Title: Fossils reveal the complex evolutionary history of the mammalian regionalized spine

Authors: K. E. Jones[1]*, K. D. Angielczyk[2], P. D. Polly[3], J. J. Head[4], V. Fernandez[5], J.K. Lungmus[6], S. Tulga[7], S. E. Pierce[1]*

Affiliations:

[1]Museum of Comparative Zoology and Department of Organismic and Evolutionary Biology, Harvard University, 26 Oxford Street, Cambridge MA 02138, USA

[2] Integrative Research Center, Field Museum of Natural History, 1400 South Lake Shore Drive Chicago, IL 60605-2496, USA

[3]Department of Earth and Atmospheric Sciences, Indiana University Bloomington, 1001 East 10th Street Bloomington, IN 47405-1405, USA

[4]Department of Zoology, University of Cambridge, Downing Street, Cambridge, CB2 3EJ, UK

[5]European Synchrotron Radiation Facility, 71 Rue des Martyrs, 38000 Grenoble, France.

[6]Department of Organismal Biology and Anatomy, University of Chicago, 1027 E 57th Street, Chicago, IL 60637, USA

[7]Department of the Geophysical Sciences, University of Chicago, 5734 S. Ellis Avenue Chicago, Illinois 60637

*Correspondence to: katrinajones@fas.harvard.edu, spierce@oeb.harvard.edu

Abstract: A unique characteristic of mammals is a vertebral column with anatomically distinct regions, but when and how this trait evolved remains unknown. Here we reconstruct vertebral regions and their morphological disparity in the extinct forerunners of mammals, the non-mammalian synapsids, to elucidate the evolution of mammalian axial differentiation. Mapping patterns of regionalization and disparity (heterogeneity) across amniotes reveals that both traits increased during synapsid evolution. However, the onset of regionalization predates increased heterogeneity. Based on inferred homology patterns, we propose a "pectoral-first" hypothesis for region acquisition. Evolutionary shifts in forelimb function in non-mammalian therapsids drove increasing vertebral modularity prior to differentiation of the vertebral column for specialized functions in mammals.

One Sentence Summary: Evolution of vertebral regions in mammal forerunners was triggered by changes in forelimb function.

Main text: The evolution of the mammalian body plan from the ancestral amniote condition is one of the most iconic macroevolutionary transitions in the vertebrate fossil record (*1, 2*). A unique feature of mammals is their specialized vertebral column, which displays constrained vertebral counts, but highly disparate morphologies (*2-4*). In therian mammals, the presacral vertebral column is traditionally divided into cervical, rib-bearing thoracic, and ribless lumbar regions (Figure 1A). In contrast, the presacral vertebrae of basal amniotes are comparatively uniform and show little differentiation (Figure 1B). The transition from an 'unregionalized' to 'regionalized' presacral column is an important step in mammalian evolution, and has been linked to the origin of specialized gaits and respiratory function (*1, 2, 5, 6*).

Recent quantitative work has detected subtle presacral regionalization in extant snakes and limbed lizards, superficially unregionalized taxa (*7*). It was hypothesized that the ancestral amniote

condition is ‘cryptic regionalization’, in which regions are present but only subtly expressed. The global-patterning *Homeobox (Hox)* genes were implicated as underlying these conserved regionalization patterns. Under this model, the degree of *regionalization* – the number of regions present – has remained constant through mammal evolution, whereas the amount of morphological disparity between regions (here termed *heterogeneity*) has increased. But, this evolutionary scenario is based solely on extant data.

The two amniote clades – Synapsida (mammals and their relatives) and Sauropsida (reptiles/birds and their relatives) – diverged over 320 Ma and have independently undergone significant morphological transformations. Therefore, to document the evolution of the mammalian vertebral column, we must examine their extinct forerunners, the non-mammalian synapsids. Here we examined the presacral vertebral columns of 16 exceptionally-preserved non-mammalian synapsids (ranging from ‘pelycosaurs’ to cynodonts), one extinct amniote outgroup, and a broad range of extant salamanders, reptiles, and mammals. Using morphometric data, we quantified patterns of regionalization and heterogeneity, and compared their evolution to elucidate when and how synapsid presacral differentiation occurred.

Using a likelihood-based segmented regression approach, we calculated a regionalization score for each taxon (an AIC-weighted average of the relative fit of one-to-six region hypotheses) producing a continuous variable reflecting the estimated number of vertebral regions (Figure S2). Similar to prior work (*7*), most reptiles and some extant mammals (e.g., monotremes) have scores around four regions (Figure 2A), whereas therians (marsupials and placentals) most frequently display five regions. Therian regionalization scores are also more variable, probably reflecting high ecomorphological diversification of their axial system (*4*). Thus, data from extant amniotes alone support the null hypothesis of conserved regionalization. However, both salamanders and

basal synapsids have lower regionalization scores than extant amniotes (Figure 2A, cool colors), demonstrating that regionalization increased independently in the sauropsid and synapsid lineages. Accordingly, we reject the hypothesis of conserved regionalization patterns in amniotes, and instead propose increasing regionalization in synapsid evolution.

Heterogeneity, the log mean variance of the morphological measures for each column, also increases during synapsid evolution (Figure 2). Lepidosaurs and salamanders have low heterogeneity, denoting relative uniformity of the axial column (*7*); therians have much higher values reflecting their extreme disparity; and crocodilians have intermediate levels. Most non-mammalian synapsids also have intermediate levels of heterogeneity. The outgroup *Diadectes* and the ophiacodontids display particularly low values, reinforcing previous assertions of homoplastic increases in mammals and archosaurs from a homogeneous ancestral condition (*7*). The cynodont *Kayentatherium* has more heterogenous morphologies than the other fossil taxa, reflecting its position close to the mammal radiation. Given the association between heterogeneity and functional specialization of the axial skeleton in therians, the more homogenous morphologies of most non-mammalian synapsids points toward functional conservatism.

Although regionalization and heterogeneity increase during synapsid evolution, they are not significantly correlated (Figure S7, Table S6, p=0.73), meaning that simple linear change is insufficient to explain these patterns. Instead, quantitative trait modeling supports evolution toward shifting adaptive optima (MultiOU models) for these data (Table S7). Based on AIC fitting, we reconstruct two major adaptive shifts in each trait during synapsid evolution (Figure 3, S8). The adaptive optimum for regionalization increases from around three regions in ‘pelycosaurs’ to around four regions at the base of Therapsida, with a later shift to five regions occurring in Theria. The adaptive optimum for heterogeneity increases first at Cynodontia, and subsequently within

therians. Taken together, our data reveal that vertebral regionalization increased prior to increasing heterogeneity, demonstrating that these two measures of axial differentiation evolved independently.

To understand how vertebral regionalization increased in synapsids, we reconstructed region boundaries recovered in the best-fit segmented regression models (Figure 4A). Region boundaries were then cross-referenced with developmental data, anatomical landmarks, and variation in extant species to identify homologies (Figure 4B). In extant tetrapods, the cervicothoracic transition is correlated with *Hox6* expression, rib morphology, and the position of the forelimb and brachial plexus (*8*). Therefore, the cervicothoracic boundary was identified by the position of the posterior branch of the brachial plexus, and the anterior sternal articulation or first long rib. Functional studies in *Mus* also show that *Hox9* patterns the transition from sternal- to non-articulating ribs, and *Hox10* controls the suppression of ribs altogether in the lumbar region (Figure 4B) (*9, 10*). In keeping with this association, dorsal regions were defined relative to their proximity to long ribs (anterior dorsal), short ribs (posterior dorsal), or absent ribs (lumbar).

Using these criteria, region homology hypotheses were constructed in key taxa for which rib or neural anatomy are known (Figure 4B). In salamanders (and the stem amniote *Diadectes*, see Supplementary text), three regions are recovered. The anterior break correlates with the posterior branch of the brachial plexus in *Ambystoma*, implying homology with the cervical region despite the lack of a true 'neck' (Figure 4B, red region). Although salamanders have poorly-developed ribs, the position of the posterior break in the mid-trunk is consistent with the anterior-posterior dorsal transition in other taxa (Figure 4A/B, yellow-pale blue). This ancestral three-region pattern is retained in the most basal synapsids. In 'pelycosaurs', the first break corresponds with the inferred cervicothoracic transition based on rib length and forelimb position (e.g., v5 in

Edaphosaurus, v7 in *Dimetrodon*), whereas the second corresponds with the gradual transition from longer to shorter dorsal ribs, signifying cervical, anterior dorsal, and posterior dorsal homologies (*Edaphosaurus* Figure 4B).

Our data point to the convergent addition of a fourth region in distinct locations in sauropsids and synapsids. In sauropsids, a fourth region is detected anterior to the brachial plexus, suggesting a novel cranial region within the neck (*Iguana* Figure 4B, purple region). Sauropsids exhibit more variation in cervical count than synapsids (*11*), providing a potential connection between neck plasticity and cervical modularity in this lineage. Conversely, in basal therapsids and cynodonts a fourth region is detected posterior to the cervicothoracic transition (*Thrinaxodon* Figure 4B, orange region). In *Thrinaxodon*, the first break corresponds with the cervicothoracic transition and first full-length rib (v7-8); the second break lies in the middle of the long rib series (v12-13); whereas the anterior-posterior dorsal boundary falls at the transition from long to short ribs (v19-20). These regions conform to the ancestral cervical region (red), a novel pectoral region (orange); and the ancestral anterior dorsal (yellow) and posterior dorsal (pale blue) regions. Therian mammals display an additional break within the posterior dorsal region that differentiates the ribless lumbar region (*Mus* Figure 4B, blue region).

Considering the pattern of region acquisition, we propose a "pectoral-first" hypothesis for the evolution of mammalian presacral regionalization (Figure 4). Under this hypothesis, 'pelycosaurs' retained the three-region ancestral amniote condition. Addition of a fourth 'pectoral module' occurred in basal therapsids accompanying the reorganization of the pectoral girdle and forelimb. Unlike 'pelycosaurs', therapsids are characterized by reduction of the pectoral girdle dermal bones and increased shoulder mobility (*1, 12*). Medial extrinsic shoulder muscles (*e.g., levator scapulae, serratus ventralis*) originating on the scapula are thought to have expanded their

axial insertions during synapsid evolution (*12*). As these vital body-support muscles attach directly onto the underlying vertebrae and ribs, shifts in pectoral morphology and function likely drove divergent neck-shoulder selective regimes in the axial skeleton, providing impetus for increased regionalization (*1, 12, 13*). Further, the vertebrae, medial extrinsic shoulder muscles, and dorsal border of the scapula all develop directly from somitic mesoderm (primaxial), signifying strong developmental ties between these structures (*14*).

Interestingly, it has been proposed that the muscular diaphragm evolved from an unmuscularized septum or 'proto-diaphragm' via cooption of shoulder muscle precursor cells – later canalized into a distinct cell population by repatterning of the posterior neck (*15*). Reorganization of the anterior column and pectoral girdle in therapsids may have facilitated this transition by increasing cervicothoracic modularity and remodeling shoulder musculature. Subsequent fixation of the cervical count at seven in non-mammalian cynodonts is hypothesized to represent the appearance of the mammalian-style muscular diaphragm (*6*). Thus, anterior regionalization initially associated with shoulder evolution in early therapsids was likely later exapted in cynodonts in response to selection for increased ventilatory efficiency (*5, 15*).

A 'lumbar module' evolved later in therian mammals. Evolution of the lumbar region in mammals is associated with *Hox10*, which functions to repress rib formation and patterns lumbar identity in *Mus* (*10*)(Figure 4B). Convergent loss/gain of lumbar ribs in multiple fossil theriiform clades suggests high plasticity of this character early in therian evolution (*16*). Within therians, lumbar count and morphology vary, and this is reflected by translocation of the (morphometrically-defined) region boundary in our sample. As the lumbar region plays a critical role in mammalian locomotion, it is predicted that region variability is related to ecological specialization caused by clade-specific functional overprinting.

Regional differentiation is "the major structural difference between reptilian and mammalian vertebral columns" (*13*), yet its evolution has never been quantitatively examined. Here we demonstrate that regionalization and heterogeneity – the two aspects of vertebral differentiation – evolved independently. Forelimb reorganization in therapsids drove initial increases in regionalization due to developmental and functional connections between the pectoral girdle and underlying vertebrae. High heterogeneity and presumed functional diversity did not appear until crown mammals. The combination of a regionalized axial skeleton with heterogenous vertebral morphologies ultimately enabled mammals to become specialized for a remarkable diversity of ecologies and behaviors.

Acknowledgments: We extend thanks to: R. Asher, B. Brainerd, D. Brinkman, T. Capellini, C. Capobianco, J. Chupasko, J. Cundiff, K. Jakata, T. Kemp, C. Mehling, A. Millhouse, M. Omura, A. Resetar, J. Rosado, B. Rubidge, R. Smith, K. Smithson, C. Tabin, R. Tykoski, I. Werneburg, and B. Zipfel. **Funding:** NSF grants EAR-1524523 (to S.E.P.) and EAR-1524938 (to K.D.A.), AAA Postdoctoral Fellowship (to K.E.J.). **Author contributions:** Study design: K.E.J., K.D.A., S.E.P. Methods: K.E.J., P.D.P., S.E.P. Data collection: K.E.J., K.D.A., S.E.P., J.J.H, V.F., J.L., S.T. Data analysis: K.E.J. Manuscript preparation: K.E.J., K.D.A., S.E.P. **Competing interests:** The authors have no competing interests. **Data and materials availability:** Data are available in Table S5 and Dryad [link]. Code is available via github (https://github.com/katrinajones/regions).

