Multiple paths to morphological diversification during the origin of amniotes

Neil Brocklehurst*[1], Roger J Benson[1]

1 – Department of Earth Sciences, University of Oxford, South Parks Road, Oxford, UK

* - Corresponding author: Neil Brocklehurst, neil.brocklehurst@earth.ox.ac.uk

Early terrestrial vertebrates (amniotes) provide a classic example of diversification following adaptive zone invasion. The initial terrestrialisation of vertebrates was closely followed by dietary diversification, as evidenced by a proliferation of craniomandibular and dental adaptations. However, morphological evolution of early amniotes has received limited study, in analyses with restricted taxonomic scope, leaving substantial questions about the dynamics of this important terrestrial radiation. We use novel analyses of discrete characters to quantify variation in evolutionary rates and constraints during diversification of the amniote feeding apparatus. We find evidence for an early burst, comprising high rates of anatomical change that decelerated through time, giving way to a background of saturated morphological evolution. Subsequent expansions of phenotypic diversity were not associated with increased evolutionary rates. Instead, variation in the mode of evolution became important, with groups representing independent origins of herbivory evolving distinctive, group-specific morphologies and thereby exploring novel character state spaces. Our findings indicate the importance of plant-animal interactions in structuring the earliest radiation of amniotes, and demonstrate the importance of variation in modes of phenotypic divergence during a major evolutionary radiation.

Key Words: Amniotes; Herbivore; Teeth; Adaptive Radiation; Quantum Evolution

Amniotes comprise mammals and reptiles (including birds), their most recent common ancestor and all of its descendants. These groups are fully terrestrialised, with reproduction and other aspect of ecology occurring independently of water. The fossil record of early amniotes documents a shift to fully terrestrial lifestyles during the Bashkirian stage of the late Carboniferous, about 325 million years ago (Mya), as indicated by a suite of skeletal anatomical traits facilitating locomotion, lung ventilation and feeding on land in early amniotes [1,2]. This was closely followed by dietary diversification. Herbivory in particular was central to the evolutionary establishment of modern-type terrestrial ecosystems, allowing access to the vast energetic resource of stored primary production in land plants [3,4]. Herbivory first appeared during the Kasimovian (late Pennsylvanian), around 306 million years ago (Mya), and had evolved independently in at least eight separate groups by the end of the Paleozoic, 252 Mya [5]. This provides a striking example of evolutionary convergence, with early herbivorous amniotes showing substantial variation in their craniodental adaptations for processing plant matter, including molariform posterior dentition in diadectids [6,7], grinding tooth plates in both edaphosaurids and captorhinids [8,9] and a toothless keratinous beak in anomodonts [10].

Vertebrate terrestrialisation is a classic example of adaptive zone invasion followed by phenotypic diversification, and provides a natural laboratory for examining how variation in rates and modes of evolution give rise to large increases in morphological diversity (disparity). Adaptive zones comprise sets of ecological opportunities or niches, and the theory of adaptive radiation [11,12] proposes that the evolutionary discovery of new or previously under-exploited adaptive zones results in divergence occurring along ecological lines [13-15], resulting in rapid increases in disparity during the early history of a group [16-18]. This has more recently been formalised by the “early burst” model, which attributes expansions of disparity primarily to a short-lived burst of high rates of anatomical change (hereafter evolutionary rates) occurring simultaneously in many lineages, which then decelerate through time as niches become filled [19-21].

Quantitative study of adaptive radiation has focussed on the role of rate variation, which has become paradigmatical as a mechanism controlling the accumulation of phenotypic disparity ([22-25], but see [26,27]). However, adaptive radiation could be envisaged more broadly as the early attainment of high disparity by lineage-specific shifts in phenotype due to variation in the dynamics of trait evolution – including variation in both evolutionary rates and constraints. For example, in addition to describing hypothesised patterns of rate variation, Simpson's foundational exposition of the theory of adaptive radiations also included the concept of quantum evolution [12]. This describes a pattern of generally constrained evolution, in which increases in disparity occur due to rare, but abrupt shifts of morphology and ecology in individual lineages, which explore new areas of phenotypic space. Evolutionary constraints may result either from intrinsic factors such as inherited developmental or functional limitations on the range of possible forms and evolutionary changes [28,29], or from extrinsic factors such as the availability of ecological niches [30,31]. Therefore, increases in disparity symptomatic of adaptive radiation could result from relaxation of developmental constraints, from evolution into new niches, or perhaps frequently, from both effects acting together [28,29].

Ultimately, the hypothesis of adaptive radiation seeks to explain variation in ecological diversity among groups, which is reflected most directly by the diversity of ecologically-relevant phenotypic traits such as the feeding apparatus. Early amniotes exhibit high disparity of jaw morphologies and functions compared to other tetrapodomorphs [32], consistent with a hypothesised adaptive radiation during vertebrate terrestrialisation. This evolutionary expansion of amniote feeding mechanisms was most likely enabled by anatomical changes to the postcranial musculoskeletal system that liberated the skull from its primitive role in lung ventilation. The origins of costal (rib-based) lung ventilation removed the need for cranial buccal pumping in ventilation [2], allowing greater evolutionary versatility of skull and jaw morphology and function [1,33,34]. In particular, reorientation of jaw-closing (adductor) muscles permitted static pressure to be exerted at the front of the jaw when nearly closed, enabling grinding and other forms of oral processing as prerequisites for a wide range of novel diets and ecologies, including 'high-fibre' herbivory of land plants [2]. Dietary diversification is therefore central to understanding the evolutionary dynamics of early amniotes.

Macroevolutionary studies of early amniotes have thus far focussed on patterns of lineage diversification (origination and extinction) and the establishment of species richness rather than patterns of morphological evolution [35-37]. This provides only an indirect and incomplete perspective on adaptive radiation because increases in species richness can occur independently of ecological or phenotypic diversification [15,38,39], and phenotype has a more direct relationship with the ecological niche than does species diversity [15,38]. So far however, few analyses of morphological evolution in early amniotes have been conducted, and these have been limited in taxonomic scope [40-44].

We examine the evolutionary dynamics of the early amniote feeding apparatus, including the macroevolutionary effects of the origins of amniote herbivory, by analysing patterns of dental evolution among early fossil amniotes. Teeth play a direct role in food acquisition, and dental variation is generally thought to reflect diversification of feeding modes [45-47]. Our analysis examines not only the changes in dental disparity and variation in rates of dental evolution, as in previous studies, but also uses a novel method to quantify variation in morphological constraints, based on the ability of lineages to explore novel regions of character space (character state saturation). We interrogate patterns across all early amniotes, and also among faunivorous and herbivorous taxa individually. Our analysis spans

~70 million years, from the Carboniferous origin of amniotes until the Olenekian (Early Triassic), employing a matrix of 114 discrete characters and a supertree of 539 taxa (Fig. 1) time-calibrated using the fossilised-birth-death model.

Results and discussion

Rates of dental evolution and disparity during the origin of amniotes.

Amniotes show clear evidence for an early burst of high rates of dental evolution among the oldest (predominantly faunivorous) lineages, with high rates during their earliest origin and a rapid fall to lower rates that persist throughout the Pennsylvanian and Cisuralian (Fig. 2a). Numerous novel dental morphologies appeared rapidly during the early evolution of amniote faunivores, including anisodonty (potentially primitive for amniotes) [48], bulbous teeth for durophagy [49] and fields of enlarged palatal teeth [49,50], all observed in the fossil record in the earliest Pennsylvanian.

The rates observed in herbivores also show a decrease from an initial peak, but the magnitude and trend are similar to those observed in amniotes as a whole, rather than representing an exceptional early burst in one dietary guild. Instead, rates of dental evolution at the first appearance of herbivorous groups during the Pennsylvanian are similar to those of contemporary faunivore lineages, and remained so until the Artinskian, about 285 Mya (Fig. 2b). Evolutionary rates were more volatile in this early interval of herbivore evolution, indicating greater between-branch variation in rates, but there is little indication of an overall rate increase sparked by the origin of herbivory.

Despite this, dental disparity of amniote herbivores increased rapidly, surpassing that of faunivores by the end of the Carboniferous (Fig. 2d). Disparity of faunivores remained relatively constant throughout the Cisuralian; by-interval Mann Witney tests, comparing the disparity values obtained in each bin to that preceding it, indicates no significant shifts in faunivore disparity between the start of the Asselian, 298 mya, and the latest Kungurian, 278 mya (Supplementary Table 1). Herbivore disparity was more variable (Supplementary Table 1) but remained consistently higher than that of faunivores throughout this interval (Fig. 2d).

Analysis using an alternative tree topology based on recent revisions to the amniote phylogeny [51] found similar patterns of disparity and rate variation (Supplementary Figures 1,2). Initial rates of herbivore dental evolution were lower than those of carnivores, while herbivore disparity rapidly overtook that of carnivores. This alternative dataset did produce a peak in herbivore rates in the latest Carboniferous, but this is short lived and for the rest of the Permian rates of dental evolution in herbivores remain consistent with those of carnivores.

Middle and Late Permian Rates of Evolution.

Rates of dental evolution in both faunivores and herbivores declined through the Carboniferous and the initial stages of the Cisuralian, consistent patterns seen across tetrapod morphology as a whole [52]. This trend ends around the Artinskian (mid- early Permian), about 285 Mya, when rates began to increase again in both guilds. Herbivore rates increased more rapidly at first, exceeding those of faunivores until the early Guadalupian (middle Permian; Fig. 2b).

This period of rate increase coincides with substantial changes in terrestrial ecosystems and the establishment of a more modern trophic structure. Earlier assemblages were overwhelmingly dominated by faunivores, with herbivores remaining rare from their first appearance in the latest Carboniferous [3,5,35]. Large subsequent increases in the abundance of herbivores occurred by the middle Permian, at which time herbivorous tetrapods dominated both in diversity and abundance [3,35]. This represents the establishment of more modern-like terrestrial ecosystems and was accompanied by a substantial faunal turnover and possible mass extinction [1,53-55]. Terrestrial ecosystems of the early Permian were dominated by pelycosaurs (a paraphyletic grade of early-diverging synapsids) and amphibians (non-amniote tetrapods) [45,46,49]. This fauna was replaced during the middle–late Permian by one dominated by therapsids (the larger synapsid clade containing mammals) and parareptiles (a stem reptile lineage) [49-51,53]. Among herbivores, edaphosaurids and diadectids declined and were replaced by lineages with entirely novel dental morphologies; initially the captorhinids, with multiple rows of maxillary and dentary teeth forming grinding tooth plates, and caseids, with multicusped spatulate teeth [35,41,56]. A second turnover of herbivores occurred across the Cisuralian/Guadalupian boundary, coinciding with the radiation of anomodont and tapinocephalian therapsids and pareiasaurian parareptiles [35,40,41,56-58]. These two turnovers coincide with, in the first case, peak disparity in herbivorous captorhinids [41] and in the second, an early peak in disparity of anomodonts [40] and a similar peak in that of parareptiles [43] (although the latter did not specifically examine patterns within the major herbivorous group of parareptiles, the pareiasaurs).

Character Saturation.

The accumulation of high relative disparity, without increased rates of evolution, among early amniote herbivores presents an apparent paradox. High dental disparity of early amniote herbivores cannot be explained by differences in the rate of evolution compared to those of faunivores, but can potentially be explained by differences in the mode of evolution. We therefore explored the possibility that faunivores exhibit more constrained evolution, repeatedly evolving similar dental morphologies via convergence.

To examine this, we used the concept of character saturation, or exhaustion. Character saturation describes the point at which further character changes are more likely to result in homoplasy than a new state [22,59]. We assess levels of character saturation using the method employed by Brocklehurst et al. [60], comparing the morphological disparity between taxon pairs (pairwise character state dissimilarity) to the inferred character state changes measured along the branches of the phylogeny (‘patristic morphological distance’), a measure of the amount of evolutionary change since the divergence of a pair of taxa from their most recent common ancestor. While dissimilarity generally increases with increasing evolutionary change, this relationship will asymptote as the lineage reaches the limit of its morphological constraint (i.e. at character state saturation). After this, further evolution will generally result in homoplasy rather than new characters states. Lower morphological disparity at the asymptote indicates greater constraint on morphological evolution.

When applying this method to dental evolution in early amniotes, the loess regression curve fitted to pairwise comparisons of faunivores asymptotes at a morphological distance of less than 0.4 (Fig. 3a). The loess curve of the herbivores has not yet reached its asymptote, despite having reached a higher morphological dissimilarity than that of faunivores. This suggests that faunivores are evolving under increased constraint relative to herbivores, and that the occurrence of relaxed constraints on dental evolution allowed early amniotes herbivores to attain higher morphological and ecological disparity.

