule and the dynamics of body mass evolution in North American fossil mammals. *Science* **280**, 731–734. (doi:10.1126/science.280.5364.731)

72. Baker J, Meade A, Pagel M, Venditti C. 2015 Adaptive evolution toward larger size in mammals. *Proc. Natl Acad. Sci. USA* **112**, 5093–5098. (doi:10.1073/pnas.1419823112)

73. Bertrand O *et al.* 2022 Brawn before brains in placental mammals after the end-Cretaceous extinction. *Science* **376**, 80–85. (doi:10.1126/science.abl5584)

74. Rowe TB. 1996 Brain heterochrony and origin of the mammalian middle ear. *Mem. CA Acad. Sci.* **20**, 71–95.

75. Rowe TB. 1996 Coevolution of the mammalian middle ear and neocortex. *Science* **273**, 651–654. (doi:10.1126/science.273.5275.651)

76. Rowe TB, Macrini TE, Luo Z-X. 2011 Fossil evidence on origin of the mammalian brain. *Science* **332**, 955–957. (doi:10.1126/science.1203117)

77. Fabre A-C, Dowling C, Portela Miguez R, Fernandez V, Noirault E, Goswami A. 2021 Functional constraints during development limit jaw shape evolution in marsupials. *Proc. R. Soc. B* **288**, 20210319. (doi:10.1098/rspb.2021.0319)

78. Tseng ZJ, Garcia-Lara S, Flynn JJ, Holmes E, Rowe TB, Dickson BV. 2023 A switch in jaw form–function coupling during the evolution of mammals. Figshare. (doi:10.6084/m9.figshare.c.6484657)

14

royalsocietypublishing.org/journal/rstb   *Phil. Trans. R. Soc. B* **378**: 20220091