Supplementary Materials:

Materials and Methods

Supplementary Text

Figures S1-S8

Tables S1-S7

References (19-50)

FIGURE CAPTIONS

Figure 1. Regionalization and heterogeneity. (A) The therian presacral column is highly regionalized and morphologically differentiated (*Mus musculus*), (B) Basal synapsids display a homogenous dorsal region with little differentiation (*Ophiacodon*).

Figure 2. Evolution of presacral differentiation in amniotes. (A) Regionalization score, (B) Log variance. Warmer colors reflect more regions and greater morphological heterogeneity, respectively. Black circles: mammals; grey circles: fossil taxa; triangles: reptiles; stars: amphibians. Greyed tips: fossil taxa excluded due to <0.75 r-squared of regionalization model. Full taxonomic names: Table S5.

Figure 3. Adaptive regime shifts in vertebral evolution. (A) Regionalization; (B) Heterogeneity. Theta: adaptive optima of each regime.

Figure 4. Best-fit region models, region homologies, and evolutionary hypothesis. (A) Best-fit region models for select taxa. Colors represent inferred region homologies. St. Dev.: Standard deviation of break locations; PS count: presacral count; R-sq: Adjusted r-squared; % complete: Total completeness. Grey boxes: Taxa with <0.75 r-squared fit were excluded from evolutionary reconstructions. (B) "Pectoral-first" hypothesis for the evolution of synapsid presacral regionalization. Taxa (left to right): *Ambystoma, Iguana, Edpahosaurus* (redrawn from (*17*)), *Thrinaxodon* (redrawn from (*2*)), *Mus*. Width of gray bars reflects relative rib-lengths and/or connection to sternum; vertical dashed lines denote cervicothoracic transition. For *Mus*, Hox bands correspond to vertebrae affected by functional gene manipulation (*18*).

REFERENCES CITED

1. T. S. Kemp, *The Origin and Evolution of Mammals*. (Oxford University Press, 2005).
2. F. A. Jenkins, Cynodont postcranial anatomy and prototherian level of mammalian organization. *Evolution* **24**, 230-252 (1970).
3. N. Schilling, Evolution of the axial system in craniates: Morphology and function of the perivertebral musculature. *Front Zool* **8**, 4-23 (2011).
4. E. J. Slijper, Comparative biologic-anatomical investigations on the vertebral column and spinal musculature of mammals. *Verh K Ned Akad Wet Afd Natuurkd Tweede Reeks* **42**, 1-128 (1946).
5. D. R. Carrier, The evolution of locomotor stamina in tetrapods: Circumventing a mechanical constraint. *Paleobiology* **13**, 326-341 (1987).
6. E. A. Buchholtz *et al.*, Fixed cervical count and the origin of the mammalian diaphragm. *Evol Dev* **14**, 399-411 (2012).
7. J. J. Head, P. D. Polly, Evolution of the snake body form reveals homoplasy in amniote Hox gene function. *Nature* **520**, 86-89 (2015).
8. A. C. Burke, C. E. Nelson, B. A. Morgan, C. Tabin, Hox genes and the evolution of vertebrate axial morphology. *Development* **121**, 333-346 (1995).
9. D. C. McIntyre *et al.*, *Hox* patterning of the vertebrate rib cage. *Development* **134**, 2981-2989 (2007).
10. D. M. Wellik, M. R. Capecchi, *Hox10* and *Hox11* genes are required to globally pattern the mammalian skeleton. *Science* **301**, 363-367 (2003).

11. J. Müller *et al.*, Homeotic effects, somitogenesis and the evolution of vertebral numbers in recent and fossil amniotes. *Proceedings of the National Academy of Sciences* **107**, 2118-2123 (2010).

12. A. S. Romer, The locomotor apparatus of certain primitive and mammal-like reptiles. Bulletin of the AMNH; v. 46, article 10. (1922).

13. A. Crompton, F. A. Jenkins Jr, Mammals from reptiles: a review of mammalian origins. *Annu Rev Earth Planet Sci* **1**, 131-155 (1973).

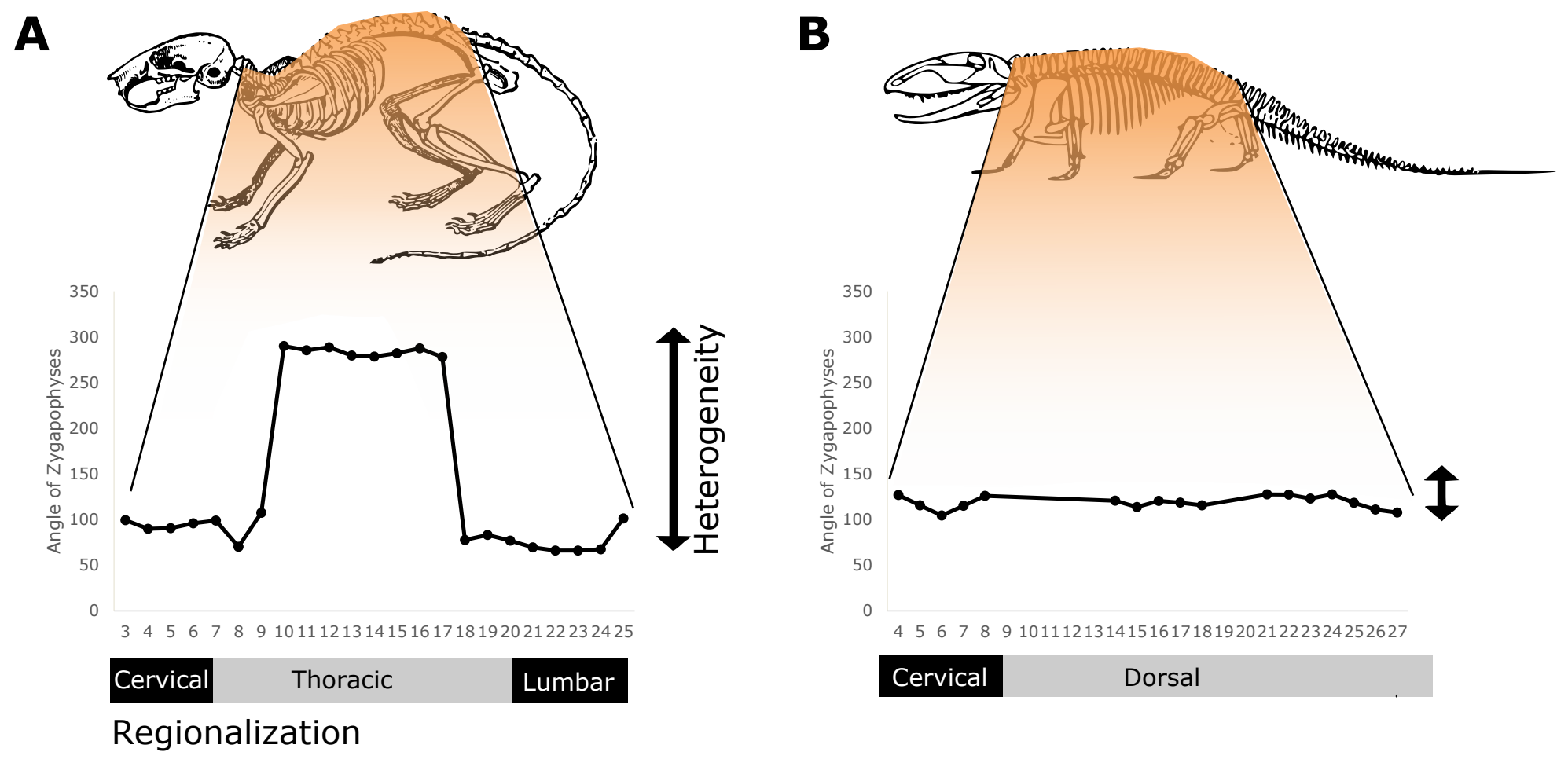
14. P. Valasek *et al.*, Somitic origin of the medial border of the mammalian scapula and its homology to the avian scapula blade. *J Anat* **216**, 482-488 (2010).

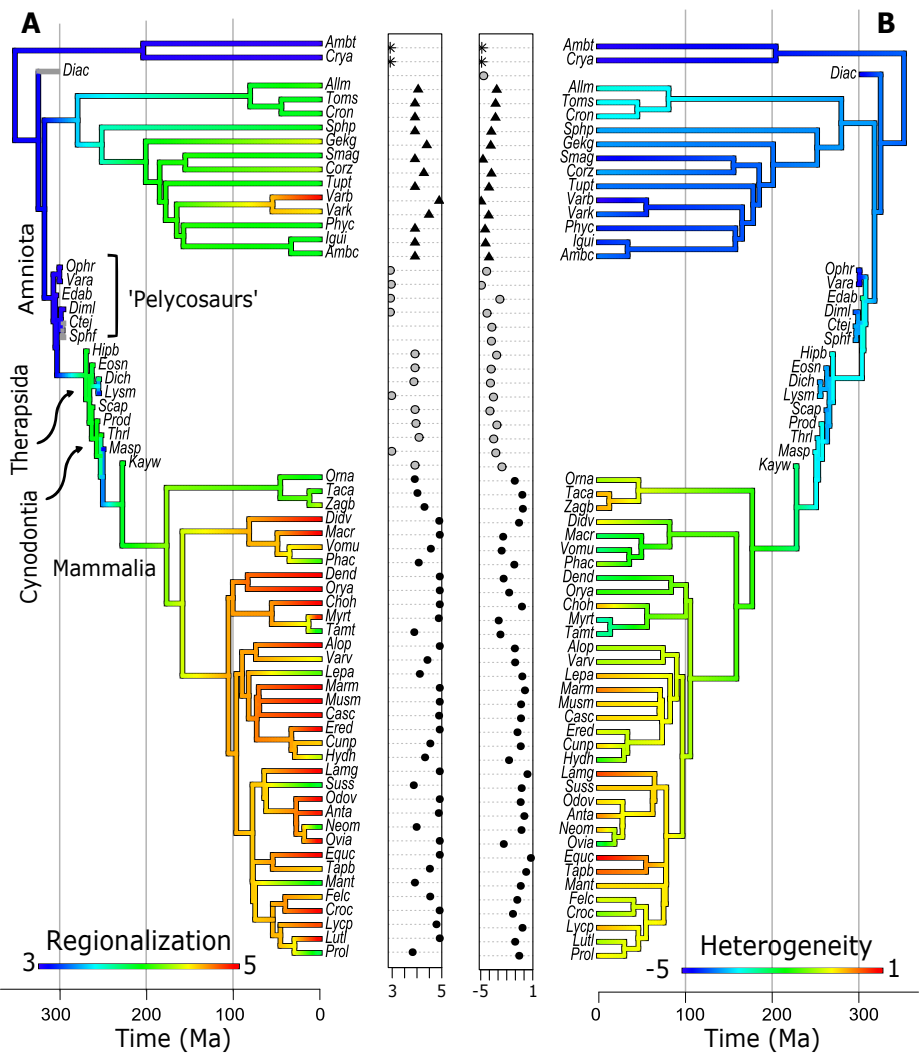
15. T. Hirasawa, S. Kuratani, A new scenario of the evolutionary derivation of the mammalian diaphragm from shoulder muscles. *J Anat* **222**, 504-517 (2013).