The apparent presence of relaxed constraints in herbivores results primarily from comparisons among herbivores from different groups, representing independent origins of herbivory (Fig 3b). Within-group comparisons between pairs of herbivorous species within a single clade, representing a single origin of herbivory, indicate similar levels of morphological constraint to those of faunivores (Fig. 3b). This indicates that each evolutionary transition to herbivory reflects the establishment of a distinctive character state space, and that this process of evolutionary ‘discovery’ of new regions of phenotypic space gave rise to most of the dental disparity of early amniotes. This pattern of constraint in herbivores is also consistent with the concept of quantum evolution [12].

We show that the dynamics of dental evolution of this interval, persisting for more than 30 million years, were dominated by eco-phenotypic shifts associated with the specialised consumption of land plants. This observation provides strong evidence for the advent of herbivory as a key innovation involved in vertebrate terrestrialisation and the assembly of modern-type ecosystems on land.

Conclusions

G.G. Simpson hypothesised that phenotypic radiations driven by the invasion of novel adaptive zones, might be responsible for much of the diversity of life on Earth [11,12]. Simpson's [11] hypothesis emphasised variation in rates of evolution; asking "How fast, as a matter of fact, do animals evolve in nature?" ([11]: pg 3). Since then, rates of diversification, either in accumulation of lineages [36,61-65], morphological diversification [22-25], or dispersal rates [66], have dominated discussions of evolutionary radiation. However, Simpson also discussed how shifts in modes of evolution can allow invasion of new adaptive zones, including release of constraints and shifts to novel regions of character state space [11].

Early shifts to high morphological diversity are common among animals [67-68]. However, these shifts can result from either high early rates of evolution, or from variation in evolutionary modes. For example, evolutionary lineages exploring constrained trait spaces will also show early establishment of their complete or near-complete phenotypic space, even in the absence of early rate bursts e.g. [67,69,70]. Indeed, evidence for early bursts of morphological evolution, in which rates are high during the early history of a group, and then subsequently decrease are rare in comparative datasets of extant species [20,71,72]. Harmon et al [20] instead found most groups are characterised by constrained evolution. In this context, expansions of phenotypic space associated with adaptive zone invasions can result either from variation in rates of evolution, or by variation in constraints or the mode of evolution, allowing exploration of novel phenotypic spaces by evolving lineages (somewhat analogous to Simpson's 'quantum evolution' [12]).

The origin of fully terrestrial vertebrates (amniotes) provides an example in which both processes operated, albeit at different times. They underwent a conspicuous early burst of dental evolution, characterised by high rates that established the phenotypic space of early amniotes, and decreased through time, giving way to a constrained pattern of subsequent evolution. Further expansions of ecomorphospace, occurred during independent origins of herbivory. These were no characterised by increases in rate, but by the origins of constraint-breaking innovations allowing access to new peaks associated with functionally disparate methods of processing land plant biomass. Subsequently, constrained evolution continued to occur within these groups, consistent with the hypothesis that phenotypic evolution is frequently constrained, as inferred in studies of some other groups e.g. [20,70].

Our findings indicate variation in rates and modes of evolution unfolding during a large-scale adaptive radiation of the land. Examinations of similar shifts, such as terrestrialisation of plants and arthropods could test the generality of these patterns. Examination of plant disparity [73] suggests an early expansion of morphological diversity at the origin of land plants, with subsequent radiations representing re-colonisations of already established regions of morphospace. While this study did not specifically test rates of evolution, the disparity patterns are consistent with an early burst followed by ecomorphological constraint and warrant further study. With regards to arthropods, there has been little examination of morphological diversity around their terrestrialisation, with more focus on the Cambrian explosion (for summary see [74]). However, lineage diversification patterns in insects, along with study of mouthpart types, do provide evidence for an early burst [75,76], providing a promising avenue for future research.

The general scarcity of early-burst like patterns of rate variation in phylogenetic comparative studies of extant taxa (e.g. [17]) raises questions about their occurrence in fossil record studies, including in early amniotes as shown here. High early rates have been documented in various palaeontological studies so far such as the Cambrian radiation of bilaterian animals [77], the early Cenozoic placental mammal radiation following the end-Cretaceous mass extinction [78] and various others [e.g. 41,79-82]. Indeed, they may be common, because high early disparity has been widely documented in many fossil record studies [22,40,52,68,72], and could result from early bursts of high rates of evolution or from the interaction of high early rates and relaxed constraints [82] (although a similar pattern is expected under constrained evolution without a role of high early rates [67]).

We know of at least two possible explanations for this discrepancy between the findings of extant-only and fossil record studies. Firstly, there may be a role for scale-dependency, in which large-scale adaptive zone invasions such as vertebrate terrestrialisation are characterised by conspicuous patterns of high early rates that are generally not present on the smaller scales often documented by comparative studies of extant radiations such as island endemics or similar [25]. Alternatively, the rarity of early bursts in analyses of extant taxa may be an artefact of the lack of fossil data. This could happen if much of the phenotypic disparity that resulted from an ‘early burst’ of morphological evolution was subsequently lost due to extinction. Indeed, very few of the earliest divergences in amniotes could be reconstructed from analysis solely of extant taxa, suggesting that palaeontological data may be central to a rigorous test of the early burst hypothesis [83] As reviewed above, studies so far suggest that early bursts may be common in the major evolutionary radiations of deep time. We suggest that analyses of the interaction between of variation in evolutionary rates and constraints would further shed light on the dynamics of these radiations, contributing further to the theory of adaptive radiations.

In a broader macroevolutionary context, our findings highlight the importance of examining both rates of evolution and evolutionary constraints. Both are necessary to explain the dynamics of dental evolution and expansion of phenotypic diversity during the early history of land vertebrates. This is also likely true across other groups and events in Earth’s history and we propose that combined studies of variation in rates and constraints of phenotypic evolution could substantially advance understanding of the macroevolutionary processes underlying attainment of the vast phenotypic diversity of life on Earth.

Methods

Supertree construction. The phylogenetic framework for this study was a formal supertree designed to maximise the taxonomic scope of the study. The source trees were selected according to criteria laid out in Brocklehurst et al. [84], with the temporal scope being limited to the end of the Olenekian and the taxonomic scope being restricted to amniotes and their two successive outgroups according to Pardo et al. [85]: Diadectomorpha and Seymouriamorpha. A single tip representing Temnospondyli was added as an outgroup to provide a calibration point for time-scaling the tree (see below). The source tree list was finalised in February 2019.

The source trees were assembled using matrix representation with parsimony (MRP) [86,87]. The MRP matrix was generated using the program Supertree0.85b [88], using the Baum and Ragan method, with all source trees receiving equal weight. The matrix was analysed in TNT [89], using the driven search at level 100. The most parsimonious trees were searched for 100 times, and then a branch-and-bound search was carried out using each most parsimonious tree previously found as a start. Lineages containing no taxa from the time period under study were collapsed into a single tip but were retained in the analysis during time calibration. The final supertree contains 539 taxa (Fig. 1). In order to test the impact of recent modifications to the amniote phylogeny on the results, a second supertree was created with relationships between major clades representing those found by Ford and Benson [50] and subjected to the same analyses of rate and disparity described below (Supplementary Figures 1,2).

Character data. A matrix of discrete dental characters was developed for use both in time calibrating the tree (see below) and for the analyses of dental evolution. These characters were drawn in part from fifteen published phylogenetic analyses. All dental characters from these lists were considered, with those invariable, unknown, or inapplicable for all taxa in the supertree being dropped. These include marginal, palatal, coronoid and prearticular teeth, as well as characters relating to beak morphology for relevant taxa. Eight new characters were also created.

Many of the characters retained required modification. Since most had been formulated for a phylogenetic analysis of a restricted portion of the amniote phylogeny, many required addition or modification of character states to make them more broadly applicable. There were also substantial issues with character non-independence that required reformulation or deletion of characters. Details of the policies employed in formulating and deleting characters, along with the character list itself, are presented in the supplementary notes. The final list contains 114 characters. Due to the substantial modification of most of the characters, all taxa were scored from scratch either from the descriptions in the literature or direct observation of specimens, rather than using the scores from the relevant matrices. 463 of the taxa in the supertree could be scored for at least one of the characters.

Time Calibration of the phylogeny. The tree was timescaled using the fossilised birth death (FBD) model [90] in MrBayes 3.2.6 [91]. The topology was constrained to the strict consensus output by TNT. To account for the uncertainty in the time of the first appearances, the ages of taxa were represented by a uniform probability distribution covering the full possible age of the formation or assemblage zone in which they first appear. Three node calibrations were used: the split between temnospondyls and amniotes (crown tetrapods), the split between synapsids and sauropsids (crown amniotes), and the node containing testudinates, archosauromorphs and lepidosauromorphs (crown saurians). Truncated normal root priors were applied with the mean and standard deviation taken from the molecular clock ages found by Pyron et al. [92], chosen as the most recent molecular clock study producing ages for all these nodes older than the oldest fossil taxon in the clades.

The matrix of dental characters was incorporated into the analysis. The independent gamma rates model was employed (an uncorrelated clock model where rates are drawn from a gamma distribution). Rate heterogeneity between characters was also modelled as a gamma distribution. The analysis was carried out with two runs containing four chains for 20000000 generations, sampling every 1000. 25% of trees were discarded as burn-in.

Disparity. The matrix of discrete dental characters was used in the analysis of morphological diversity (disparity). A pairwise morphological distance matrix comparing all taxa that could be scored for dental characters was calculated using the R package Claddis [93]. The Maximum Observable Rescaled Distance (MORD) was used, having been shown to be more suitable for datasets with large amounts of missing data. In addition to the observed tip taxa, reconstructed ancestral taxa were included in the pairwise distance matrix, their character states estimated using likelihood. This allows as-yet-unsampled morphologies to be incorporated into the disparity estimates. As recommended by Lloyd [93], missing data was not reconstructed. To account for uncertainty in the phylogenetic relationships and the divergence dates, 100 trees were selected at random from the posterior distribution to form the basis of the analyses.

The median pairwise morphological distance between taxa present at a particular time was used as the disparity metric. For calculating the disparity through time, the time-slicing approach [94] was used, rather than binning the data, as the latter approach potentially allows taxa that did not overlap in time to be included in the same bin. The tree was sliced at million-year intervals, and the branches present at each point in time were included in the disparity calculations of that time.

The branches were assigned to either a faunivorous or herbivorous category, in order to compare disparity between the two. This assignment was based on stochastic character mapping using the simmap procedure [95]. 100 evolutionary histories were generated in this manner over the 100 trees using the R package phytools [96]. Tips were assigned to categories based on opinions in the published literature regarding their ecology, and where the diet of a species was uncertain, the tip would be assigned to a dietary category at random for each of the stochastic maps generated.

The time slicing approach of studying disparity requires models of morphological evolution to infer a branch’s morphology at the point it is sliced. The gradual splits model, where the branch sliced will be assigned at random the morphology of either its ancestral or descendant node, with the probability of selecting one or the other being defined by the distance along the branch the slice occurs, was used here [94].

Rates. The FBD method of time calibrating the phylogeny infers rates of character evolution along each branch simultaneously while inferring topology and branch lengths. As MrBayes only ouputs rates for the maximum clade credibility tree, only this tree was used in this section of analysis. These rates of evolution between faunivores and herbivores were compared using the stochastic mapping approach described above to assign branches in the maximum clade credibility tree to the dietary categories. Rates through time were examined using the same time-slicing approach applied to disparity.

Character Saturation. Character saturation was tested by comparing the MORD morphological dissimilarity between pairs of taxa to the patristic morphological distance [97] as done by Brocklehurst et al. [60]. Patristic morphological distance represents the number of morphological changes that have occurred along the lineages leading to each taxon since their most recent common ancestor. This was inferred using an undated Bayesian phylogenetic analysis under the MkV model, carried out using the matrix of dental characters and constraining the topology to the maximum clade credibility tree identified by the FBD analysis described above. The branch lengths from this undated analysis represent amount of character change along phylogenetic lineages. The pairwise patristic distance between pairs of taxa were therefore inferred from this tree using the distTips function in the package adephylo [97]. Pairwise morphological dissimilarity and patristic morphological distances were compared via a loess regression, fitted in R.