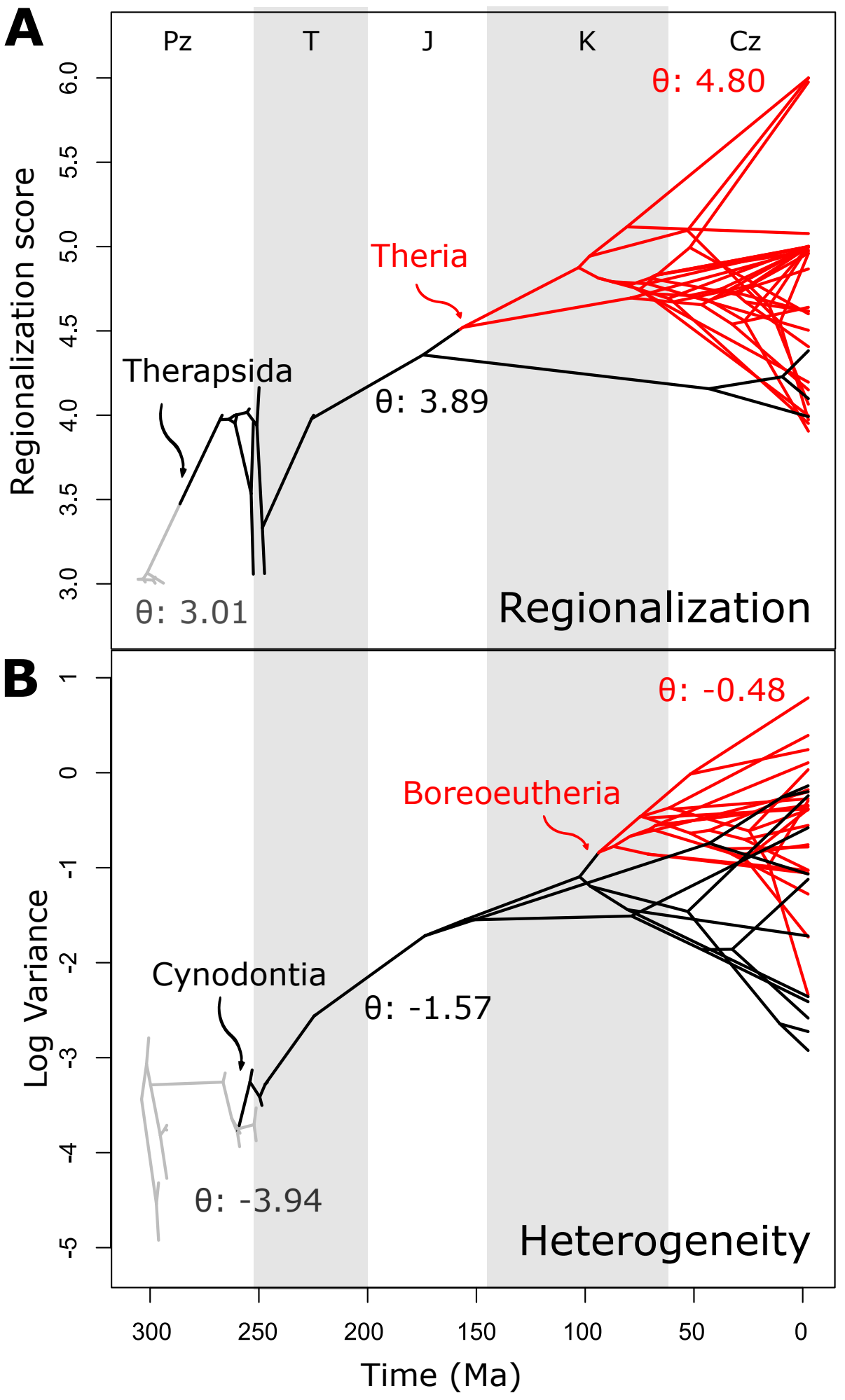
16. Z. X. Luo, P. J. Chen, G. Li, M. Chen, A new eutriconodont mammal and evolutionary development in early mammals. *Nature* **446**, 288-293 (2007).

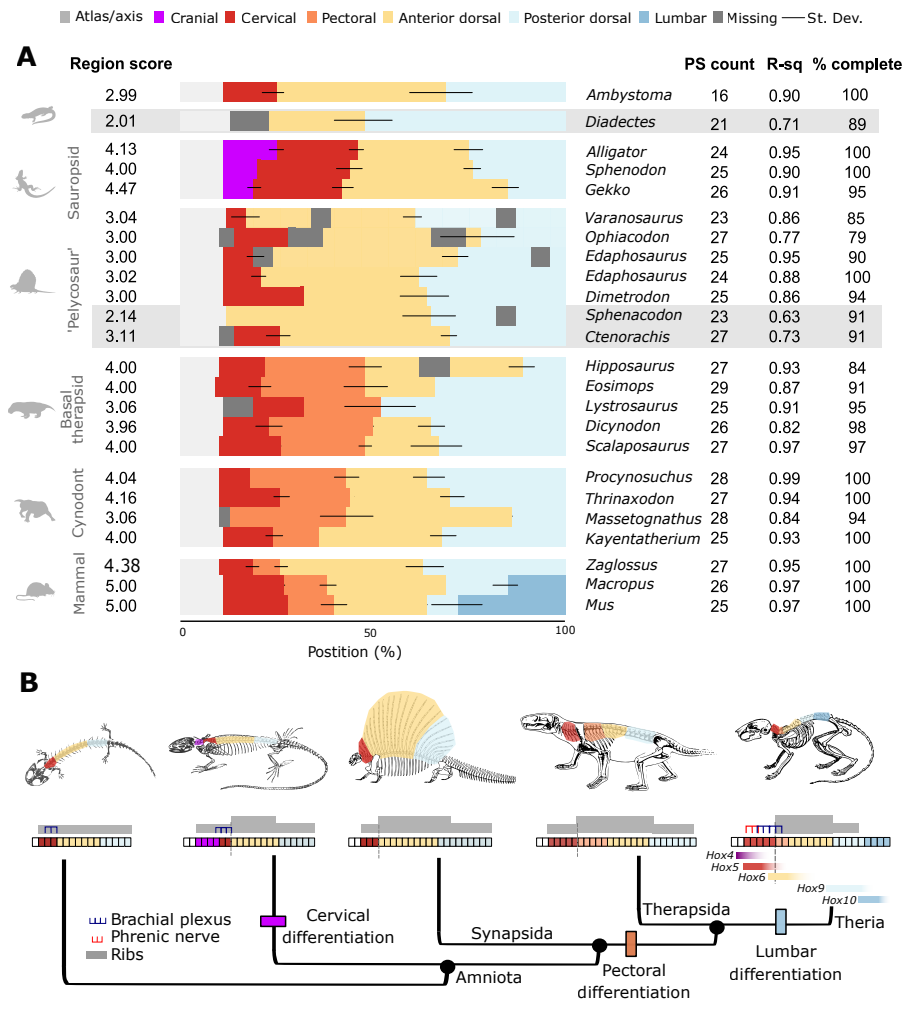
17. H.-D. Sues, R. R. Reisz, Origins and early evolution of herbivory in tetrapods. *Trends Ecol. Evol.* **13**, 141-145 (1998).

18. D. M. Wellik, *Hox* patterning of the vertebrate axial skeleton. *Dev Dyn* **236**, 2454-2463 (2007).


A
B
Angle of Zygapophyses
0
50
100
150
200
250
300
350
3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25
4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27
Heterogeneity
Cervical
Thoracic
Lumbar
Cervical
Dorsal
Regionalization


A
B
Ambt
Crya
Diac
Allm
Toms
Cron
Sphp
Gekg
Smag
Corz
Tupt
Varb
Vark
Phyc
Igui
Ambc
Amniota
Ophr
Vara
Edab
Diml
Ctej
Sphf
'Pelycosaurs'
Therapsida
Hipb
Eosn
Dich
Lysm
Scap
Prod
Thrl
Masp
Kayw
Cynodontia
Mammalia
Orna
Taca
Zagb
Didv
Macr
Vomu
Phac
Dend
Orya
Choh
Myrt
Tamt
Alop
Varv
Lepa
Marm
Musm
Casc
Ered
Cunp
Hydh
Lamg
Suss
Odov
Anta
Neom
Ovia
Equc
Tapb
Mant
Felc
Croc
Lycp
Lutl
Prol
Regionalization
3
5
Heterogeneity
-5
1
300
200
100
0
Time (Ma)


A
Pz
T
J
K
Cz
θ: 4.80
Theria
Therapsida
θ: 3.89
θ: 3.01
Regionalization score
6.0
5.5
5.0
4.5
4.0
3.5
3.0
Regionalization
B
θ: -0.48
Boreoeutheria
Cynodontia
θ: -1.57
θ: -3.94
Log Variance
1
0
-1
-2
-3
-4
-5
Heterogeneity
300
250
200
150
100
50
0
Time (Ma)


Atlas/axis
Cranial
Cervical
Pectoral
Anterior dorsal
Posterior dorsal
Lumbar
Missing
St. Dev.
A
Region score
PS count
R-sq
% complete
2.99 Ambystoma 16 0.90 100
2.01 Diadectes 21 0.71 89
Sauropsid
4.13 Alligator 24 0.95 100
4.00 Sphenodon 25 0.90 100
4.47 Gekko 26 0.91 95
'Pelycosaur'
3.04 Varanosaurus 23 0.86 85
3.00 Ophiacodon 27 0.77 79
3.00 Edaphosaurus 25 0.95 90
3.02 Edaphosaurus 24 0.88 100
3.00 Dimetrodon 25 0.86 94
2.14 Sphenacodon 23 0.63 91
3.11 Ctenorachis 27 0.73 91
Basal therapsid
4.00 Hipposaurus 27 0.93 84
4.00 Eosimops 29 0.87 91
3.06 Lystrosaurus 25 0.91 95
3.96 Dicynodon 26 0.82 98
4.00 Scalaposaurus 27 0.97 97
Cynodont
4.04 Procynosuchus 28 0.99 100
4.16 Thrinaxodon 27 0.94 100
3.06 Massetognathus 28 0.84 94
4.00 Kayentatherium 25 0.93 100
Mammal
4.38 Zaglossus 27 0.95 100
5.00 Macropus 26 0.97 100
5.00 Mus 25 0.97 100
0
50
100
Postition (%)
B
Hox4
Hox5
Hox6
Hox9
Hox10
Brachial plexus
Phrenic nerve
Ribs
Cervical differentiation
Synapsida
Therapsida
Theria
Amniota
Pectoral differentiation
Lumbar differentiation

Supplementary Materials for

Title: Fossils reveal the complex evolutionary history of the mammalian regionalized spine

Authors: K. E. Jones[1]*, K. D. Angielczyk[2], P. D. Polly[3], J. J. Head[4], V. Fernandez[5], J. K. Lungmus[6], S. Tulga[7], S. E. Pierce[1*]

*Correspondence to: katrinajones@fas.harvard.edu, spierce@oeb.harvard.edu

This PDF file includes:

Materials and Methods

Sample

Extant sample - Vertebral morphology was examined in a wide range of extant mammals (n=46), salamanders (n=2), and non-avian reptiles (n=14, Table S1). All specimens were adults with no visible vertebral pathology, and complete or nearly complete vertebral columns. Only articulated lepidosaur vertebral columns were used when skeletonized because they are difficult to seriate when disarticulated. For larger animals, measurements were taken directly from the skeleton, whereas measurements from smaller animals were taken from Computed-Tomographic scans of specimens (Bruker Skyscan 1173, Harvard University; GE v|tome|x scanner, University of Chicago).

Fossil sample - Only the most pristinely preserved fossils were selected for analysis. We obtained CT scans for 32 exceptionally preserved specimens. Of these, eight were excluded due to poor CT scan quality or insufficient contrast, and a further nine were excluded due to insufficient completeness, damage, or distortion. The remaining 16 specimens are listed in Table S2. They represent a broad taxonomic sampling of non-mammalian synapsids including six 'pelycosaurs,' five basal therapsids, and four cynodonts. One stem amniote (*Diadectes tenuitectes*) was included as an outgroup.

Fossil specimens were CT scanned using a variety of devices and locations (Table S2). *Edaphosaurus boangeres* DMNH 2011-04-01 was digitized with structured-light surface scanner (0.5 mm resolution) because its elongate spines made CT scanning challenging. An additional specimen of *Edaphosaurus boangeres* (AMNH FARB 4015), a mounted exhibit specimen, was also included for comparison and was measured directly using calipers and tape. Most specimens were preserved in articulation, or partial articulation, allowing for accurate seriation of the vertebrae. Where disarticulated, specimens had associated vertebral numbers from their original preparation, had been mounted in sequence, or had original descriptions in the literature to which we referred. Generally, undeformed specimens were selected. However, *Eosimops newtoni* was preserved on a flat block which was compressed dorsoventrally. To correct for this distortion, the entire block was retrodeformed uniformly by applying a 1.42 scaling factor. Though specimens were generally selected with very limited damage, vertebrae that were broken or split into pieces were digitally reconstructed in 3Matic software (Materialise, Leuven, Belgium) prior to measurement.

Fossil provenance – All fossils included in this study are accessioned in museums (Table S2), and all but one have been described previously *(Eosimops newtoni: 19; Dicynodon huenei: 20; Thrinaxodon liorhinus: 21; Kayentatherium wellesi: 22; Procynosuchus delaharpaea:23; Ctenorachis jacksoni: 24; Dimetrodon limbatus: 25; Scalaposaurus punctatus: 26; Diadectes tenuitectes: 27; Ophiacodon retroversus: 28; Sphenacodon ferox: 29; Hipposaurus boonstrai: 30; Massetognathus pasculi: 31; Varanosaurus acutirostris: 32; Lystrosaurus murrayi: 33; Edaphosaurus boanerges* AMNH FARB 4015*: 34).*

DMNH 2011-04-01 (*Edaphosaurus boanerges*) consists of a nearly complete skeleton that was collected by Mr. David Williams from the Putnam Formation of Archer Country, Texas, in October, 1994 (detailed locality information available to qualified researchers on request from DMNH). The specimen was donated to DMNH in 2011, at which time the vertebral column (preserved in a nearly solid ironstone concretion) was prepared.

Measurements

The presacral column (excluding atlas and axis) for each specimen was measured using a traditional morphometric approach (Step 1, Figure S2). When taking measurements directly from the skeleton, linear measures were taken using digital calipers (Mitutoyo, 0.02mm precision) and angles were measured from photographs in ImageJ (*35*). When using 3D models from scans, linear measures were taken in Mimics software (Materialise, Leuven, Belgium), whereas angles were taken from orthogonal screenshots and measured in ImageJ (*35*).

Fifteen linear and four angular measures were taken on most specimens (Figure S1, Table S3). We follow a functionally-relevant definition of the transverse process, homologizing cervical transverse processes, laterally projecting diapophyses, and lumbar transverse processes. The inferior lamellae of the cervical transverse processes were not included in transverse process measurements, but were included in total width, where appropriate. No diapophyses were present in some lepidosaurs, so these measurements were excluded. Where a structure was absent for some portion of the column (but not missing), the structure was coded as zero for a linear measure or ninety degrees for an angular measure. This allows the lack of a structure to be incorporated as a morphological trait in the analysis, and reflects the serial homology of axial structures. Measures were scaled to unit variance and log transformed (zeros converted to 0.01 to enable logging) prior to analysis to extract patterns of variation, not magnitudes.

Missing data

Fossils are prone to missing data due to damage, distortion or incomplete preservation. Well-preserved, complete vertebral columns that demonstrably represent single individuals, and that are amenable to CT-scanning or detailed direct measurement, are exceedingly rare. Therefore, to maximize our dataset we included specimens with limited missing data (Table S4). We optimized our methodology to allow inclusion of fossil data by: 1) using linear and angular measurements (instead of e.g., geometric morphometric approaches), which are easier to collect on damaged specimens and are more forgiving of missing data; 2) interpolating some missing data; 3) using analysis techniques amenable to missing data (see below); and 4) applying regionalization analysis to each individual specimen separately, thereby allowing different combinations of measures based on preservation.

For short strings of missing data (two or less adjacent vertebrae missing the same measurement), the data were interpolated by taking the mean of their neighboring measurements. For longer strings, missing data were coded as NA, and thus excluded from subsequent analysis. Variables with large numbers of missing data were excluded entirely. Missing vertebrae were excluded, but were considered when numbering the vertebral positions for the independent variable of the regionalization analysis (described below). For a full review of the impact of missing data on our statistical analysis, see the Sensitivity Analysis section below.

Ordination

Data ordination was used to extract key aspects of variation from the dataset and to remove noise from the data (Step 2, Figure S2). A distance-based Principal Coordinates Analysis (PCO) was implemented in 'R' using a Gower distance matrix from the package *'cluster'* and custom R scripts (See package *'regions'*). When calculating mean dissimilarity between pairs of vertebrae, missing data were excluded by applying a weighting of zero.

Regionalization

To calculate the number of regions within a vertebral column, we used a segmented regression approach (*36*) modified for fossil data. Unlike traditional cluster-based methods, which neglect vertebral integration, our approach models the vertebral column as a series of morphological gradients and regions based on changes in slope or elevation. Details of the method are provided below and the workflow is summarized in Figure S2. The code used to perform the analysis has been developed into an R package called *'regions'*, which is freely-available on GitHub (link: to be uploaded to github at publication).

Model fitting - Each of the PCO's was fit successively to a series of segmented regression models (Step 3, Figure S2). An exhaustive search was conducted of all possible regressions containing two vertebrae or more, beginning with a single slope (one-region model) and progressing up to six separate slopes (six-region model). Regression lines were not constrained to be continuous with adjacent regression lines. To estimate the goodness-of-fit of the model to the data, the residual sums of squares was calculated for each line, and then summed for each PCO axis to give the total residual sums of squares (RSS) for each region model. For taxa with very short vertebral columns, the ratio of parameters to variables was too great to calculate AIC support for a six-region model (see below). Therefore, it was necessary to restrict the maximum number of regions of *Diadectes* and *Ambystoma* to five and four, respectively. Both taxa recovered values significantly lower than this in the analysis, suggesting they did not approach the upper threshold.

PCO selection – The number of PCOs used in the regionalization score calculation was selected to maximize its value (Step 4, Figure S2). The regionalization procedure described below was repeated using the cumulative residuals with a cutoff of PCO one through five iteratively. The final PCO cutoff was then selected *a posteriori* to yield the maximum possible regionalization score using the *PCOmax* function in the package 'r*egions'*. This method provides the most conservative approach for confirming the lower regionalization scores obtained for fossil synapsids as it removes potential downward bias based on PCO selection. Comparison of results with other PCO cut-off methods is provided in Figure S5.

Model selection - To estimate the number of vertebral regions in each specimen, we used a likelihood-based approach to select between the six regionalization hypotheses (1 region, 2 regions, 3 regions, etc.). First, the best segmented regression model – consisting of a series of slopes and breakpoints – was selected for each hypothesis by minimizing the total residual sums of squares (Step 5, Figure S2). Next, the six hypotheses were compared using the Akaike Information Criterion (AIC). The corrected AIC value was calculated for each hypothesis as follows:

$$AICc = nlog(\sigma^2) + 2K + \frac{2K(K+1)}{n-K-1}$$

Where σ^2 is the total RSS divided by *n*, *n* is the number of variables (number of PCOs multiplied by the number of vertebrae), and K is the number of parameters estimated. K was calculated as:

$$K = 2rv + (r-1)$$

where *r* is the number of regions, and *v* is the number of PCOs. This reflects two parameters for each regression (slope and intercept), and one for each breakpoint between regions.

Regionalization score - Since there may be multiple region hypotheses that fit the data well, we calculated a 'regionalization score' – a continuous variable that represents the degree of regionalization (Step 6, Figure S2). The regionalization score was calculated from the AICc difference of each hypothesis relative to the best hypothesis (ΔAICc). From this, the Akaike weight *w* was calculated:

$$w_i = \frac{exp\left(\frac{-1}{2}\Delta i\right)}{\sum_{r=1}^{R} exp\left(\frac{-1}{2}\Delta i\right)}$$

The regionalization score is the sum of the region hypotheses (region number) weighted by the Akaike weight (*w*) for each region hypothesis*:*

$$Regionalizationscore = 1w_1 + 2w_2 + 3w_3 + 4w_4 + 5w_5 + 6w_6$$

The following is a worked example from *Didelphis virginiana* MCZ 62069:

Hypothesis	RSS	AICc	ΔAICc	w_i	Score
4	0.103	-755.4	0.0	0.578	2.311
5	0.064	-754.8	0.6	0.422	2.111
3	0.190	-724.0	31.4	<0.001	~0.0
6	0.043	-721.9	33.4	<0.001	~0.0
2	0.319	-695.5	59.9	<0.001	~0.0
1	0.777	-613.2	142.2	<0.001	~0.0
				Total Regionalization Score	**4.42**

If all six region hypotheses are equally likely based on Akaike weight, a regionalization score of 3.5 would be recovered. Though this situation is unlikely to occur, we verified this in our dataset by examining the Akaike weight support of the best model. For six regions, this can vary from 0.17 (equal fit of all models), to 1 (perfect fit of best model). In our dataset, the lowest weighting recovered was 0.43, with an average of 0.90 (Table S5). This can be further confirmed by calculating the standard deviation of the Akaike weights for each specimen. A standard deviation of zero indicates equal fit of all models, whereas a standard deviation of 0.41 indicates that only the best model is supported (weight of one, all others have weight of zero). For taxa fit with six regions, standard deviations ranged from 0.2 to 0.41, indicating that equal-weighting bias was not an issue (Table S5).

Region breaks - The position of the region breaks was determined as the position of the segmented regression breaks in the best model based on the maximized regionalization score (Step 7, Figure S2). Confidence intervals on the location of region breaks was calculated from the standard deviation of the position of each break in the top 5% of models for the given region hypothesis (e.g., five regions).

Goodness-of-fit – To compare how well the data fit the segmented regression model between taxa, we calculated an adjusted r-squared value for the best-fit model of each taxon based on the summed sums of squares from the individual regressions, using the function *'multvarrsq'* in *'regions'*. Taxa with r-squared values of less than 0.75 were excluded from

evolutionary modeling analyses as they may indicate high levels of noise (e.g., taphonomy in fossil specimens) or a poor fit of the model.

Heterogeneity

The heterogeneity of each specimen was measured as its multivariate intracolumnar disparity. The calculation takes the average of the variance of the logged, unscaled data for each specimen across the measures.

Evolutionary patterns

Phylogeny - A composite, time-calibrated phylogeny was constructed for the sampled taxa. Relationships and branch lengths of extant taxa were gathered from *'timetree'* (www.timetree.org, (*37*)).). Non-mammalian synapsids were added based on conventional relationships from the literature (*38*), with branch lengths based upon first occurrences in the fossil record (Figure S3).

Trait mapping - Maximum-likelihood ancestral state reconstructions were calculated using the fastAnc function in the *'phytools'* package for R (*39*). Evolution of these traits on the tree was visualized using a heat map in the *'cont.map'* function. To aide with visualizing patterns, the data were cropped at a regionalization score of five for Figure 2, as only a few taxa exceeded this value. A version with the full data range can be found in Figure S6.

PGLS - Correlated evolution of regionalization and heterogeneity was tested using phylogenetic least-squares regressions (PGLS) in the *'gls'* function in the package *'nlme',* weighted to correct for the non-ultrametric topology (*40*). Evolution was modeled using a correlation structure based on Pagel's Lambda (partial phylogenetic influence) using the *corLambda'* function of the *'ape'* package (*41*).. Since relationships between traits may vary between phylogenetic groups, the PGLS was first run as a total effects model with group (sauropsid, synapsid) as a factor. The interaction term was dropped as it was insignificant.

Adaptive shifts – To test the hypothesis of shifting adaptive regimes in synapsid evolution, we used Ornstein-Uhlenbeck (OU) modeling. The fit of both traits separately was compared to Brownian motion (random-walk, BM), and OU (pull toward an optimum) models, including both single optima (OU1) and multiple optima with adaptive shifts (multiOU). BM and OU models were fit using the package '*ouch'* (*42*). The positions of shifts in the multiOU models were determined both using a stepwise AIC procedure in '*Surface'*, and a bayesian reversible-jump Markov chain Monte Carlo approach in *'bayou'*(*43, 44*). Priors for the 'bayou' model were set as follows: alpha and sigma squared, half-cauchy distribution; theta, normal distribution; number of shifts, conditional Poisson distribution; and run over 100,000 generations. Significant shifts were determined using a posterior-probability cutoff of 0.3. The 'surface' models generated included some minor shifts occurring on terminal tips. These were excluded in the final hypothesis tests, in which we mapped the primary shifts occurring on internal nodes, because we are primarily interested in patterns along the synapsid lineage.

Monte Carlo simulation was used to select the best fitting model following (*45*). Using the package *'pmc',* pairwise comparisons of models were made by simulating data under the null (simpler) model and test (more complex) hypotheses 1000 times. This provides a more robust approach than interpretation of AICc alone, which can lead to high error rates, particularly on smaller phylogenies, and can be biased by the underlying structure of the tree (*45*). Data from both sets of simulations were then fit to both null and test models to produce a distribution of likelihood ratios for each, thereby considering any inherent biases in the phylogeny (*46*).

Comparing the actual likelihood ratio to that of the simulated distributions provides p-values for the test, along with confidence intervals for each of the estimated parameters (*45*).

Sensitivity Analysis

Sensitivity analyses were conducted to assess the influence of missing data on the measurement of vertebral regionalization. All analyses followed the segmented regression method described above, and were conducted on datasets with simulated missing data with one hundred replicates for each set of parameters. Five extant specimens were used as the basis for the sensitivity analysis: *Alligator missippiensis* (MCZ 81457), *Varanus bengalensis* (MCZ 43739), *Sphenodon punctatus* (MCZ 4702), *Zaglossus bruijnii* (MCZ 12414), *Mus musculus* (MCZ 59560), and one complete fossil specimen, *Thrinaxodon liorhinus* (BP/1/7199). For each species, datasets were simulated at 10%, 20%, and 30% missing data, as described below.

Missing vertebrae and variables - Missing vertebrae were simulated by removing complete rows from the dataset, per the proportion of missing data parameter, rounded to the nearest number of whole rows. Missing variables were simulated by removing complete columns from the dataset, per the proportion of missing data parameter, rounded to the nearest number of whole columns.

Removal of vertebrae had a significant effect on regionalization score (Figure S4). The effect was greatest when 20% or more of the vertebrae were removed, and had the largest impact on highly regionalized taxa (e.g., *Mus*). This likely reflects the fact that removing vertebrae is more likely to disrupt the regionalization signal when there are many short regions instead of few long regions. Therefore, we focused our fossil sampling on taxa with 80% or greater vertebral completeness, with most species displaying 90% or greater completeness (Table S4, exceptions: *Diadectes, Ophiacodon*).

Removal of variables had a much more limited effect on regionalization score. Despite the random removal of variables, all mean regionalization scores were within 10% of the true value. The limited effect of removing variables is likely due to the high integration across vertebral measures, such that even when variables are removed the PCO space remains relatively unchanged.

Missing cells - Missing data points were simulated by randomly replacing data cells with "NA"s, per the proportion of missing data parameter, rounded to the nearest number of whole cells. Analyses were conducted both with interpolation of missing data (filled) or with missing data left as "NA"s in the final analysis.