Acknowledgements

I would like to thank David Button, Martin Ezcurra, David Ford, Christian Kammerer, Jasper Ponstein and Xiao-Chun Wu for helpful discussions and assistance with character scores. Peter Wagner and an anonymous reviewer provided helpful comments that greatly improved the manuscript. I am also grateful to Amy Scott Murray and Michael Day (Natural History Museum London), Daniella Schwarz (Museum für Naturkunde, Berlin), Nour-Eddine

Jalil (Muséum National d'Histoire Naturelle, Paris), Sifelani Jirah and Bernhard Zipfel (Evolutionary Studies Institute, Johannesburg), Carl Mehling and Mark Norrell (American Museum of Natural History, New York), Jessica Cundiff (Museum of Comparative Zoology, Harvard), and William Simpson (Field Museum of Natural History, Chicago) for access to specimens and assistance during visits to collections. NB’s research was funded by Deutsche Forschungsgemeinschaft grant number BR 5724/1-1, Palaeontological Association Research grant number PA-RG201901, and a Collections Study Grant from the American Museum of Natural History. Parts of this work were funded by the European Union’s Horizon 2020 research and innovation program 2014–2018 under grant agreement 677774 (European Research Council [ERC] Starting Grant: TEMPO) to RBJB.

Author Contributions

Conception: NB
Methods: NB, RBJB
Data: NB
Analysis: NB
Writing: NB, RBJB

Competing Interests Statement

The authors declare no competing interests.

Data Availability

All data analysed in this study are available in Supplementary Data 1-6

Code Availability

The analysis code is available in Supplementary Data 7

References

1. Frazzetta TH. Adaptive problems and possibilities in the temporal fenestration of tetrapod skulls. *J. Morph.* **125**, 145-158 (1968)
2. Janis CM, Keller JC. Modes of ventilation in early tetrapods: costal aspiration as a key feature of amniotes. *Acta Palaeontol. Pol.* **46**, 137-170 (2001)
3. Olson, E. C. Community evolution and the origin of mammals. *Ecology* **47**, 291-302. (1966)
4. DiMichele, W. A. & Aronson, R. B. The Pennsylvanian-Permian vegetational transition: a terrestrial analogue to the onshore-offshore hypothesis. *Evolution* **46**, 807-824 (1992)
5. Sues, H.-D. & Reisz, R. R. Origins and early evolution of herbivory in tetrapods. *Trends Ecol. Evol.* **13**, 141-145
6. Olson, E. C. The family Diadectidae and its bearing on the classification of reptiles. *Fieldiana, Geol.* **7**, 2-53 (1947)
7. Hotton, N., Olson, E. C. & Beerbower, R. In *Amniote Origins* (eds Sumida, S. S. & Martin K. L. M.) 207-264 (San Diego Academic Press, California, 1996)
8. Dodick, J. T. & Modesto. S. The cranial anatomy of the captorhinid reptile *Labidosaurikos meachami* from the Lower Permian of Oklahoma. *Palaeontology* **38**, 687-711 (1995).
9. Modesto, S.P. The skull of the herbivorous synapsid *Edaphosaurus boanerges* from the Lower Permian of Texas. *Palaeontology* **38**, 213-239 (1995)
10. King, G. M., Oelofsen, B. W. & Rubidge, B. S. The evolution of the dicyndont feeding system. *Zool. J. Linn. Soc.* **96**, 185-211 (1988)

11. Simpson, G. G. *Tempo and Mode in Evolution*. New York: Columbia University Press (1944).
12. Simpson, G. G. *The Major Features of Evolution*. New York: Columbia University Press (1953).
13. Schluter, D. *The Ecology of Adaptive Radiation*. Oxford University Press, Oxford (2000).
14. Givnish, T. J. Ecology of plant speciation. *Taxon* **59**, 1326-1366 (2010).
15. Givnish, T. J. Adaptive radiation versus ‘radiation’ and ‘explosive diversification’: why conceptual distinctions are fundamental to understanding evolution. *New Phytol.* **207**, 297-303 (2015).
16. Valentine, J. W. Determinants of diversity in higher taxonomic categories. *Paleobiology* **6**, 444-450 (1980)
17. Gould, S. J. The disparity of the Burgess Shale arthropod fauna and the limits of cladistic analysis: why we must strive to quantify morphospace. *Paleobiology* **17**, 411-423 (1991)
18. Briggs, D. E. G., Fortey, R. A., Wills, M. A. Morphological disparity in the Cambrian. *Science* **256**, 1670-1673 (1992)
19. Ingram, T., & Kai, Y. The geography of morphological convergence in the radiations of Pacific Sebastes rockfishes. *Am Nat* **184**, 115-131 (2014).
20. Harmon, L. J. et al. Early bursts of body size and shape evolution are rare in comparative data. *Evolution* **64**, 2385-2396 (2010)
21. Puttick, M. N. Mixed evidence for early bursts of morphological evolution in extant clades. *J. Evol. Biol.* **31**, 502-515 (2018)
22. Foote, M. Morphological disparity in Ordovician-Devonian crinoids and the early saturation of morphological space. *Paleobiology* **20**, 320-344 (1994)

23. Grant, V. *Plant Speciation*. Columbia University Press: New York (1981)

24. Blomberg, S. P., Garland, T. & Ives, A. R. Testing for phylogenetic signal in comparative data: behavioural traits are more labile. *Evolution* **57**, 717-745 (2003)

25. Mahler, D. L., Revell, L. J., Glor, R. E. & Losos, J. B. Ecological opportunity and the rate of morphological evolution in the diversification of Greater Antillean anoles. *Evolution* **64**, 2731-2745 (2010)

26. Harmon, L. J., Schulte J. A., Larson, A., & Losos, J. B. Tempo and mode of evolutionary radiation in iguanian lizards. *Science* **301**, 961-964 (2003)

27. Sidlaukas, B. Continuous and arrested morphological diversification in sister clades of characiform fishes: a phylomorphospace approach. *Evolution* **62**, 3135-3156 (2008)

28. Valentine, J.W., Erwin, D.H. Interpreting great developmental experiments: the fossil record. In (R.A. Raff, E.C. Raff, eds) *Development as an Evolutionary Process*. Pp 71-107 (New York: Liss, 1987)

29. Erwin, D.H., Valentine, J.W. Jablonski, D. The origin of animal body plans. *Am. Sci.*, **85**, 126–137 (1997).

30. Stanley, S. M. Paleozoic mass extinctions: shared patterns suggest global cooling as a common cause. *Am. J. Sci.* **288**, 334-352 (1988)

31. Bengston, S. The advent of animal skeletons. In (Bengston, S. ed) *Early Life on Earth*. Pp 412-425 (New York: Columbia University Press, 1994)

32. Anderson, P. S. L., Friedman, M. & Ruta, M. Late to the Table: Diversification of Tetrapod Mandibular Biomechanics Lagged Behind the Evolution of Terrestriality. *Integr. Comp. Biol.* **53**, 197-208

33. Olson EC. Jaw mechanisms: rhipidisteans, amphibians, and reptiles. *Am. Zool.* **1**, 205-215 (1961)

34. Carroll RL. Problems of the origin of reptiles. *Biol. Rev.* **44**. 393-432 (1968)

35. Pearson, M. R., Benson, R. B. J., Upchurch, P., Fröbisch, J., & Kammerer, C. F. Reconstructing the diversity of early terrestrial herbivorous tetrapods. *Palaeogeogr. Palaeoclim. Palaeoecol.* **372**, 42-49 (2013).

36. Brocklehurst, N., Ruta, M., Müller, J. & Fröbisch, J. Elevated extinction rates as a trigger for diversification rate shifts: early amniotes as a case study. *Sci. Rep.* **5**, 17104 (2015).

37. Brocklehurst, N., Kammerer, C. F., Benson, R. B. J. The origin of tetrapod herbivory: effects on local plant diversity. *Proc. R. Soc. B* **287**, 20200124 (2020)

38. Neige, P., Dera, G. & Dommergues, J. L. Adaptive radiation in the fossil record: a case study among Jurassic ammonoids. *Palaeontology* **56**, 1247-1261 (2013).

39. Simões, M. et al. The evolving theory of evolutionary radiations. *Trends Ecol. Evol.* **31**, 27-34 (2016).

40. Ruta, M., Angielczyk, K., Fröbisch, J., & Benton, M. J. Decoupling of morphological disparity and taxic diversity during the adaptive radiation of anomodont therapsids. *Proc. R. Soc. B* **280**, 20131071 (2013)

41. Brocklehurst, N. Rates of morphological evolution in Captorhinidae: an adaptive radiation of Permian herbivores. *PeerJ* **5**, e3200 (2017)

42. Grunert HR, Brocklehurst N, Fröbisch J. Diversity and disparity of Therocephalia: macroevolutionary patterns through two mass extinctions. *Sci. Rep.* **9**, 5063 (2019)

43. MacDougall, M.J., Brocklehurst, N., & Fröbisch, J. Species richness and diversity of parareptiles across the end-Permian mass extinction. *Proc. R. Soc. B* **286**, 20182572 (2018)

44. Brocklehurst, N. Morphological evolution in therocephalians breaks the hypercarnivore ratchet. *Proc. R. Soc. B* **286**, 20190590 (2019)

45. Van Valkenberg, B. Carnivore dental adaptations and diet: a study of trophic diversity within guilds. In *Carnivore Behaviour, Ecology and Evolution* (Gittleman, J. ed) pp 410-436 (Boston: Springer, 1989)

46. Jernvall, J., Hunter, J. P., & Fortelius, M. Molar tooth diversity, disparity, and ecology in Cenozoic ungulate radiations. *Science* **274**, 1489-1492 (1996).

47. Schwenk, K. *Feeding: Form, Function, and Evolution in Terrestrial Vertebrates*. San Diego: Academic Press (2000)

48. Müller J, Reisz RR. The phylogeny of early eureptiles: comparing parsimony and Bayesian approaches in the investigation of a basal fossil clade. *Syst. Biol.* **55**, 503-511 (2006)

49. Carroll RL, Gaskill P. *The Order Microsauria*. American Philosophical Society (1978)

50. Brocklehurst N, Reisz RR, Fernandez V, Fröbisch J. A re-description of *'Mycterosaurus' smithae*, an early Permian eothyridid, and its impact on the phylogeny of pelycosaurian-grade synapsids. *PLoS One*, **11**, e60156810 (2016)

51. Ford, D. P. Benson, R. B. J. The phylogeny of early amniotes and the affinities of Parareptilia and Varanopidae. *Nature Eco Evo* **4**, 57-65 (2020)

52. Ruta, M., Wagner, P. J., Coates, M. I. Evolutionary patterns in early tetrapods. I. Rapid initial diversification followed by decrease in rates of character change. *Proc. R. Soc. B.* **273**, 2113-2118 (2006)

53. Sahney S, Benton MJ. Recovery from the most profound mass extinction of all time. *Proc. R. Soc. B.* **275**, 759-765 (2008)

54. Brocklehurst N, Day MO, Rubidge BS, Fröbisch J. Olson's extinction and the latitudinal biodiversity gradient of tetrapods in the Permian. *Proc. R. Soc. B.* **284**, 20170231 (2017)

55. Brocklehurst N. An examination of the impact of Olson's extinction on tetrapods from Texas. *PeerJ*, **6**, e4767 (2018)

56. Brocklehurst N, Kammerer CF, Fröbisch J. The early evolution of synapsids, and the influence of sampling on their fossil record. *Paleobiology* **39**, 470-490 (2013)

57. Fröbisch J. Composition and similarity of global anomodont-bearing tetrapod faunas. *Earth-Sci. Rev.* **95**, 119-157 (2009)

58. MacDougall MJ, Brocklehurst N, Fröbisch J. Species richness and disparity of parareptiles across the end-Permian mass extinction. *Proc. R. Soc. B.* **286**, 20182572 (2018)

59. Brocklehurst, N., Panciroli, E., Benevento, G. L., Benson, R. B. J. Mammaliaform extinctions as a driver of the morphological radiation of Cenozoic mammals. *Curr. Biol.* (In Press)

60. Wagner, P. J. Exhaustion of morphologic character states among fossil taxa. *Evolution* **54**, 365-386 (2000)

61. Nee, S., Mooers, A. O. & Harvey, P. H. Tempo and mode of evolution revealed from molecular phylogenies. *Proc. Natl. Acad. Sci.* **89**, 8322-8326 (1992)

62. Seehausen, O. African cichlid fish: a model system in adaptive radiation research. *Proc. Natl. Acad. Sci.* **273**, 1987-1998 (2006)

63. Phillimore, A. B. & Price, T. D. Density-dependent cladogenesis in birds. *PLoS Biol.* **6**, e71 (2008)

64. McPeek, M. A. The ecological dynamics of clade diversification and community assembly. *Am. Nat.* **172**, E270-E284 (2008)

65. Rabosky, D. L. & Lovette, I. J. Explosive evolutionary radiations: decreasing speciation or increasing extinction through time? *Evolution* **62**, 1866-1875 (2009)

66. O'Donovan, C., Meade, A. & Venditti, C. Dinosaurs reveal the geographic signal of an evolutionary radiation. *Nat. Ecol. Evol.* **2**, 452-458 (2018)

67. Foote, M. The evolution of morphological diversity. *Annu. Rev. Ecol, Syst.* **28**, 129-152 (1997)

68. Hughes, M., Gerber, S., & Wills, M. A. Clades reach highest morphological disparity early in their evolution. *Proc. Nat. Acad. Sci.* **110**, 13875-13879

69. Sidlaukas, B. Testing for unequal rates of morphological diversification in the absence of a detailed phylogeny: a case study from characiform fishes. *Evolution* **61**, 299-316.