Missing data points had a significant effect on regionalization score (Figure S4). When missing data remained in the final analysis, there was a decrease in regionalization score as the amount of missing data increased. Effects were greatest in the most regionalized taxa (e.g., *Mus*) where regions tend to be shorter. In contrast, when short strings of missing data (two or less) were filled using interpolation, this effect was reduced and the mean bootstrapped score fell within 10% of the true value. For some taxa, there was up to around 10% inflation of regionalization score when data were interpolated. In our analysis, we interpolate short strings of missing data because this slight inflation is more conservative when testing the hypothesis of reduced regionalization in fossil taxa.

Intraspecific variability

Sensitivity analyses were also conducted to examine intraspecific variability in regionalization patterns. A sample of 10 cats (*Felis catus*, for specimen number see Figure S5)

was examined to provide a robust estimate of the intraspecific variability within a species. In addition, two specimens were analyzed for two other mammal species: mouse (*Mus musculus*) and echidna (*Tachyglossus aculeatus*), and one fossil taxon, *Edaphosaurus boanerges*. Finally, a juvenile specimen of alligator (*Alligator mississippiensis*) was examined to assess the impact of ontogenetic intraspecific variation.

The 10 cats examined all displayed five regions as the best-fit model, with regionalization scores ranging from 4.59 to 5.09, and a mean regionalization score of 4.88±0.06 s.e.m (Figure S5). The two mouse specimens and two echidna specimens displayed regionalization scores of 5.00 and 5.00, and 4.00 and 4.20 respectively, whereas the two *Edaphosaurus* specimens displayed regionalization scores of 3.00 and 3.02. The adult and juvenile alligator specimens displayed regionalization scores of 4.13 and 4.00 respectively. This suggests that intraspecific variability in regionalization patterns is limited compared to the taxonomic variation examined in the main analyses.

Supplementary Text

Variations in regionalization score between taxa

Mammals – Within mammals, regionalization scores ranged from 3.91-6, though most were between four and five (median: 4.76). In monotremes, we recover cervical, pectoral, anterior dorsal, and posterior dorsal regions, reflecting the poorly defined lumbar region in these taxa. Only three taxa (*Equus*, *Choloepus* and *Orycteropus*) approached 6 regions. *Choloepus* exhibits an extra region in the anterior column and is one of very few mammals to exhibit an aberrant cervical count (*47*). *Equus* exhibits an additional region anterior to the sacrum, coincident with the development of unique accessory articulations between the transverse processes that have been linked to elevated thoracolumbar counts and lumbosacral stabilization (*48*). The additional region in *Orycteropus* is in a similar position and similarly reflects antero-posteriorly expanded transverse processes.

Sauropsids – Four regions were consistently favored for extant sauropids, except for the two monitor lizards (genus: *Varanus)* which displayed higher scores. The limited sampling here precludes any detailed conclusions regarding variation in sauropsids, but it is interesting to note high regionalization in these highly-active lepidosaurs.

Anamniotes - The fossil outgroup taxon for amniotes in our analysis was the diadectomorph *Diadectes tenuitectes*. Obtaining complete material appropriate for CT scanning for diadectomorphs proved challenging, and our specimen, which consists of an articulated series of 17 presacral vertebrae, is likely incomplete. The presacral count for *Diadectes* is thought to be 21; therefore this specimen likely consists of postaxial vertebrae five to 21, meaning cervicals three and four are missing from the analysis (*49*). Regionalization analysis recovered two presacral regions for this specimen (regionalization score=2.01, Table S5). However, the cervical region in diadectomorphs is quite short - consisting of five or six vertebrae (*50*). Therefore, we predict that *Diadectes* has three regions, and that the cervical region has been missed due to the missing vertebrae. This is supported by the patterns recovered in the two extant salamander taxa sampled.

Non-mammalian synapsids - Though three regions were most common in 'pelycosaurs', *Sphenacodon* was an outlier with a lower regionalization score of 2.14 (Table S5, Figure 2). The relatively high levels of missing data (76% data completeness) and the poor fit of the model (r-squared: 0.63) indicate that the quality of this specimen (YPM 818) may be insufficient to detect

the full regionalization pattern. Four regions were supported for non-mammalian therapsids, except for *Massetognathus* and *Lystrosaurus.* In both cases four regions was the next best supported model, and long strings of missing data (e.g., *Massetognathus*: neural spines largely missing, v4-v11 zygapophyses damaged, *Lystrosaurus*: v12-v20 neural spines missing) may have influenced this result.

Evolutionary modelling

To examine the relationship between regionalization and heterogeneity we used a phylogenetically-corrected regression analysis (PGLS). Superficially, regionalization and heterogeneity both increase in synapsids, reflected by a significant ($p=<0.001$) but weak ($r\text{-squared}=0.54$) relationship when using raw data. (Figure S7, Table S6). However, this relationship disappears when phylogenetic covariation between taxa is considered. Maximum likelihood estimation suggests moderate-to-high phylogenetic signal ($lambda=0.84$) and no significant effect of group or slope (Table S6). Comparing non-mammalian synapsids and mammals in Figure S7 illustrates that although mammals generally have higher values than non-mammalian synapsids on both axes, there is little covariation in the traits within the groups. This suggests that these two traits have both increased, but are not evolving in a coordinated manner along branches of the phylogeny.

By modeling the evolution of each trait separately, we can test for more complex evolutionary scenarios. We tested the hypothesis that vertebral regionalization and heterogeneity may be evolving toward an adaptive optimum (OU1), and that this optimum may have shifted over the course of synapsid evolution (multiOU). Based on the AICc, the multiOU models outperformed the single optimum OU1 model and the Brownian motion (BM) model for both regionalization and heterogeneity (Table S7). Two alternative multiOU models were considered, in which the locations of adaptive shifts were obtained either by Bayesian (Bayou model) or likelihood (Surface model) approaches. Model parameters and their confidence intervals, obtained by simulation, can be found in Table S7. For regionalization, primary shifts were reconstructed at Therapsida and Theria using Surface, whereas Bayou supported only the earlier shift (Figure 3). For heterogeneity, primary shifts were reconstructed at the base of Cynodontia and Boreotheria using Surface, whereas Bayou supported a single shift at Therapsida (Figure 3). It should be noted that although the early shifts are well constrained by fossil taxa, the locations of the later shifts are considered preliminary due to the lack of fossil sampling near the base of mammals.

Although the Surface model fit the data best when evaluated using AICc, we conducted pairwise hypothesis tests to confirm this finding. We compared the relative likelihood of the Surface model to null distributions relative to the other models, determined using simulation (*45*). Null distributions for each test can be found in Figure S8. In each case the simulated distributions of the likelihood ratios produce clear and separate peaks, indicating that this topology contains sufficient information to distinguish between the hypotheses (*45*). Further, the actual likelihood ratio of the models recovered in our analyses (vertical line, Figure S8) clearly lies within the test distribution (Surface model, blue peak) and not the null distribution (BM/OU1/Bayou, red peak). This provides clear support for the Surface model over the other models, producing highly significant p-values (Table S7).

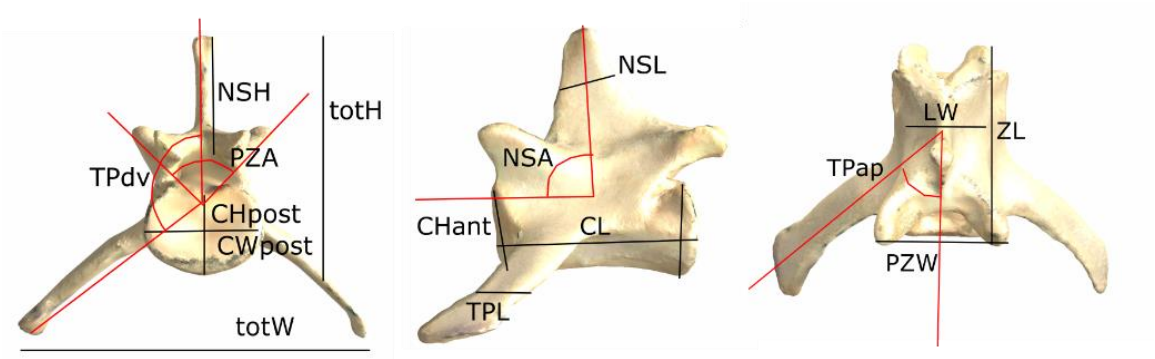


Fig. S1. Measurements

Linear and angular measurements. See Table S3 for measurement key.

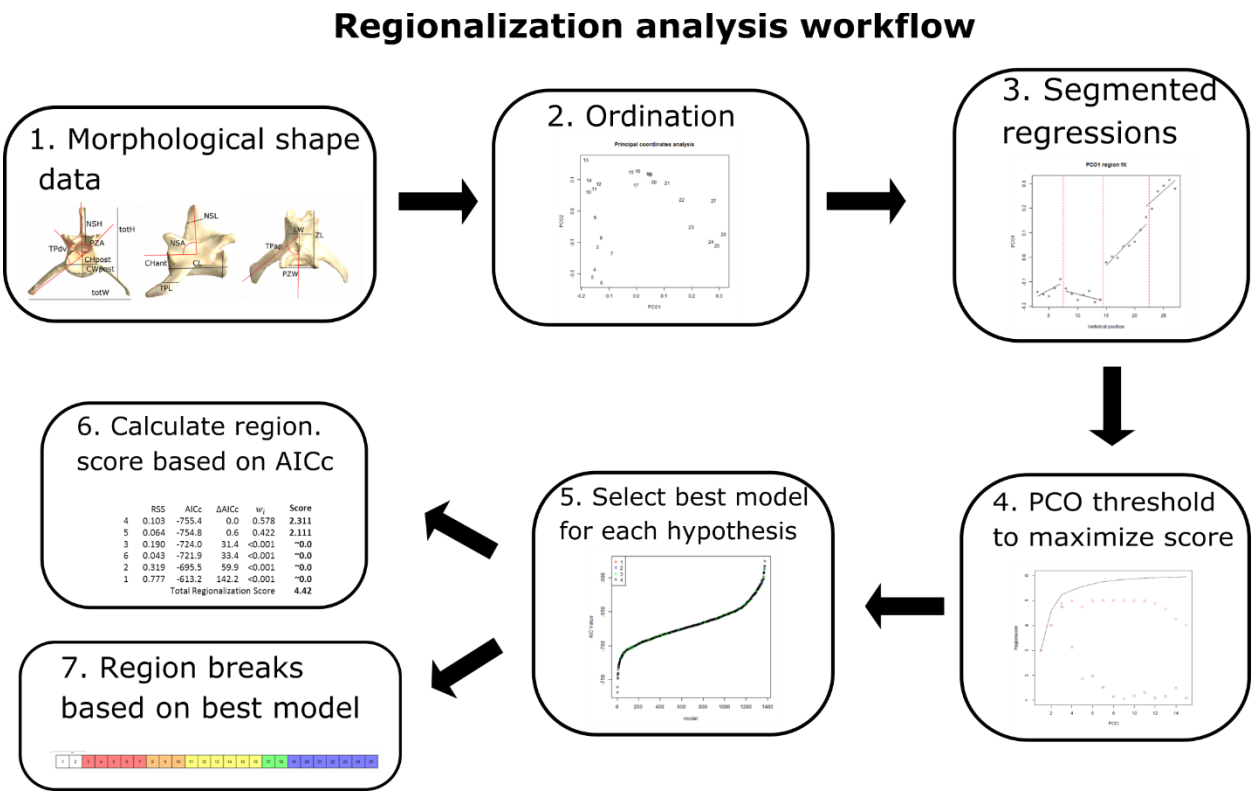


Fig. S2. Regionalization analysis workflow

Summary of workflow for regionalization analysis. See text for details.

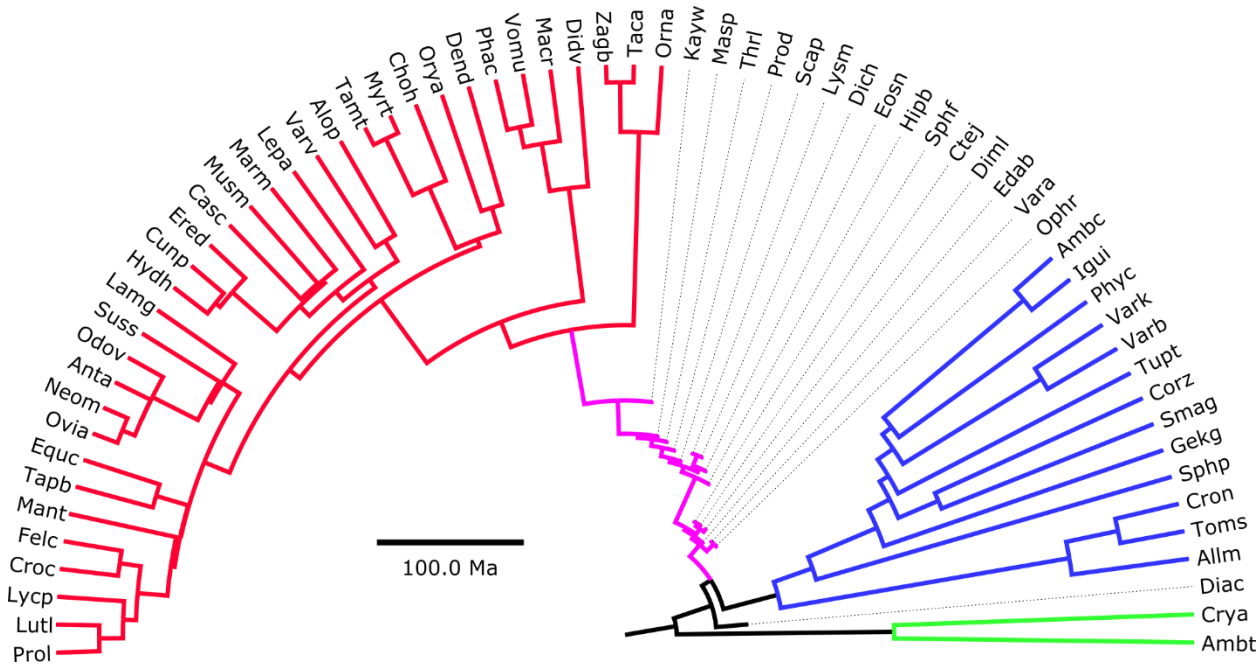


Fig. S3. Phylogeny

Composite phylogeny used for analyses. Backbone for fossils based on (*38*). Abbreviations are as follows: Genus species– [Gen][s]. For full taxonomic names see Tables S1 and S2. Red: mammals; pink: non-mammalian synapsids; blue: sauropsids; green: amphibians.

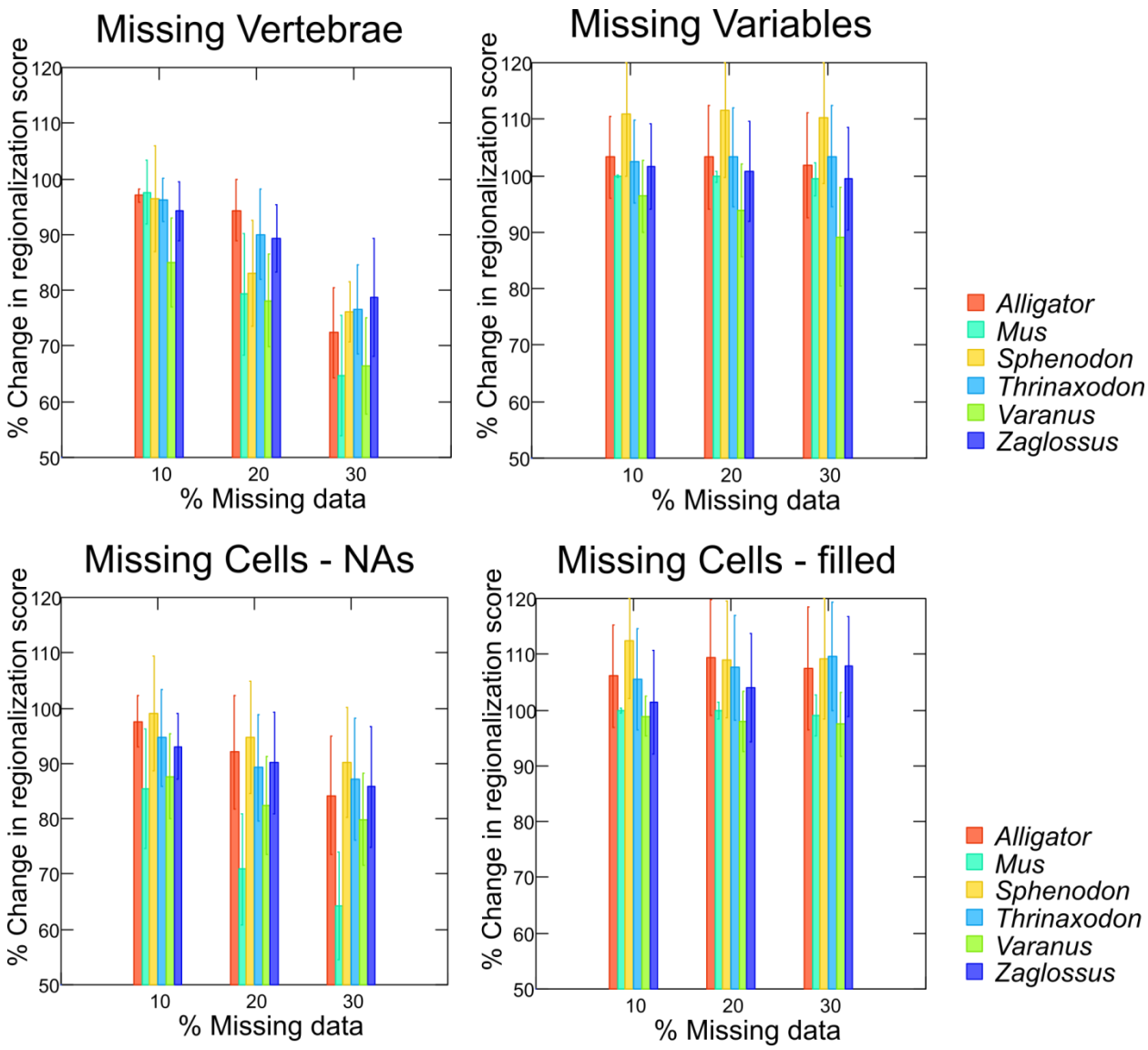


Fig. S4. Sensitivity analysis

Upper: Mean regionalization score from 100 replicates with the random removal of vertebrae and variables. Scores expressed as a percentage of the regionalization score of the full dataset. Error bars indicate standard deviation, also expressed as a percentage of the full score.
Lower: Mean regionalization score after random removal of data points (cells), both with and without interpolation of missing data. Scores expressed as a percentage of the regionalization score of the full dataset. Error bars indicate standard deviation, also expressed as a percentage of the full score.

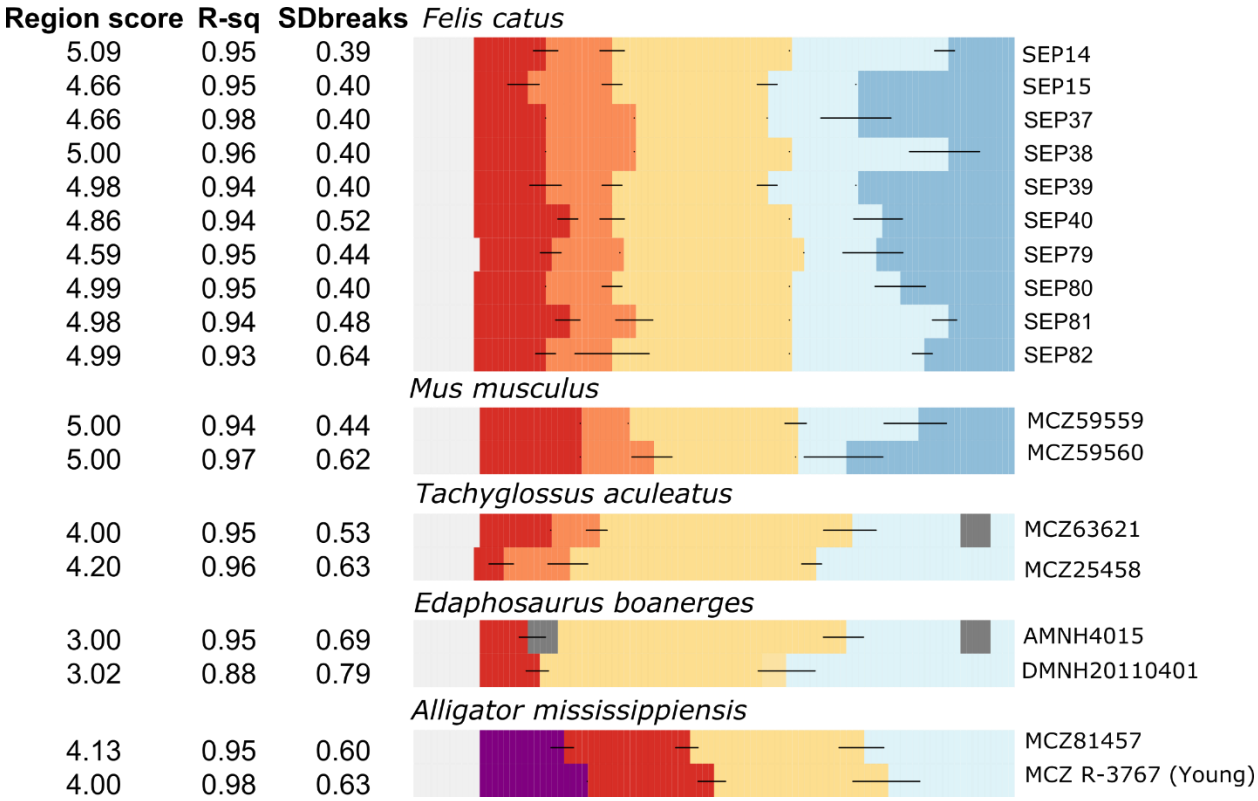


Fig. S5. Intraspecific variation

Intraspecific variability in regionalization score and boundaries. Intraspecific variability was examined in five species. Colors indicate vertebral regions regions, as in Fig 4 of the main text, and bars represent standard deviation across top 5% of models, expressed as a percentage of column length. R-sq: r-squared of best model; SDbreaks: standard deviation of breaks in units of vertebrae.

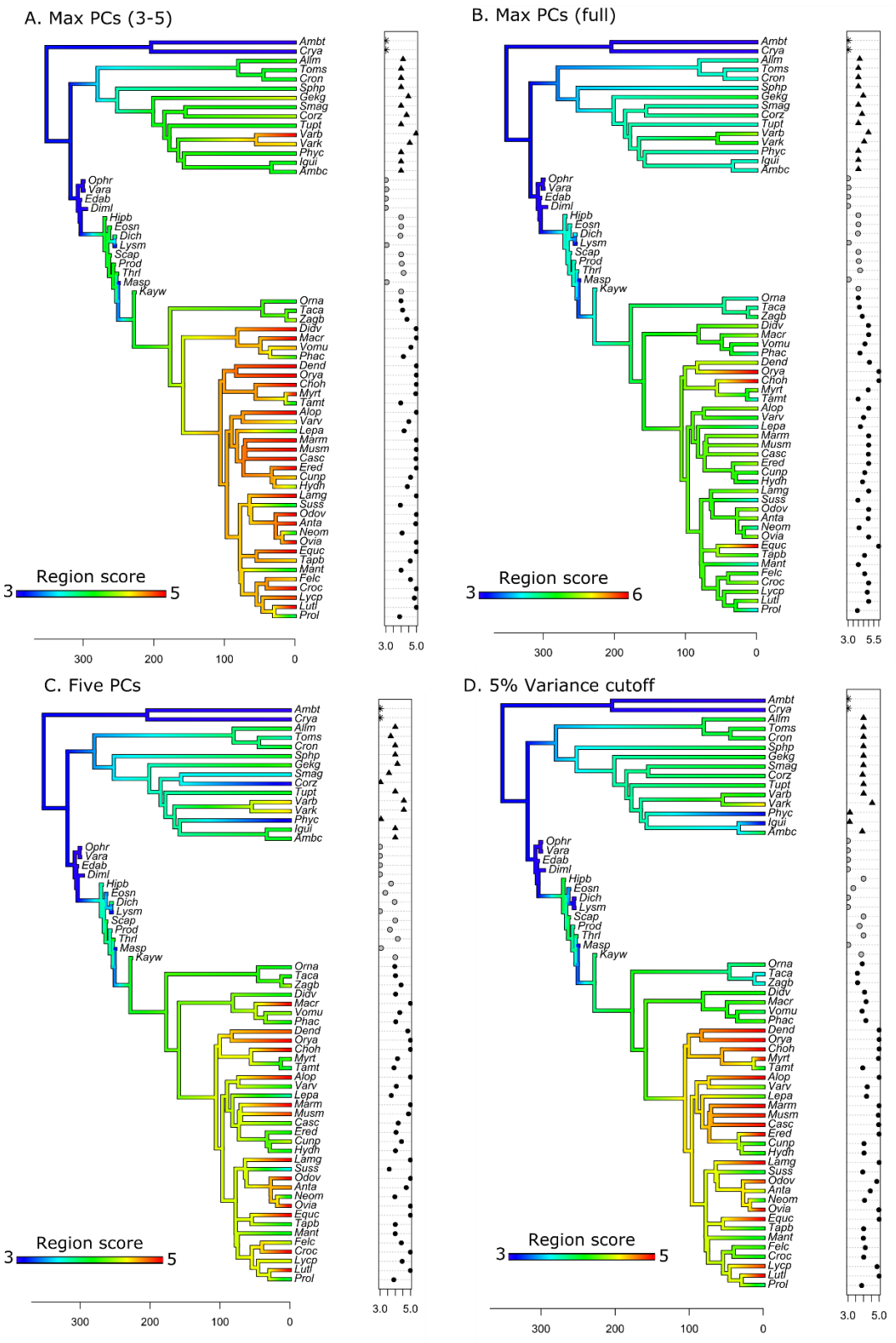

A. Max PCs (3-5)
B. Max PCs (full)
C. Five PCs
D. 5% Variance cutoff
Region score
3
5
6
300
200
100
0
3.0
5.0
5.5
Ambt Crya Allm Toms Cron Sphp Gekg Smag Corz Tupt Varb Vark Phyc Iguí Ambc
Ophr Vara Edab Diml
Hipb Eosn Dich Lysm Scap Prod Thrl Masp Kayw
Orna Taca Zagb Didv Macr Vomu Phac Dend Orya Choh Myrt Tamt Alop Varv Lepa Marm Musm Casc Ered Cunp Hydh Lamg Suss Odov Anta Neom Ovia Equc Tapb Mant Felc Croc Lycp Lutl Prol

Fig. S6. Selection of PCOs.