70. Benson, R. B. J., Hunt, G., Carrano, M. T., & Campione, N. Cope's rule and the adaptive landscape of dinosaur body size. *Palaeontology* **61**, 13-48 (2018)

71. Hill, J. J., Puttick, M. N., Stubbs, T. L., Rayfield, E. J., Donoghue, P.C. Evolution of jaw disparity in fishes. *Palaeontology* **61**, 847-854 (2018)

72. Simões T. R., Vernygora, O., Caldwell, M. W., Pierce, S. E. Megaevolutionary dynamics and the timing of evolutionary innovation in reptiles. *Nat Comms* **11**, 322 (2020)

73. Lee, M. S. Y., Soubrier, J., Edgecombe, G. D. Rates of phenotypic and genomic evolution during the Cambrian Explosion. *Curr. Biol.* **23**, 1889-1895 (2013)

74. Halliday, T. J. D., Upchurch, P., & Goswami, A. Eutherians experienced elevated evolutionary rates in the immediate aftermath of the Cretaceous-Paleogene mass extinction. *Proc. R. Soc. B.* **283**, 20153026 (2016)

75. Oyston, J. W., Hughes, M., Gerber, S., Wills, M. A. Why should we investigate the morphological disparity of plant clades? *Ann. Bot.* **117**, 859-879

76. Lofgren, A. S., Plotnick, R. E., Wagner, P. J. Morphological diversity of Carboniferous arthropods and insights on disparity patterns through the Phanerozoic. *Paleobiology* **29**, 349-368 (2003)

77. Condamine, F. L., Clapham, M. E., Kergoat, G. J. Global patterns of insect diversification: towards a reconciliation of fossil and molecular evidence. *Sci Rep*, **6**, 19208 (2016)

78. Labandeira, C. C. & Sepkoski, J. J. Insect diversity in the fossil record. *Science*, **261**, 310-315 (1993)

79. Schaeffer, B. Rates of evolution in the coelacanth and dipnoan fishes. *Evolution*, **6**, 101-111 (1952)

80. Lloyd, G. T., Wang, S. C., & Brusatte, S. L. Identifying heterogeneity in rates of morphological evolution: discrete character change in the evolution of lungfish (Sarcopterygii; Dipnoi). *Evolution*, **66**, 330-348 (2012).

81. Wright, D. F. Phenotypic innovation and adaptive constraints in the evolutionary radiation of Palaeozoic crinoids. *Sci Rep*, **7**, 13745 (2017).

82. Wagner, P. J. Early bursts of disparity and the reorganization of character integration. *Proc R Soc B*, **285**, 20181604. (2018).

83. Slater, G. J., Harmon, L. J., & Alfaro, M. E. Integrating fossils with molecular phylogenies improves inference of trait evolution. *Evolution*, **66**, 3931-3944 (2012).

84. Brocklehurst, N., Dunne, E. M., Cashmore, D. D. & Fröbisch, J. Physical and environmental drivers of Paleozoic tetrapod dispersal across Pangaea. *Nat. Comm.* **9**, 5216 (2018)

85. Pardo, J. D., Szostakiwskyj, M., Ahlberd, P. E. & Anderson, J. S. Hidden morphological diversity among early tetrapods. *Nature* **546**, 642-645 (2017)

86. Baum, B. Combining trees as a way of combining data sets for phylogenetic inference, and the desirability of combining gene trees. *Taxon* **41**, 3-10 (1992)

87. Ragan, M. Phylogenetic inference based on matrix representation of trees. *Mol. Phylogenet. Evol.* **1**, 53-58 (1992)

88. Salamin, N., Hodkinson, T. R., & Savolainen, V. Building supertrees: an empirical assessment using the grass family (Poaceae). *Syst. Biol.* **51**, 134-150 (2002)

89. Goloboff, P. A., Farris, J. S., & Nixon, K. C. TNT, a free program for phylogenetic analysis. *Cladistics* **24**, 774-786

90. Heath, T. A., Huelsenbeck J. P., & Stadler T. The fossilized birth-death process for coherent calibration of divergence-time estimates. *Proc. Nat. Acad. Sci.* **111**, E2957-2966 (2014)

91. Ronquist, F. & Huelsenbeck, J. P. MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* **19**, 1572-1574 (2003)

92. Pyron, R. A. A likelihood method for assessing molecular divergence time estimates and the placement of fossil calibrations. *Syst. Biol.* **59**, 185-194 (2010)

93. Lloyd GT. Estimating morphological diversity and tempo with discrete character-taxon matrices: implementation, challenges, progress and future directions. *Biol. J. Linn. Soc.* **118**, 131-151 (2016)

94. Guillerme, T. & Cooper, N. Time for a rethink: time sub-sampling methods in disparity-through-time analyses. *Palaeontology* **61**, 481-493 (2018)

95. Bollback, J. P. SIMMAP: stochastic character mapping of discrete traits on phylogenies. *BMC Bioinformatics* **7**, 88 (2006)

96. Revell, L. J. phytools: an R package for phylogenetic comparative biology (and other things). *Methods Ecol. Evol.* **3**, 217-223. (2012)

97. Jombart, T., Balloux, F. & Dray, S. Adephylo: new tools for investigating the phylogenetic signal in biological traits. *Bioinformatics* **26**, 1907-1909 (2010)

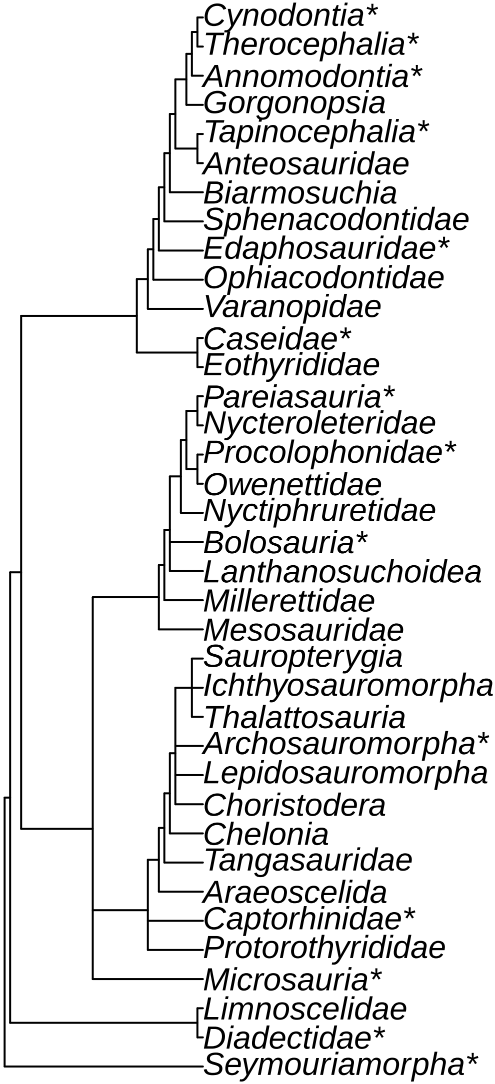
Figure Captions

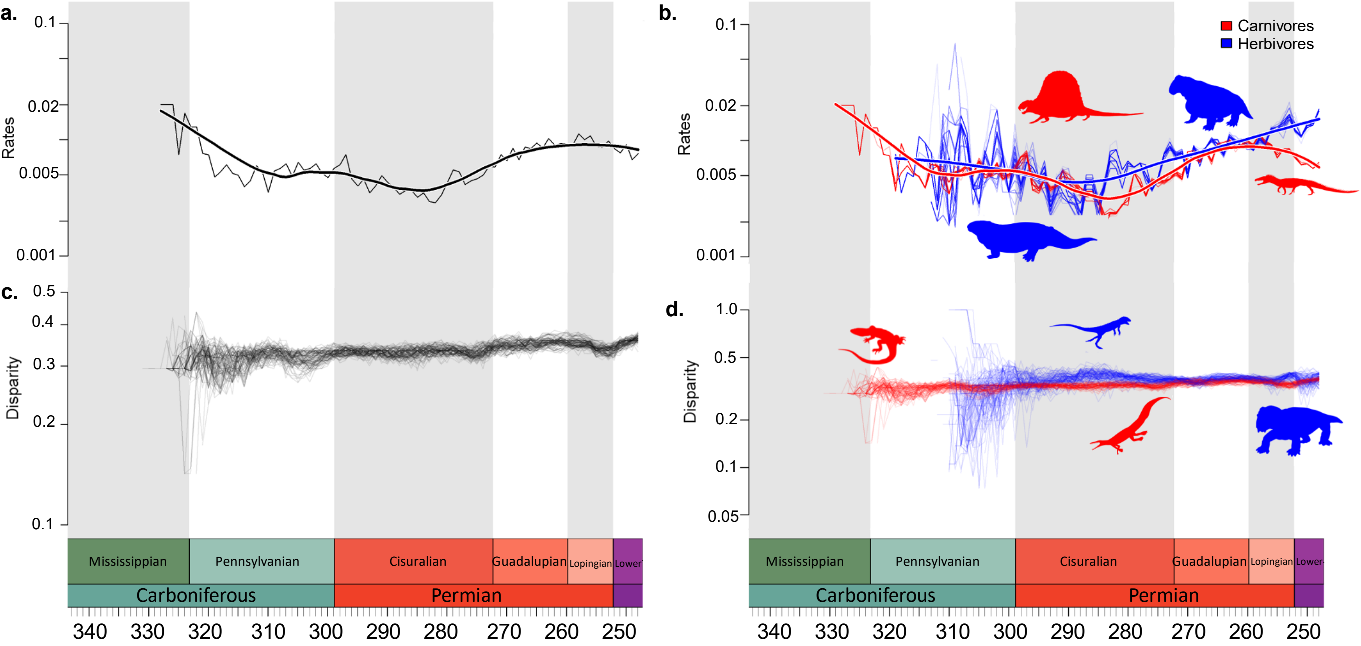
Figure 1: Summary of the supertree used in the study. Clades marked * evolved, or have been suggested to have evolved, herbivory during the period of time under study (Carboniferous-Early Triassic)