Comparison of regionalization scores calculated with different parameters. A. Based on top five PCOs with maximum of five regions (as in main text), B. Based on top five PCOs with maximum of six regions. Note that even with an increased maximum number of regions, most mammals have scores close to five with only two taxa approaching six. Therefore, Figure 2 of the main text displays a maximum of five to allow clearer visualization of variation within mammals (i.e., A). C. Based on PCOs exceeding five percent contribution to total variation. D. Based on number of PCOs that produces the maximum regionalization score. Note that although individual values may vary, trends remain similar despite parameters used in the analysis. Even when regionalization score is maximized, basal synapsids display reduced regionalization relative to mammals

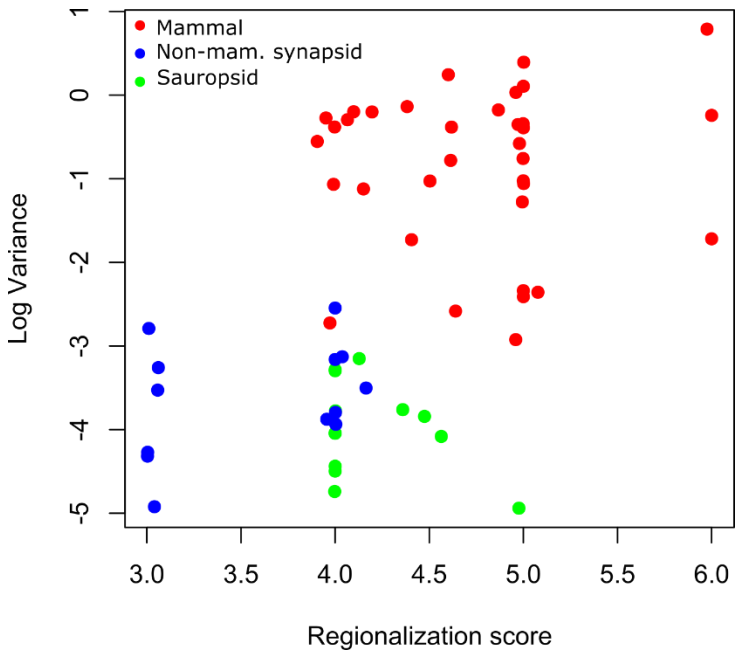


Fig. S7. Heterogeneity versus regionalization

Relationships between regionalization and heterogeneity. PGLS tests are described in Table S6.

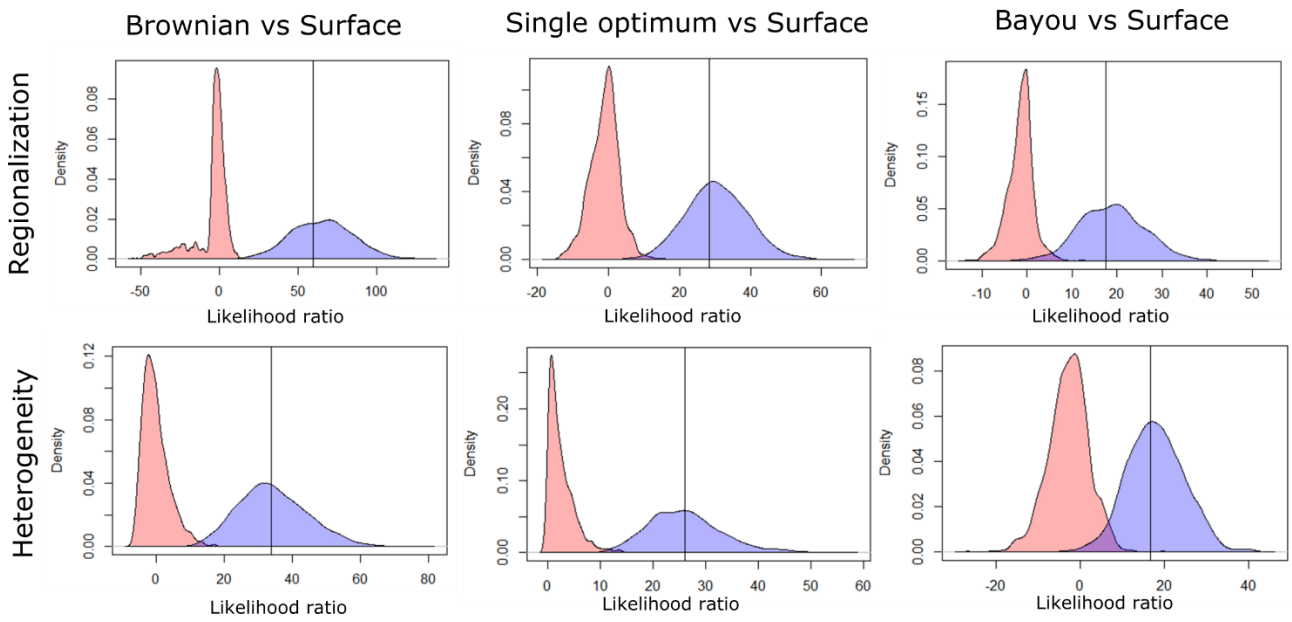


Fig. S8. Likelihood ratio tests

Hypothesis testing evolutionary models. The simpler null hypotheses of Brownian motion, single optimum or Bayou shifts are compared to the more complex Surface hypothesis. Monte Carlo simulations (n=1000) were performed under both the null (red, BM, OU1 or Bayou respectively) model and test (blue, Surface) hypothesis, and the relative fit was assessed using likelihood in both scenarios. Separate null and test peaks indicate that there is power to discern between the models given the phylogeny and taxonomic sampling, whereas strong overlap between peaks indicates weak power. The observed likelihood ratio is indicated by the vertical line. The observed ratio falls within the test distribution (blue), indicating support for the test model over the null model. P-values and confidence intervals can be found in Table S7.

Table S1. Extant sample

Extant sample. For abbreviations see Table S2 footnotes.

Species	**Common name**	**Specimen no.**
Anamniotes		
Ambystoma tigrinum	Tiger salamander	TNHC 17991
Cryptobranchus alleganiensis	Hellbender	GMUCL W16
Sauropsids		
Amblyrhynchus cristatus	Marine iguana	MCZ 2006
Iguana iguana	Green iguana	MCZ 182895
Physignathus cocincinus	Chinese water dragon	MCZ 43732
Varanus bengalensis	Bengal monitor	MCZ 43739
Varanus komodoensis	Komodo dragon	MCZ 24907
Tupinambis teguixin	Gold tegu	FMNH 217382
Corucia zebrata	Solomon Islands skink	MCZ 72918
Smaug giganteus	Giant girdled lizard	MCZ 173607
Gekko gecko	Tokay gecko	MCZ 13158
Sphenodon punctatus	Tuatara	MCZ 4702
Crocodylus niloticus	Nile crocodile	RVC 'Flunch'
Tomistoma schlegelii	False gharial	MCZ 12459
Alligator missippiensis	American alligator	MCZ 81457, MCZ 3767
Mammals		
Didelphis virginiana	Virginia opossum	MCZ 62096
Phascolarctos cinereus	Kaola	MCZ 58136
Vombatus ursinus	Common wombat	MCZ 24974
Macropus robustus	Common wallaroo	MCZ 63609
Crocuta crocuta	Spotted hyeana	MCZ 20968
Felis Catus	Housecat	SEP14, SEP15, SEP 37, SEP 38, SEP 39, SEP 40, SEP 79, SEP 80, SEP 81 SEP 82
Lycaon pictus	African wild dog	MCZ 13233
Procyon lotor	Racoon	MCZ 7101
Lutra lutra	Eurasian otter	UMZC K2768
Manis temminckii	Ground pangolin	MCZ 34184
Equus caballus	Horse	MCZ 14915
Tapirus bairdii	Baird's tapir	MCZ 1076
lama glama	Llama	MCZ BOM1881
Ovis aries	Sheep	MCZ 6338
Neotragus moschatus	Suni	MCZ 53804
Antilocapra americana	Pronghorn	MCZ 1773
Odocoileus virginianus	White-tailed deer	MCZ 46590
Sus scrofa	Pig	MCZ 6246
Lepus americanus	Snowshoe hare	MCZ 852

Hydrochoerus hydrochaeris	Capybara	MCZ 6013
Cuniculus paca	Lowland paca	MCZ 829
Erethizon dorsatum	North American porcupine	MCZ 965
Mus musculus	House mouse	MCZ 59560, MCZ 59559
Castor canadensis	North American beaver	MCZ 64159
Marmota monax	Groundhog	MCZ 377
Varecia veregata	Black-and-white ruffed lemur	MCZ 18740
Alouatta palliata	Mantled howler monkey	MCZ 47267
Tamandua tetradactyla	Southern tamandua	MCZ 20965
Myrmecophaga tridactyla	Giant anteater	MCZ 20969
Choloepus hoffmani	Hoffmann's two-toed sloth	MCZ 12348
Dendrohyrax dorsalis	Western tree hyrax	MCZ 6069
Orycteropus afer	Aadvark	MCZ 20970
Ornithorhynchus anatinus	Platypus	USNM 221110
Zaglossus bruijnii	Western long-beaked echidna	MCZ 12414
Tachyglossus aculeatus	Short-beaked echidna	MCZ 63621, MCZ 25458

Table S2. Fossil sample

Fossil sample and scanning information.

Family	Species	Sp. No	Scan facility	Scanner type	Power	Filter (mm)	Voxel size (mm)	Digital data accession
Anomodontia	*Dicynodon huenei*	TSK 14	UMZC	Nikon XT 225 ST	150kV, 146µA	0.5Cu	0.0305-0.0636	UMZC/MCZ
Anomodontia	*Eosimops newtoni*	BP/1/6674	ESI	Nikon Metrology XTH 225/320 LC	110kV, 110µA	0.5Cu	0.1297	ESI
Anomodontia	*Lystrosaurus murrayi*	UMCZ T763	UMZC	Nikon XT 225 ST	160kV, 160µA	0.5Cu	0.0219-0.0745	UMZC
Biarmosuchia	*Hipposaurus boonstrai*	SAM-PK-8950	ESI	Nikon Metrology XTH 225/320 LC	215kV, 525µA	1.2Cu	0.0444-0.0499	ESI
Cynodontia	*Kayentatherium wellesi*	MCZ 8812	CNS	Nikon Metrology (X-Tek) HMXST225	124kV, 165µA	1Cu	0.127	MCZ
Cynodontia	*Massetognathus pascuali*	MCZ 3691	CNS	Nikon Metrology (X-Tek) HMXST225	175kV, 46µA	0.01Cu	0.1272	MCZ
Cynodontia	*Procynosuchus delaharpaea*	TSK 34	UMZC	Nikon XT 225 ST	150kV, 150µA	0.5Cu	0.0179-0.0433	UMZC/MCZ
Cynodontia	*Thrinaxodon liorhinus*	BP/1/7199	ESRF	Beamline ID17	96keV	-	0.0455	ESI
Diadectidae	*Diadectes tenuitectes*	FMNH UC 650	UC	GE custom-built dual tube CT	200kV, 220µA	0.5Sn	0.0734	FMNH
Edaphosauridae	*Edaphosaurus boanerges*	DMNH 2011-04-01	Structured light scanner	Creaform Go Scan 20	-	-	-	MCZ
Ophiacodontidae	*Ophiacodon retroversus*	FMNH UC 458	UC	GE custom-built dual tube CT	200kV, 300µA	0.5Sn	0.0872-0.1125	FMNH
Ophiacodontidae	*Varanosaurus acutirostris*	AMNH FARB 4174	CNS	Nikon Metrology (X-Tek) HMXST225	165kV, 150µA	1Cu	0.0561	AMNH
Sphenacodontidae	*Ctenorhachis jacksoni*	USNM 437710	Varian medical systems	Varian M3 accelerator, BIR 800D-Linear array	1mV	3.3Steel	0.4	FMNH
Sphenacodontidae	*Dimetrodon limbatus*	AMNH FARB 4008	CNS	Nikon Metrology (X-Tek) HMXST225	200mV, 160µA	1Cu	0.127	AMNH
Sphenacodontidae	*Sphenacodon ferox*	YPM 818	Brown University	Philips Medical System	110kV, 31µA	-	0.4861	YPM

Therocephalia	*Scalaposaurus punctatus*	UMZC T837	UMZC	Nikon XT 225 ST	165kV, 469µA	0.01Cu	0.047	UMZC

UMZC: University Museum of Zoology, Cambridge University; ESI: Evolutionary Studies Institute, University of Witwatersrand; CNS: Center for Nanoscale Systems, Harvard University; ESRF: European Synchrotron Radiation Facility; UC: University of Chicago; NHM: Natural History Museum, London; FMNH: Field Museum of Natural History; AMNH: American Museum of Natural History; YPM: Yale Peabody Museum; TSK: Tom Kemp Collection (being accession into the Natural History Museum, London); BP: Evolutionary Studies Institute; SAM: Iziko Museums of South Africa; DMNH: Dallas Museum of Natural History, GMUCL: Grant museum University College London.