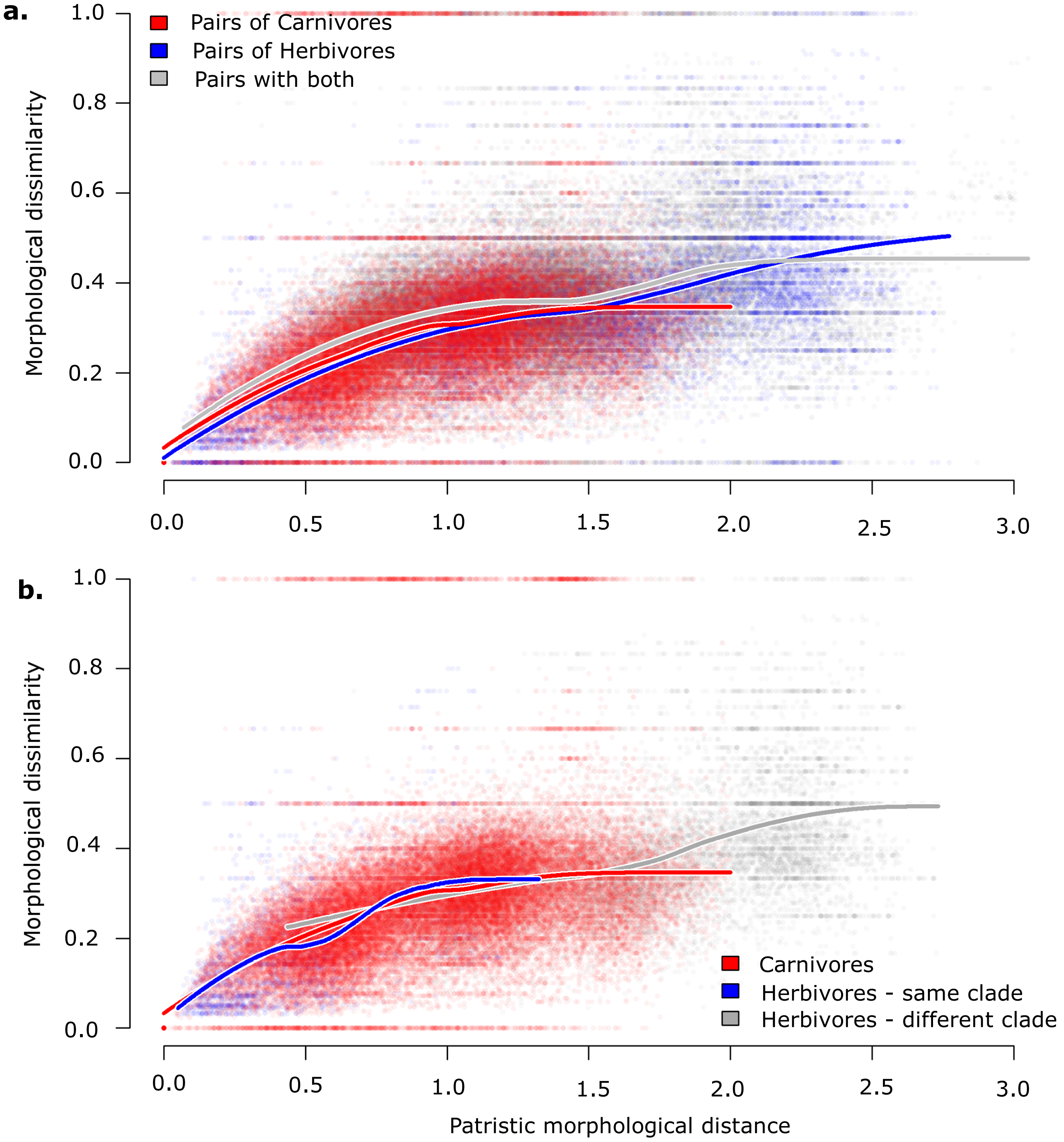
Figure 2: A) Rates of dental evolution in amniotes, inferred from maximum clade credibility tree produced by the fossilised birth death model. The thin line represents median rate of each branch at each time slice. The thick line represents a Loess fitted regression curve. B) Rates of dental evolution of faunivorous (red) and herbivorous (blue) amniotes. Thin lines represent rates inferred from the maximum clade credibility tree. Each thin line represents results from one stochastic map used to assign a diet to each branch. The thick lines represent Loess fitted regression curves. C) Dental disparity of amniotes (median pairwise distance between all taxa present at each time slice. Each line represents a disparity curve inferred from a randomly selected tree from the Bayesian posterior distribution. Each line represents a disparity curve inferred from a randomly selected tree from the Bayesian posterior distribution. D) Dental disparity of faunivorous (red) and herbivorous (blue) amniotes. Each line represents Each thin line represents results from one stochastic map used to assign a diet to each branch in one selected tree from the Bayesian posterior distribution. Silouettes from phylopic.or (public domain), not to scale.

Figure 3: Comparisons of patristic morphological distances (x axis) to pairwise morphological dissimilarity for faunivorous and herbivorous amniotes. Each point represents

a comparison of distances between two taxa. The lines represent loess fitted regression curves. A) Red points represent comparisons between pairs of faunivorous amniotes; blue points represent pairs of herbivorous amniotes. B) Blue points represent comparisons between pairs of herbivores from within the same clade; grey points represent comparisons between pairs of herbivores from different clades.


Cynodontia*
Therocephalia*
Annomodontia*
Gorgonopsia
Tapinocephalia*
Anteosauridae
Biarmosuchia
Sphenacodontidae
Edaphosauridae*
Ophiacodontidae
Varanopidae
Caseidae*
Eothyrididae
Pareiasauria*
Nycteroleteridae
Procolophonidae*
Owenettidae
Nyctiphruretidae
Bolosauria*
Lanthanosuchoidea
Millerettidae
Mesosauridae
Sauropterygia
Ichthyosauromorpha
Thalattosauria
Archosauromorpha*
Lepidosauromorpha
Choristodera
Chelonia
Tangasauridae
Araeoscelida
Captorhinidae*
Protorothyrididae
Microsauria*
Limnoscelidae
Diadectidae*
Seymouriamorpha*


a.
b.
c.
d.
Rates
Rates
Disparity
Disparity
0.1
0.02
0.005
0.001
0.1
0.02
0.005
0.001
0.5
0.4
0.3
0.2
0.1
1.0
0.5
0.2
0.1
0.05
🟥 Carnivores
🟦 Herbivores
Mississippian
Pennsylvanian
Cisuralian
Guadalupian
Lopingian
Lower
Carboniferous
Permian
Mississippian
Pennsylvanian
Cisuralian
Guadalupian
Lopingian
Lower
Carboniferous
Permian
340
330
320
310
300
290
280
270
260
250
340
330
320
310
300
290
280
270
260
250


a.
Pairs of Carnivores
Pairs of Herbivores
Pairs with both
Morphological dissimilarity
0.0
0.2
0.4
0.6
0.8
1.0
0.0
0.5
1.0
1.5
2.0
2.5
3.0
b.
Carnivores
Herbivores - same clade
Herbivores - different clade
Morphological dissimilarity
0.0
0.2
0.4
0.6
0.8
1.0
0.0
0.5
1.0
1.5
2.0
2.5
3.0
Patristic morphological distance

Supplementary Dataset 1

Character List

Because the matrices forming the basis of this character list are designed to examine relationships of specific clades, many include characters with specific morphological descriptions that are not relevant to the taxa in other clades. The most obvious example is the use of incisors/incisiform teeth and molars/molariform teeth in therapsid/diadectid character lists, which are not relevant to taxa without mammal-like differentiation. Characters referring to canine, precanine and postcanine teeth are not relevant to taxa without a caniniform tooth/region. In the case of characters referring to the canine and caniniform teeth the obvious solution is to score taxa not possessing such teeth as inapplicable. But for precanine/incisor teeth and molar/postcanine teeth, one must attempt to avoid issues of redundancy and character non-independence with characters referring to premaxillary/maxillary teeth or lateral dentition in general. One cannot, for example, have characters referring to the number/morphology of incisors, score taxa without incisiform teeth as inapplicable, but then include a character referring to the number of premaxillary teeth so that variation in taxa without heterodonty may be accounted for.

In compiling this character list, the following practices are observed when choosing and rewording characters. These practices have been formulated with eco-functionality in mind, rather than homology (as would be used when formulating characters used in phylogeneteic analysis).

- The upper teeth will be defined by the bone in which they are implanted (premaxillary and maxillary). With dentary teeth, this becomes more problematic, as incisors, canines and molars are all found within a single bone, so different practices are applied (see below).
- The maxillary teeth are divided between caniniform and non-caniniform; if a caniniform region/tooth is not present, all characters relating to caniniforms are scored as inapplicable, and all maxillary teeth are treated as non-caniniforms
- Canines/caniniforms are defined as being a single tooth or region with teeth noticeably longer than those before or after them in the maxilla. This definition does not include precanine maxillary teeth, which in some descriptions of therapsids in the 1910s-1940s were referred to as canines/micro-canines.
- Characters referring to molars/molariform teeth have been reworded to refer to non-caniniform maxillary teeth. In taxa where there is no canine or caniniform region this will account for all maxillary teeth. In cases where there is a canine or caniniform region this will account only for postcanines
- Some characters specify postcanine maxillary teeth. No characters in the source matrix referred to the precanines except one referring to their number. Since characters relating to the total number of maxillary teeth and the number of postcanine maxillary teeth will by proxy allow variation in the number of precanine maxillary teeth to be accounted for, this character was ignored.
- Characters referring to upper incisors have been reworded and if necessary rescored to relate to teeth implanted in the premaxilla.

- Characters relating to overall tooth morphology are no longer included. Variation in tooth morphology in taxa with no heterodonty or heterodonty only in that they have canines will be accounted for by characters referring to the morphology of the non-canine maxillary teeth and the premaxillary teeth being scored for the same character state.
- Characters relating to dentary incisors/molars are reworded and rescored as precanine and postcanine teeth and are scored as inapplicable in taxa without a dentary canine. If a character refers to both upper and lower incisors or anterior dentary teeth, the dentary teeth included in the scoring of the character are those that cover a portion of the dentary equivalent to the fraction of length of the tooth row covered by premaxillary teeth. For characters referring to both upper and lower molars the same logic is applied; the dentary teeth included in the scoring of the character are those that cover a portion of the dentary equivalent to the fraction of length of the tooth row covered by non-canine maxillary teeth
- Palatal tooth characters are arranged by tooth field. This has the biggest effect when considering the teeth on the anterior ramus of the pterygoid. Sometimes this tooth field will extend laterally onto the ectopterygoid and palatie, or even anteriorly onto the vomer, but this still counts as the anterior ramus tooth field and so should not be considered as representing the presence of an ectopterygoid/palatine/vomerine tooth field. Tooth fields in close proximity or which seem confluent with each other may be considered separate tooth fields if there is a noticeable and abrupt shift in size/organisation coinciding with the transition to a new bone. E.g. palatal dentition of *Seymouria sanjuanensis* – the tooth field of the anterior ramus of the pterygoid forms numerous denticle rows which extend anteriorly onto the posteriormost portion of the vomer. These posterior vomerine denticles are of similar size to the pterygoid denticles and clearly form rows that are extensions of those on the pterygoid. Therefore, these denticles are still considered part of the tooth field on the anterior ramus of the pterygoid (a character is included indicating whether the pterygoid tooth field extends onto the vomer). However, there are also denticles on the anteriormost vomer. These are separated from those more posteriorly by a noticeable gap, and also are not organised in rows in the same was that the posterior denticles are. Therefore, these represent a separate, vomerine tooth field. Also present are vomerine fangs, which are again treated separately.
- For purposes of coding for the presence or absence of teeth on the coronoids, the pattern of loss of the individual coronoids is judged to be from anterior to posterior i.e. where two coronoids are present, they are the posterior and middle coronoid; where one is present it is the posterior coronoid.

1) Tooth implantation: subthecodont (=protothecodont) (0); ankylothecodont (1); pleurodont (2); acrodont (3); thecodont (4)
    - [1] character 299

2) Tooth pedicely: absent (0); present (1).
    - [2], character 80

3) Number of lateral premaxilliary teeth: premaxilla edentulous (0); 1-4 (1); 5-6 (2); 7-20 (3); >20 (4)
    - Boundaries between states were determined by K means clustering
    - Lateral premaxillary teeth are specified due to the suggestion of premaxillary teeth on the medial process in *Haptodus* [3]

4) Number of lateral maxillary teeth: maxilla edentulous (0); 1-4 (1); 5-11 (2); 12-21 (3); >22 (4)
    - Boundaries between states were determined by K means clustering
    - In taxa with multiple tooth rows, only the lateral-most row is counted

5) Maxillary caniniforms present as single large member of tooth series (0); present as two or more enlarged members of tooth series (1); absent (2); or present as tusk (3).
    - Modified from [4], character 22

6) Number of postcanine maxilliary teeth: 0 (0); 1-4 (1); 5-9 (2); 10-17 (3); >18 (4)
    - Boundaries between states were determined by K means clustering

7) Number of tooth rows in the maxilla: one (0); two to four (1); five (2); six or more (3), irregular pavement (4).
    - States from [5], character 9, and [6], character 133

8) If multiple rows present, number of teeth on maxillary dental field is: Less than 40 (0); 40 or more (1).
    - Modified from [5], character 7

9) If multiple rows of maxillary teeth present, double row of teeth extends far anteriorly on maxilla: absent (0); present (1).
    - Modified from [5], character 11

10) Number of lateral dentary teeth: dentary edentulous (0); 1-9 (1); 10-20 (2); 20-34 (3); >35 (4)
    - Boundaries between states were determined by K means clustering
    - In taxa with multiple tooth rows, only the lateral-most row is counted

11) Dentary, number of tooth rows: one (0); two (1); more than two (2)
    - [1], character 279

12) Dentary: caniniform teeth absent (0); caniniform region present anteriorly (1); Single caniniform tooth present anteriorly (2)
    - [5], character 56

13) Dentary postcanines: present (0); absent (1).
    - [7], character 44

14) Lower precanine number: 0 (0); 1 (1); 2-3 (2); 4 (3); >5 (4)
   - Boundaries between states were determined by K means clustering

15) Dental tooth wear: absent (0); present, modest (1); present, saddle-shaped (2); present, produces cutting ridges (3).
   - Modified from [5], character 12
   - Character state 3 new

16) Premaxillary tooth proportions: anteriormost tooth approximately same size as other teeth (0); anteriormost tooth reduced relative to other premaxillary teeth (1); anteriormost tooth larger than other premaxillary teeth (2).
   - Modified from [2], character 289

17) Premaxilliary tooth morphology: relatively straight and conical (0); spatulate (1); transversely compressed and recurved (2); chisel shaped (3); bulbous and ogival (4), with broad flattened tip (5); fan shaped (6);
   - States drawn from multiple matrices
   - State 5 new, representing the unique morphology found in *Anomocephalus*
   - The distinction between spatulate teeth and fan-shaped teeth has not been robustly defined in previous matrices (it has mostly been used in analyses of pareiasaurs) beyond representing antero-poserior lengthening of the tooth. A fan-shaped tooth is here distinguished from a spatulate tooth in the crown being antero-posteriorly longer than high

18) Location of premaxillary teeth lateral (0), medial (1)
   - Modified from [4], character 8

19) Premaxilla, orientation of the tooth series of the occlusal surface of premaxilla in ventral view: approximately parasagittal (0); strongly transverse and anterior teeth covering each other in lateral view (1)
   - [1], character 43

20) Premaxillary teeth with longitudinal facets or fluting: absent (0); present (1).
   - Modified from [8], character 97

21) Procumbant anterior teeth: absent (0); present in the lower jaw only (1); present in both upper and lower jaws (2); Present in upper jaw only (3)
   - Modified from [9], character 29
   - Character state 3 new

22) Premaxilliary and anterior dentary teeth possess serrated cutting margins (0); are smoothly ridged (1).
   - Modified from [10], character 100

23) Premaxillary teeth intermesh with anterior dentary teeth: absent (0), present in anteriormost premaxillary teeth (1), present in all premaxillary teeth (2).

    - Modified from [11], character 62

24) Heel/ridge on premaxillary and anterior dentary teeth: absent (0); weak heel (1); strong heel (heel width 30% or more larger than the base of the tooth apex) (2); ridge narrowing towards towards the tip (3)
    - Modified from [12], character 8

25) Premaxillary fangs: absent (0); present (1).
    - Modified from [13], character 66
    - Distinct from character 16; premaxillary fang cannot be anteriormost tooth
    - Fang cannot be accounted for by a general trend towards larger teeth posteriorly; fangs have to be exceptionally enlarged tooth/teeth among other homogenously sized teeth
    - Since enlarged anterior dentary tooth/teeth is accounted for by dentary canine character, dentary fangs are not included, unlike in the original character.

26) Diastema between premaxillary and maxillary teeth: absent (0); present (1)
    - [12], character 26

27) Maxilla, canniniform teeth size: less than twice the height of non-canniniform teeth (0); more than twice the height (1).
    - [14], character 35
    - Where there are multiple caniniformes, use the largest

28) Maxillary caniniform curvature: straight or slightly recurved (<50°) (0); strongly recurved (>50°) (1); curves anteriorly (2)
    - Modified from [12], character 7
    - State 2 new

29) Maxillary caniniform serrations: present (0), absent (1).
    - [10], character 106

30) Prominent ridge running vertically along the labial surface of the dominant upper caniniform tooth: absent (0); present (1).
    - Modified from [8], character 101
    - The original formulation of this character related to the presence of a deep groove on the canine and was supposed to represent the morphology observed in *Euchambersia*. However, the canine of *Euchambersia* does not possess a deep groove, rather a prominent ridge on the labial surface. The earliest descriptions make this clear both in drawings and text [15-17]. The assumption that the canine was grooved appears to have derived from a drawing by Lehman [18] where the shading makes the ridge appear to be a prominent groove.

31) Lower caniniforms: fits into choana (0); into fossa roofed by premaxilla and maxilla (1); passes anterior and external to upper canine (2); passes through fossa in skull roof so is visible dorsally (3).
    - Modified from [11], character 66
    - State 3 new

32) Upper and lower canines: without heels (0); small heels present (1).
    - [11], character 67

33) Postcanine diastema on upper jaw: absent (0), present (1)
    - [11], character 68

34) Upper postcanine teeth confluent with premaxillary row medial to canine: absent (0), present (1).
    - [11], character 71

35) Maxilla, posterior extent of lateral dentition: anterior to posterior orbit margin (0); ventral to postorbital bar (1); posterior to postorbital bar (2); anterior to orbital midlength (3).
    - [14], character 30
    - In the absence of a temporal fenestra, the postorbital represents the postorbital bar

36) Number of maxillary positions ventral to the jugal: Three or less (0); four or more (1)
    - [6], character 138

37) Axis of posterior part of maxillary tooth row: directed lateral to subtemporal fossa (0), directed toward centre of fossa (1), directed toward medial rim of fossa and curved (2)
    - [10], character 107

38) Maxilla, secondary enlargement of teeth posterior to canniniforms, if present: absent (0); present (1).
    - [14], character 39

39) Maxilla, non-caniniform tooth morphology: relatively straight and conical (0); spatulate (1); transversely compressed and recurved (2); chisel shaped (3); bulbous and ogival (4); molariform, bucco-lingually expanded (5); sectorial with incipient lingual cingulum (6); sectorial with a well-developed lingual cingulum (7), fan-shaped (8).
    - States drawn from multiple matrices
    - See note from character 17 on the distinction between spatulate and fan-shaped teeth

40) Non caniniform maxillary dentition, recurvature: completely absent (0); at least slightly recurved (1), strongly recurved, apex approximately 80-90 degrees from vertical (2).

    - Modified from [14], character 41

41) Posterior-most maxillary teeth canted posterolaterally relative to more anterior teeth: absent (0); present (1)
    - Modified from [12], character 6

42) Non-caniniform maxillary teeth located near lateral margin of maxilla (0); located more medially (1)
    - Modified from [4], character 23

43) Posterior dentary teeth: same size as non-caniniform maxillary teeth (0); larger than non-caniniform maxillary teeth (1); smaller than non-caniniform maxillary teeth (2)
    - Modified from [2], character 333

44) Posterior extent of mandibular and maxillary tooth rows: subequal (0); maxillary teeth extending further posteriorly (1); mandibular teeth extending further posteriorly (2)
    - Modified from [1], character 298
    - Character state 2 new

45) Non caniniform maxillary and posterior dentary alveolar ridges: straight (0); twisted (helical), distal teeth inclined laterally (1).
    - [14], character 31

46) Serrated carinae on the non-caniniform maxillary and posterior dentary teeth: absent (0); distinctly present on the mesial margin only (1); distinctly present on the distal margin only (2); present and distinct on both margins (3)
    - Modified from [1], character 304
    - Character state 1 new

47) Non-caniniform maxillary teeth: serrated cusps antero-posteriorly orientated relative to the long axis of the tooth (0); serrations angled anterolingually-posterolabially (1).
    - Modified from [12] character 4

48) Non-caniniform maxillary teeth, denticles: density high, 14–24 denticles/mm (0); denticle density low, <10 serrations/mm (1).
    - Modified from [14], character 43
    - Taxa without serrations scored as inapplicable

49) Presence of denticles on non-caniniform maxillary and dentary teeth independent of carinae: absent (0), present (1).
    - [14] character 245

50) Non caniniform maxillary teeth, number of apical cusps; one (0); central cusp, additional labial and lingual cusps poorly developed, represented by shoulders (1); labial and lingual cusps well developed in addition to central cusp (2); transverse crest with two cusps (3); transverse crest with more than two cusps (4); two-four cusps

arranged longitudinally (5); five-seven cusps arranged longitudinally (6); more than seven cusps arranged longitudinally (7); multiple cusps arranged around the edge of the crown (8)

- States drawn from multiple matrices
- Character state 8 new

51) If present, position of upper transverse cusp row on crown of non-caniniform maxillary teeth: on anterior half of crown (0); midcrown almost to posterior margin (1)

- Modified from [10], character 114

52) If transverse row present, central cusp is: absent (0), midway between buccal and lingual cusps (1), close to labial cusp (2)

- Modified from [10], character 115

53) If multiple cusps arranged longitudinally, is upper posterobuccal accessory cusp: present (0), absent (1).

- [10], character 118

54) If multiple cusps arranged longitudinally, is upper anterolingual accessory cusp: absent (0), present (1).

- [10], character 119

55) Where present, upper anterior transverse (cingulum) ridge: low (0), high (1).

- Modified from [10], character 120

56) Cutting ridge on lingual surface of molars: absent (0), present (1).

- [10], character 121

57) Non-caniniform maxillary and dentary teeth orientation: vertical (0); turned lingually (1); turned labially (2)

- Modified from [2], character 64
- Character state 2 new

58) Dentary: tooth row greater than 50% of total jaw length (0); tooth row less than 50% of total jaw length (1)

- [2], character 76

59) Teeth present on dorsal surface of dentaries (0); medially displaced, sometimes on a swelling or shelf (1)

- [4], character 117

60) Lingual cingulum in posterior dentary teeth: absent (0), small (1), well developed (2).

- Modified from [10], character 124

61) If present, cingulum on dentary teeth denticulated: absent (0) or present (1)

- Modified from [4], character 119

62) Anterior dentary teeth: not with distinct shape from rest of tooth row (0); distinct division in morphology similar to that observed between premaxillary and maxillary teeth (1); bulbous base giving teardrop-shaped morphology (2); enlarged and incisiform (3).
    - States taken from [4], character 120, and [14], character 33
    - Character state 1 new
    - Size related heterodonty not included in this.

63) Jaw symphysis terminates in dorsal platform bearing the anterior teeth elevated above level of posterior dentary ramus (0); Symphyseal region of lower jaw smoothly rounded and at same level as rest of dentary ramus in lateral view (1), with an upturned beak that is raised above the level of the dorsal surface of the jaw rami and has a scooped-out depression on its posterior surface (2), drawn into a sharp, spiky beak (3), or shovel-shaped beak with a rounded or squared-off edge and a weak depression on its posterior surface (4).
    - Modified from [4], character 121

64) Dentary, canniniform tooth size: less than twice the height of post-canine teeth (0); more than twice the height (1).
    - Modified from [19], character 22

65) Posterior dentary teeth: undifferentiated (0), differentiated into premolariforms and molariforms (1).
    - Modified from [10], character 109
    - This differentiation is called “premolariforms and molariforms” for want of anything better but scoring of character state 1 should not be limited to those with mammal-like differentiation. The character is intended to indicate any posterior shape (not size) heterogeneity, and character 112 is used to indicate the actual morphology

66) Tall, dorsally-convex cutting blade on medial edge of dorsal surface of dentary absent (0) or present (1)
    - [4], character 126

67) Degree of molarization of the largest preserved, midseries dentary cheek teeth: absent (0); low (mediolateral width:anteroposterior length and mediolateral width:height not >0.5) (1); high (2)
    - [9], character 30

68) Number of lower cusps in transverse row, if present: 1 (0), 2 (1), 3 or more (2)
    - [10], character 123

69) Widest lower cusp in transverse row, where present: lingual (0), buccal (1).
    - [10], character 127

70) Lower posterior basin: absent (0), present (1).
  - [10], character 126

71) Posterior dentary teeth occlusion with upper teeth: single-sided overlap (0); flat occlusion (1); occlude with palatine (2); maxillary teeth pass into spaces between dentary teeth (3)
  - [1], Character 280
  - Character states 2 and 3 new

72) Parasphenoid cultriform process with shagreen (0), with patch of denticles (1), with denticle row or rows (2), smooth (3).
  - [20], character 55

73) Parasphenoid, teeth on ventral plate: present along edges (0); present across width of plate, forming central denticle field (1); teeth absent from ventral plate (2).
  - Modified from [14]

74) Vomerine tooth field: small denticles (0); large conical teeth (1); absent (2)
  - Character states taken from [21] character 121 and [20] character 59
  - Note, vomerine tooth field is separate from anterolateral field of pterygoid, which sometimes extends onto the vomer

75) Organisation of vomerine teeth: Single row (0); single row but multiple teeth anterior to contact with pterygoid (1); multiple rows or field (2)
  - Modified from [1], character 187

76) Vomerine fang(s): absent (0); present (1)
  - Modified from [22], character 69

77) Palatine dental field: multiple rows with numerous teeth/denticles (0); single row with four or more teeth (1), single row with less than four teeth (2); restricted to palatine boss (3); absent (4).
  - Modified from [13], character 140
  - Note: palatine dental field is separate from anterolateral teeth of pterygoid, which sometimes extend onto palatine
  - Character state 3 new

78) Palatine teeth size: larger than marginal teeth (0); equal to marginal (1); smaller than marginal (2).
  - [2], character 81
  - To judge from the original scorings in [2], palatine fangs are included in this, so this policy is followed here

79) Dentition on palatine boss, if present: extensive (0); elongate single row (1); a few teeth in a restricted position (2).

- Modified from [7], character 10

80) Presence (0) or absence (1) of fangs on palatine
    - [20], character 62

81) Bone texture of the palatine: primarily smooth, without evidence of keratinized covering (0); relatively smooth but with fine pitting and texturing suggestive of a keratinized covering (1); rugose and textured (2).
    - [4], character 76

82) Ventral surface of pterygoid palatal ramus with radiating, densely spaced low ridges with denticle rows (0), with radiating sharp ridges with denticles (1), with densely spaced denticle rows radiating from posterior midlength of palatal ramus (2), with teeth in two distinct fields, one extending anteriorly, one laterally (3), with a single tooth field covering the width of the process (4), dense covering of large teeth forming tooth plates (5); edentulous (6); teeth restricted to medial ridge or boss (7); with teeth in three distinct fields, one directed anteriorly, one anterolaterally towards suture with palatine, one laterally towards suture with ectopteygoid (8).
    - States drawn from [14], character 122, [20], character 47 and [1], character 195
    - Character state 7 and 8 new; note the midline field (T4) is treated separately from character state 7 (see notes for character 86). State 8 represents condition seen in *Brouffia* and *Australothyris*).

83) Pterygoid, number of rows on palatal process tooth field T2 (laterally extending): more than two or do not dispose on distinct rows (0); two rows parallel to each other (1); single row (2)
    - Modified from [1] character 196

84) Pterygoid, number of rows on palatal process tooth field T3 (medial): more than two or not disposed in distinct rows (0); two parallel rows (1); single row (2)
    - Modified from [1] character 197

85) Pterygoid, a row of teeth ramus on the medial edge of the anterior ramus (T4) larger than those the anterior ramus fields: absent (0); present (1)
    - Modified from [1], character 199
    - The identification in the literature of field T4 appears thus far to have only been based on comparisons of size and arrangement relative to the T3 field; there don’t appear to be any archosaurs with T4 but no T3.
    - This does not hold for other clades. Some pelycosaurs e.g. *Cotylorhynchus*, have uniformly large palatal teeth, with their T3 field forming a single row along the midline, so in appearance (and potentially functionally) their T3 field is more similar to the T4 field. The plesiomorphic diadectid *Orobates* has no teeth on the palatal ramus but a single row of large teeth along the pterygoid midline. While potentially analogous to the T4 field in archosaurs, the difficulty is that in more derived diadectids this is reduced to a midline of

denticles, so a distinction of being large and fanglike is not necessarily useful as far as homology goes.
- Solution: T4 field is defined relative to the T3 field, as previously in the archosaur literature i.e. this character can only be scored if the T3 field is present, otherwise should be treated as inapplicable. Character state 7 has been added to character 83 to represent the condition seen in diadectomorphs and some therapsids where the only pterygoid teeth are restricted to a midline ridge or boss (the T4 field in archosaurs is not elevated in such a way)

86) Pterygoid, teeth in anterolaterally oriented field: extends posteromedially to basicranial area (0); does not extend as far as basicranial area (1)
    - Modified from [14], character 122

87) Pterygoid, teeth in anteriorly oriented field: extends anteriorly onto vomer (0); does not extend onto vomer (1)
    - Modified from [20], character 70

88) Pterygoid, teeth in anterolaterally oriented field: extends latterally onto palatine (0); does not extend onto palatine (1)
    - Modified from [20], character 71

89) Pterygoid, teeth in anterolaterally oriented field: extends laterally onto ectopterygoid (0); does not extend onto ectopterygoid (1)
    - Modified from [20], character 72

90) Pterygoid, teeth arranged along posterior margin of ventral surface of pterygoid flange: present (0); absent (1).
    - Modified [14], character 121

91) Pterygoid, teeth on transverse flange: shagreen of very small teeth (0); single row of large teeth (1); single row of large teeth, with a shagreen of very small teeth anterior to it (2); Durophagous plate (3)
    - Modified from [22], character 71
    - Character state 3 new

92) Quadrate flange of pterygoid dentition: absent (0); present (1).
    - [22], character 72

93) Dentition on boss of pterygoid, if present: extensive (0); reduced (1).
    - Modified from [7], character 13

94) Ectopterygoid field of teeth: present (0); absent (1)
    - Multiple matrices
    - Note: ectopterygoid dental field is separate from pterygoid tooth field, which sometimes extend onto ectopterygoid

95) Ectopterygoid teeth organised into row of three or more of teeth: present (0); absent (1).
    - Modified from [2], character 284

96) Presence (0) or absence (1) of fangs on ectopterygoid
    - [20], character 66

97) Denticle field on posterior coronoid absent (0) or present (1).
    - [20], character 121

98) Denticle field on middle coronoid absent (0) or present (1).
    - [20], character 120

99) Coronoid dentition, if present, forms organized tooth row: yes (0); no (1)
    - Modified from [2], character 328

100) Coronoids: size of teeth on anterior and middle coronoids relative to dentary tooth size: (0) about the same; (1) half the height or less.
    - [2], character 330

101) Separate field of denticles lining the exochoanal margin absent (0) or present (1).
    - Modified from [20], character 73

102) Prearticular dentition absent (0) present (1)
    - Modified from [20], character 106

103) Premaxillary tooth size: subequal to non-caniniform maxillary teeth (0); larger than non-caniniform maxillary teeth (1); Smaller than non-caniniform maxillary teeth (2)
    - Modified from [11], character 64
    - Character state 2 new

104) Posteriormost dentary teeth occlude with the pterygoid teeth: absent (0); present (1)
    - New
    - If either are edentulous, score inapplicable

105) Upper caniniform implantation, if present: almost vertical (0); emergent portion of tooth/teeth points anteriorly (1)
    - New

106) Ventral bulging of tooth-bearing portion of maxilla ventrally offsets the maxillary tooth-row from the posteriormost pre-maxillary tooth: absent (0); present (1)
    - Modified from [14]

- Rewording to distinguish the “precanine step” visible in, for example, sphenacodontids (ventral bulging of maxilla) from taxa where the premaxillary toothrow is offset by an upturning of the premaxilla, which sometimes also produces a step in the lateral margin of the upper jaw, but which causes procumbent premaxillary teeth and is therefore already incorporated into character 21. Note that the two are not mutually exclusive; the premaxillary tooth row may be dorsally offset from the maxillary tooth row both by ventral bulging of the maxilla and an upward turn of the premaxilla.

107) Premaxillary tooth row displaced ventromedially by downturned premaxilla: absent (0); present (1)
- Modified from [5]
- Reworded to distinguish between the morphology observed in some archosaurs and captorhinids (where a downturned premaxilla changes the orientation of the teeth) and the toothless beak of some anomodonts (where a sharp ventral projection at the tip of the premaxilla forms the beak that itself represents the food processing equipment)

108) Diastema separating posteriormost maxillary tooth from those anterior to it: absent (0); present (1)
- New

109) Heel/ridge on non-caniniform maxillary and posterior dentary teeth: absent (0); present lingually, weak (1), present lingually, strong (2); present distally on upper jaw, medially on lower jaw (3); Ridge, narrowing towards the tip (4)
- New
- Characters referring to heeled morphology of the entire tooth-row have been included in matrices referring to parareptiles (to account for the morphology in bolosaurids) but have not been included here due to the need to account for the pattern observed in some therapsids, where precanine and canine can have combinations heels, but no postcanine teeth with heels are observed. Therefore, an extra character is added to incorporate presence or absence of heels on maxillary teeth.
- Character state 3 new represents the unique morphology observed in *Gansurhinus*

110) Lower premolar morphology: Conical (0); sectorial, with anterior and posterior accessory cusps (1); Bulbous (2); Spatulate (3); Molariform (4)
- New

111) Labial parapet on lower jaw: absent (0); low (not projecting higher than the bases of the cheek teeth (1); tall (as tall as the occlusal surface of the teeth) (2)
- [9]; character 22

112) Anteriormost tooth-bearing portion of the maxilla turns dorsally, so that the anterior maxillary teeth are directed anteriorly: absent (0); present (1)

- New

113) If maxillary/premaxillary tooth-row offset present due to ventral bulging of maxilla, caniniform (if present) is posterior to step (0); anterior to step (1)

- New

114) Row of enlarged teeth on lateral margin of palatal process of pterygoid: absent (0); present (1)

- New

Supplementary Dataset 2

Source Trees for supertree

Backbone: [2]

Seymouriamorpha: [20]

Diadectomorpha: [23]

Stem Eureptiles: [21,24-26]

Archosauromorpha: [1,27,28]

Lepidosauromorpha: [29]

‘Euryapsids’: [6,13,30,31]

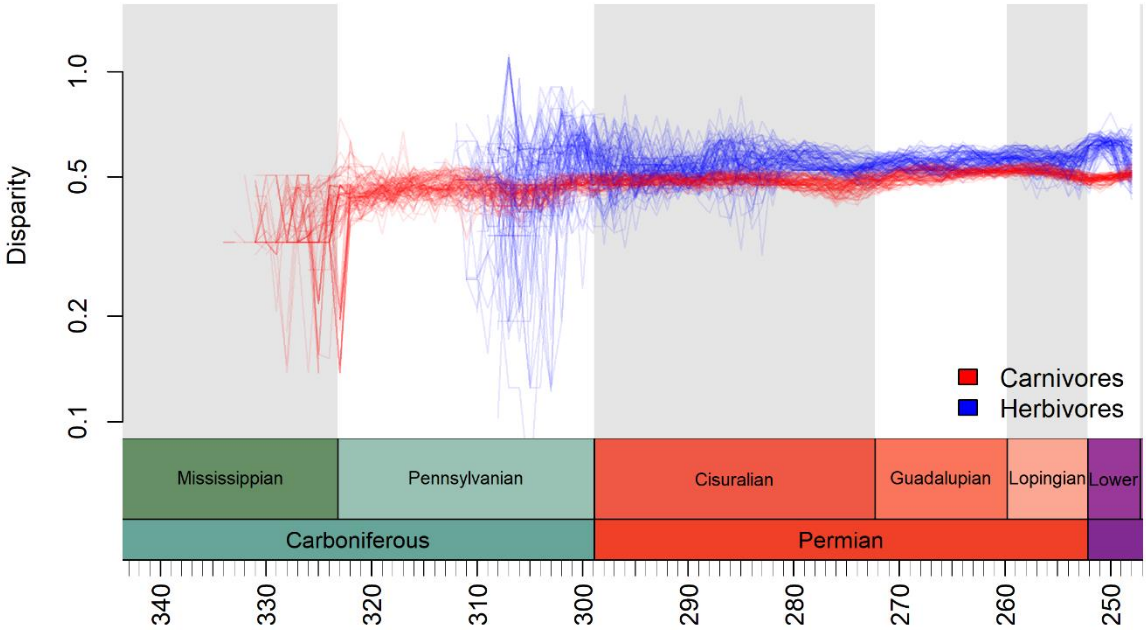
Parareptilia: [32-38]

‘Pelycosaurs’: [14,39-43]

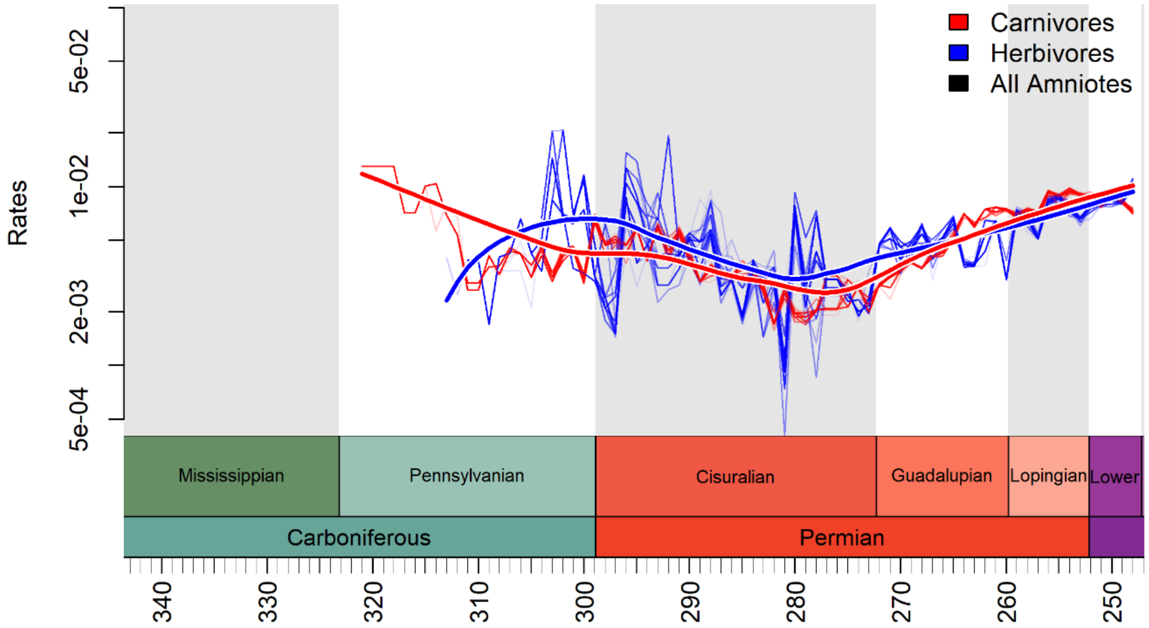
Therapsida: [4,7,8,11,12,19,44,45]

Microsauria: [46,47]

Supplementary Figures



Supplementary Figure 1: Dental disparity of faunivorous (red) and herbivorous (blue) amniotes, obtained analysing the supertree based on the backbone relationships of Ford & Benson [48]. Each line represents Each thin line represents results from one stochastic map used to assign a diet to each branch in one selected tree from the Bayesian posterior distribution.



Supplementary Figure 2: Rates of dental evolution of faunivore (red) and herbivorous (blue) amniotes, obtained analysing the supertree based on the backbone relationships of Ford & Benson [48]. Thin lines represent rates inferred from the maximum clade credibility tree. Each thin line represents results from one stochastic map used to assign a diet to each branch. The thick lines represent Loess fitted regression curves.

References

1. Ezcurra, MD. The phylogenetic relationships of basal archosauromorphs, with an emphasis on the systematics of proterosuchian archosauriforms. *PeerJ*, 4: e1778 (2016)
2. Pardo, JD, Szostakiwskyj, M, Ahlberg, PE & Anderson, JS. Hidden morphological diversity among early tetrapods. *Nature*, 546: 642-645 (2017)
3. Laurin, M. Anatomy and relationships of *Haptodus garnettensis*, a Pennsylvanian synapsid from Kansas. *J Vert Paleontol*, 13: 200-229 (1993)
4. Angielczyk, KD & Kammerer CF. The cranial morphology, phylogenetic position and biogeography of the upper Permian dicynodont *Compsodon helmoedi* van Hoepen (Therapsida, Anomodontia). *Pap Palaeontol*, 3: 513-545 (2017).
5. Modesto, SP, Lamb, AJ & Reisz, RR. The captorhinid reptile *Captorhinikos valensis* from the lower Permian Vale Formation of Texas, and the evolution of herbivory in eureptiles. *J Vert Paleontol*, 34: 291-302 (2014).
6. Moon, BC. A new phylogeny of ichthyosaurs (Reptilia: Diapsida). *J Syst Paleontol*, 17:129-155 (2019)
7. Kammerer, CF & Masyutin V. Gorgonopsian therapsids (*Nochnitsa* gen. nov. and *Viatkogorgon*) from the Permian Kotelnich locality of Russia. *PeerJ*, 6: e4954 (2018).
8. Kammerer, CF & Masyutin V. A new therocephalian (*Gorynychus masyutinae* gen. et sp. nov.) from the Permian Kotelnich locality, Kirov Region, Russia. *PeerJ*, 6: e4933 (2018).
9. Kissel, RA & Reisz RR. *Ambedus pusillus*, new genus, new species, a small diadectid (Tetrapoda: Diadectomorpha) from the Lower Permian of Ohio, with a consideration of diadectomorph phylogeny. *Annal Carnegie Mus*, 73, 197-212 (2004).

10. Martinelli, AG, Kammerer, CF, Melo, TP, Paes Neto, VD, Ribeiro, AM, Da-Rosa, ÁA, Schultz, CL & Soares, MB. The African cynodont *Aleodon* (Cynodontia, Probainognathia) in the Triassic of southern Brazil and its biostratigraphic significance. *PLoS One*, 12: 0177948.
11. Amson, E & Laurin, M. On the affinities of *Tetraceratops insignis*, an Early Permian synapsid. *Acta Palaeontol Polon*, 56: 301-312 (2011).
12. Liu, J. Osteology, ontogeny, and phylogenetic position of Sinophoneus yumenensis (Therapsida, Dinocephalia) from the Middle Permian Dashankou Fauna of China. *J Vert Paleontol*, 33: 1394-1407 (2013).
13. Neenan, JM, Li, C, Rieppel, O, & Scheyer, TM. The cranial anatomy of Chinese placodonts and the phylogeny of Placodontia (Diapsida: Sauropterygia). *Zool J Linn Soc*, 175: 415-428 (2015).
14. Brocklehurst, N & Fröbisch, J. A reexamination of Milosaurus mccordi, and the evolution of large body size in Carboniferous synapsids. *J Vert Paleontol*, 38: e1508026 (2018).
15. Broom, R. Notices of some new genera and species of Karroo fossil reptiles. *Rec Albany Mus*, 41: 161-166 (1931).
16. Broom, R. *The mammal-like reptiles of South Africa and the origin of mammals*. H. F. and G. Witherby, London (1932).
17. Boonstra, LD. A contribution to the morphology of the mammal-like reptiles of the suborder Therocephalia. *Ann S Afr Mus*, 31: 215-267 (1934).
18. Lehman J-P. Therocephalia. In *Traité de Paléontologie, tome 6, partie 1* (Piveteau J, Ed.). Masson et Cie., Paris, 224-245.
19. Rubidge B.S. & Van Den Heever J.A. Morphology and systematic position of the dinocephalian Styracocephalus platyrhynchus. *Lethaia*, 30: 157-168 (1997).

20. Klembara J. The cranial anatomy, ontogeny, and relationships of *Karpinskiosaurus secundus* (Amalitzky) (Seymouriamorpha, Karpinskiosauridae) from the Upper Permian of European Russia). *Zool J Linn Soc*, 161: 184-212 (2010)
21. Sobral G, Sues HD, Müller J. Anatomy of the enigmatic reptile *Elachistosuchus huenei* Janesch 1949 (Reptilia: Diapsida) from the Upper Triassic of Germany and its relevance for the origin of Sauria. *PloS One*, 10: e0135114 (2015)
22. MacDougall MJ, Scott D, Modesto SP, Williams SA and Reisz RR. (2017). New material of the reptile *Colobomycter pholeter* (Parareptilia: Lanthanosuchoidea) and the diversity of reptiles during the Early Permian (Cisuralian). *Zool J Linn Soc*, 180: 661-671 (2017).
23. Liu, J. Bever, G.S. The last diadectomorph sheds light on Late Palaeozoic tetrapod biogeography. *Biol Lett*. **11**, 20150100 (2015)
24. DeBraga, M., Reisz, R. R. A new diapsid reptile from the uppermost Carboniferous (Stephanian) of Kansas. Palaeontology 38, 199-212 (1995)
25. Müller, J., Reisz, R. R. The phylogeny of early eureptiles: comparing parsimony and Bayesian approaches in the investigation of a basal fossil clade. Syst. Biol. 55, 503-511 (2006)
26. Modesto SP, Scott D, Reisz RR. A new small captorhinid reptile from the lower Permian of Oklahoma and resource partitioning among small captorhinids in the Richards Spur Fauna. *Pap Paleontol*, 4, 293-307 (2018)
27. Pinheiro, F. L., Franca, M. A. G., Lacerda, M. B., Butler, R. J., Schultz, C. L. An exceptional fossil skull from South America and the origins of the archosauriform radiation. *Sci Rep* **6**, 22817 (2016).
28. Butler, RJ, Brusatte, SL. Reich, M., Nesbitt, SJ, Schoch RR, Hormung JJ. The sail-backed reptile *Ctenosauriscus* from the latest Early Triassic of Germany and the

timing and biogeography of the early archosaur radiation. *PLoS ONE*, 6: e25693 (2011)

29. Simões TR, Caldwell MW, Talanda M, Bernardi M, Palci A, Vernygora O, Bernardini F, Mancini L, Nydam RL. The origin of squamates revealed by a Middle Triassic lizard from the Italian Alps. *Nature*, 557: 707-709 (2018)
30. Chen, X-H, Motani R, Cheng L, Jiang D-Y, Rieppel O. A new specimen of Carroll's mystery hupehsuchian from the Lower Triassic of China. *PLoS ONE*, 10, e0126024
31. Liu, J., Zhao, L-J., Li, C., He, T. Osteology of *Concavispina biseridens* (Reptilia, Thalattosauria) from the Xiaowa Formation (Carnian), Guanling, Guizhou, China. *J Paleontol*, 87, 341-350
32. Liu, J. Bever GS. 2018. The tetrapod fauna of the Upper Permian Naobaogou Formation of China: a new species of *Elginia* (Parareptilia, Pareiasauria). *Pap Paleontol*, 4, 197-209
33. Cisneros, J.C., Rubidge, B.S., Mason, R. and Dube, C. Analysis of millerettid parareptile relationships in the light of new material of *Broomia perplexa* Watson, 1914, from the Permian of South Africa. *J. Syst. Paleontol* **6**, 453-462 (2008)
34. Falconnet, J. First evidence of a bolosaurid parareptile in France (latest Carboniferous-earliest Permian of the Autun basin) and the spatiotemporal distribution of the Bolosauridae. *Bull. Soc. Géol. Fr.* **183**, 495-508 (2012).
35. MacDougall, M. J., Modesto, S. P. and Botha-Brink, J. The postcranial skeleton of the Early Triassic parareptile *Sauropareion anoplus*, with a discussion of possible life history. *Acta Palaeontol. Pol.* **58**, 737-749 (2013).
36. MacDougall, M.J., Modetso, S.P. and Reisz, R.R. A new reptile from the Richard's Spur locality, Oklahoma, U.S.A., and patterns of Early Permian parareptile diversification. *J. Vert Paleontol.* **36**, e1179641 (2016)

37. Modesto, S. P., Scott, D. M. and Reisz, R. R. A new parareptile with temporal fenestration from the Middle Permian of South Africa. *Can. J. Earth Sci.* **46**, 9-20 (2009).
38. Cisneros JC, Damiani R, Schultz C, da Roda A, Schwanke C, Neto LW, Aurélio PLP. A procolophonoid reptile with temporal fenestration from the Middle Triassic of Brazil. *Proc R Soc B,* 271, 1541-1546 (2004)
39. Berman, D. S., Henrici, A. C., Sumida, S. S., Martens, T., Pelletier, V. First European record of a varanodontine (Synapsida: Varanopidae): member of a unique Early Permian upland paleoecosystem, Tambach Basin, Central Germany. In Kammerer, C. F., Angielczyk, K. D. and Frobisch, J. (ed.) *Early Evolutionary History of the Synapsida*. Springer Science, Dordrecht, p69-86 (2014).
40. Brocklehurst, N., Brink. K.S. Selection towards larger body size in both herbivorous and carnivorous synapsids during the Carboniferous. *Facets* **2**, 68-84 (2017)
41. Mazierski, D. M., Reisz, R. R. 2010. Description of a new specimen of *Ianthasaurus hardestiorum* (Eupelycosauria: Edaphosauridae) and a re-evaluation of edaphosaurid phylogeny. *Can. J. Earth. Sci.* **47**, 901-912
42. Romano, M., Ronchi, A., Maganuco, S., Nicosia, U. New Material of *Alierasaurus ronchii* (Synapsida, Caseidae) from the Permian of Sardinia (Italy), and its phylogenetic affinities. *Paleontol Electr* 20.2.26A: 1-27 (2018)
43. Spindler, F., Wenberg, R., Schneider, J.W., Luthardt L., Annacker, V., Rößler R. First arboreal ‘pelycosaurs’ (Synapsida: Varanopidae) from the early Permian Chemnitz fossil Lagerstätte, SE Germany, with a review of varanopid phylogeny. *PalZ*, 92,315-364

44. Fröbisch, J., Reisz, R. R. 2011. The postcranial anatomy of *Suminia getmanovi* (Synapsida: Anomodontia), the earliest known arboreal tetrapod. *Zool. J. Linn. Soc.* **162**, 661-698 (2011).
45. Day, M. O., Smith, R. M. H., Fernandez, V. A new species of burnetiid (Therapsida, Burnetiamorpha) from the early Wuchiapingian of South Africa and implications for the evolutionary ecology of the family Burnetiidae. *Pap Paleontol,* 4, 453-475 (2018)
46. Glienke, S. A new "microsaur (Ambibia; Lepospondyli) from the Rotliegens of the Saar-Palatinate region (Carboniferous/Permian transition; West Germany). *Paläontol. Z.* **86**, 297-311 (2012)
47. Huttenlocker, A.K., Pardo, J.D., Small, B.J., Anderson, J.S. Cranial morphology of recumbirostrans (Lepospondyli) from the Permian of Kansas and Nebraska, and early morphological evolution inferred by
48. Ford, D. P. & Benson, R. B. J. The phylogeny of early amniotes and the affinities of Parareptilia and Varanopidae. *Nature Eco Evo* **4**, 57-65 (2020)