Table S3. Measurements.

Description of measurements

Measurement	Description
CL	Centrum Length
CHPost	Centrum height (posterior)
CWPost	Centrum width (posterior)
CHant	Centrum height (anterior)
CWant	Centrum width (anterior)
ArchH	Arch height
ArchW	Arch width
MidLW	Mid-lamina width
NSL	Neural spine length
NSH	Neural spine height
NSA	Neural spine angle
TotH	Total height
TotW	Total width
PreZw	Pre-zygapophysis width
PreZA	Pre-zygapophysis joint angle
InterZL	Inter-zygapophyseal length
TPL	Transverse process/diapophysis length
TPDV	Transverse process/diapophysis dorsoventral angle
TPAP	Transverse process/diapophysis anteroposterior angle

Table S4. Completeness of fossils

Completeness of fossil samples update. PS: Presacral; var: variables; vert: vertebrae; data: data points.

Species	Sp.no.	PS count	N (var)	%var	N (vert)	%vert	%data (filled)	%data	%mean
Ctenorhachis jacksoni	USNM 437710	27	16	84.2	24	96.0	93.0	78.6	91.1
Diadectes tenuitectes	FMNH UC 650	21	15	78.9	17	89.5	98.8	88.2	89.1
Dicynodon huenei	TSK 14	26	19	100.0	24	100.0	95.0	87.7	98.3
Dimetrodon limbatus	AMNH FARB 4008	25	17	89.5	23	100.0	93.6	85.7	94.4
Edaphosaurus boanerges	DMNH 2011-04-01	24	19	100.0	22	100.0	100.0	95.9	100.0
Edaphosaurus boanerges	AMNH 4015	25	15	78.9	21	91.3	100.0	100.0	90.1
Eosimops newtoni	BP/1/6674	29	15	78.9	27	100.0	95.1	83.7	91.3
Hipposaurus boonstrai	SAM-PK-8950	27	13	68.4	23	92.0	92.3	77.9	84.2
Kayentatherium wellesi	MCZ 8812	25	19	100.0	23	100.0	99.3	90.8	99.8
Lystrosaurus murrayi	UMZC T763	25	19	100.0	21	91.3	93.2	82.7	94.8
Massetognathus pascuali	MCZ 3691	28	19	100.0	25	96.2	85.5	78.3	93.9
Ophiacodon retroversus	FMNH UC 458	27	12	63.2	20	80.0	93.8	74.2	79.0
Procynosuchus delaharpaea	TSK 34	28	19	100.0	26	100.0	98.8	92.9	99.6
Scalaposaurus punctatus	UMZC T837	27	19	100.0	25	100.0	90.1	87.8	96.7
Sphenacodon ferox	YPM 818	23	17	89.5	20	95.2	87.1	76.8	90.6
Thrinaxodon liorhinus	BP/1/7199	27	19	100.0	25	100.0	100.0	100.0	100.0
Varanosaurus acutirostris	AMNH FARB 4174	23	13	68.4	19	90.5	97.6	84.2	85.5

Table S5. Regionalization and heterogeneity results

Sp. No: specimen number; Reg. score: regionalization score; RSS: residual sums of squares; AICc: corrected AIC score; Ak.weight: Akaike weight; SD (Ak. W): Standard deviation of Akaike weights; R-sq: R-squared of best model; Hetero: heterogeneity. Breaks indicate the vertebral position of region breaks for the best fit model, with the break occurring posterior to the vertebra named. Akaike weights reflect relative probabilities of the best fit hypothesis. Bold: fossil species.

Species	Sp. No	Reg. score	Best model	Breaks					RSS	AICc	Ak. weight	SD (Ak.W)	R-sq	Hetero.
Anamniotes														
Diadectes tenuitectes	**FMNH UC 650**	**2.01**	**2**	**10**	**0**	**0**	**0**	**0**	**0.169**	**-255**	**0.99**	**0.44**	**0.71**	**0.009**
Ambystoma tigrinum	TNHC 17991	2.99	3	4	11	0	0	0	0.053	-327	0.99	0.50	0.90	0.008
Cryptobranchus alleganiensis	LDUCL3AF??	3.00	3	3	13	0	0	0	0.079	-312	1.00	0.41	0.86	0.007
Sauropsids														
Alligator missippiensis	MCZ 81457	4.13	4	6	11	18	0	0	0.045	-548	0.87	0.35	0.95	0.043
Alligator missippiensis	MCZ 3767	4.00	4	7	12	19	0	0	0.013	-288	1.00	0.41	0.98	0.634
Amblyrhynchus cristatus	MCZ 2006	4.00	4	7	14	21	0	0	0.064	-676	1.00	0.41	0.92	0.018
Corucia zebrata	MCZ 72918	4.14	4	6	12	22	0	0	0.041	-293	0.43	0.20	0.96	0.024
Crocodylus niloticus	RVC 'Flunch'	4.00	4	4	9	13	0	0	0.033	-536	1.00	0.41	0.96	0.037
Gekko gecko	MCZ 13158	4.47	4	5	11	22	0	0	0.084	-564	0.59	0.25	0.91	0.021
Iguana iguana	MCZ 182895	4.00	4	6	8	16	0	0	0.033	-496	1.00	0.41	0.95	0.012
Physignathus cocincinus	MCZ 43732	4.00	4	7	11	17	0	0	0.016	-278	1.00	0.41	0.98	0.011
Smaug giganteus	MCZ 173607	4.00	4	5	9	17	0	0	0.029	-253	1.00	0.41	0.94	0.009
Sphenodon punctatus	MCZ 4702	4.00	4	5	11	19	0	0	0.075	-540	1.00	0.41	0.90	0.023
Tomistoma schlegelii	MCZ 12459	4.00	4	8	12	21	0	0	0.012	-291	1.00	0.41	0.98	0.038
Tupinambis teguixin	FMNH 217382	4.00	4	7	12	22	0	0	0.041	-594	1.00	0.41	0.95	0.018
Varanus bengalensis	MCZ 43739	4.98	5	6	9	14	22	0	0.042	-698	0.98	0.40	0.94	0.007
Varanus komodoensis	MCZ 24907	4.56	5	6	8	12	22	0	0.045	-899	0.56	0.26	0.95	0.017
Synapsids														

Ctenorhachis jacksoni	**USNM 437710**	**3.11**	**3**	**7**	**19**	**0**	**0**	**0**	**0.247**	**-654**	**0.89**	**0.36**	**0.73**	**0.024**
Dicynodon huenei	**TSK 14**	**3.96**	**4**	**6**	**13**	**17**	**0**	**0**	**0.125**	**-688**	**0.96**	**0.39**	**0.82**	**0.021**
Dimetrodon limbatus	**AMNH FARB 4008**	**3.00**	**3**	**8**	**16**	**0**	**0**	**0**	**0.071**	**-256**	**1.00**	**0.41**	**0.86**	**0.014**
Edaphosaurus boanerges	**AMNH 4015**	**3.00**	**3**	**5**	**18**	**0**	**0**	**0**	**0.027**	**-265**	**1.00**	**0.41**	**0.95**	**0.048**
Edaphosaurus boanerges	**DMNH 2011-04-01**	**3.02**	**3**	**5**	**15**	**0**	**0**	**0**	**0.094**	**-527**	**0.98**	**0.40**	**0.88**	**0.074**
Eosimops newtoni	**BP/1/6674**	**4.00**	**4**	**6**	**14**	**19**	**0**	**0**	**0.090**	**-660**	**1.00**	**0.41**	**0.87**	**0.022**
Hipposaurus boonstrai	**SAM-PK-8950**	**4.00**	**4**	**6**	**13**	**24**	**0**	**0**	**0.059**	**-396**	**1.00**	**0.41**	**0.93**	**0.042**
Kayentatherium wellesi	**MCZ 8812**	**4.00**	**4**	**6**	**9**	**17**	**0**	**0**	**0.062**	**-727**	**1.00**	**0.41**	**0.93**	**0.078**
Lystrosaurus murrayi	**UMZC T763**	**3.03**	**3**	**8**	**19**	**0**	**0**	**0**	**0.012**	**-130**	**0.97**	**0.39**	**0.97**	**0.028**
Massetognathus pascuali	**MCZ 3691**	**3.06**	**3**	**12**	**24**	**0**	**0**	**0**	**0.150**	**-754**	**0.94**	**0.38**	**0.83**	**0.034**
Ophiacodon retroversus	**FMNH UC 458**	**3.00**	**3**	**8**	**21**	**0**	**0**	**0**	**0.102**	**-194**	**1.00**	**0.41**	**0.77**	**0.013**
Procynosuchus delaharpaea	**TSK 34**	**4.04**	**4**	**5**	**12**	**18**	**0**	**0**	**0.006**	**-176**	**0.95**	**0.38**	**0.99**	**0.044**
Scalaposaurus punctatus	**UMZC T837**	**4.00**	**4**	**7**	**13**	**18**	**0**	**0**	**0.018**	**-139**	**1.00**	**0.41**	**0.97**	**0.020**
Sphenacodon ferox	**YPM 818**	**2.14**	**2**	**15**	**0**	**0**	**0**	**0**	**0.190**	**-190**	**0.86**	**0.35**	**0.63**	**0.023**
Thrinaxodon liorhinus	**BP/1/7199**	**4.16**	**4**	**7**	**12**	**19**	**0**	**0**	**0.063**	**-816**	**0.83**	**0.33**	**0.94**	**0.030**
Varanosaurus acutirostris	**AMNH FARB 4174**	**3.04**	**3**	**4**	**14**	**0**	**0**	**0**	**0.052**	**-204**	**0.96**	**0.39**	**0.86**	**0.007**
Alouatta palliata	MCZ 47267	5.00	5	6	10	19	22	0	0.024	-672	1.00	0.41	0.98	0.347
Antilocapra americana	MCZ 1773	4.96	5	6	8	14	18	0	0.004	-350	0.96	0.39	0.99	1.032
Castor canadensis	MCZ 64159	4.97	5	6	11	17	21	0	0.022	-449	0.97	0.39	0.97	0.704
Choloepus hoffmani	MCZ 12348	6.00	6	5	8	13	27	30	0.036	-1072	1.00	0.41	0.97	0.785
Crocuta crocuta	MCZ 20968	4.99	5	7	10	18	22	0	0.045	-799	0.99	0.41	0.95	0.279
Cuniculus paca	MCZ 829	4.62	5	9	15	20	24	0	0.005	-337	0.62	0.27	0.99	0.682
Dendrohyrax dorsalis	MCZ 6069	5.08	5	6	9	27	31	0	0.077	-604	0.90	0.36	0.95	0.095
Didelphis virginiana	MCZ 62096	4.98	5	5	14	21	24	0	0.012	-319	0.98	0.40	0.98	0.560
Equus caballus	MCZ 14915	5.98	6	7	9	16	25	29	0.028	-1002	0.98	0.40	0.97	2.200
Erethizon dorsatum	MCZ 965	5.00	5	7	10	17	24	0	0.008	-342	1.00	0.41	0.99	0.468
Felis Catus	MCZ 68415	5.09	5	6	9	17	24	0	0.046	-608	0.91	0.36	0.95	0.436
Felis Catus	MCZ 68416	4.66	5	5	9	16	20	0	0.047	-607	0.66	0.28	0.95	0.382

Felis Catus	SEP 37	4.66	5	6	10	16	20	0	0.018	-300	0.82	0.32	0.98	0.472
Felis Catus	SEP 38	5.00	5	6	10	17	24	0	0.032	-455	1.00	0.41	0.96	0.469
Felis Catus	SEP 39	4.98	5	6	9	16	20	0	0.060	-762	0.98	0.40	0.94	0.475
Felis Catus	SEP 40	4.86	5	7	9	17	21	0	0.059	-583	0.52	0.21	0.94	0.460
Felis Catus	SEP 79	4.59	5	6	9	17	20	0	0.048	-739	0.59	0.26	0.95	0.472
Felis Catus	SEP 80	4.99	5	6	9	17	22	0	0.054	-776	0.99	0.41	0.95	0.454
Felis Catus	SEP 81	4.98	5	7	10	17	24	0	0.054	-592	0.98	0.40	0.94	0.466
Felis Catus	SEP 82	4.99	5	6	9	17	23	0	0.064	-754	0.99	0.40	0.93	0.444
Hydrochoerus hydrochaeris	MCZ 6013	4.41	4	7	15	21	0	0	0.063	-628	0.59	0.26	0.94	0.177
lama glama	MCZ BOM1881	5.00	5	6	8	15	19	0	0.029	-611	1.00	0.41	0.97	1.481
Lepus americanus	MCZ 852	4.20	4	5	12	18	0	0	0.059	-562	0.80	0.32	0.94	0.817
Lutra lutra	UMZC K2768	5.00	5	9	17	19	24	0	0.031	-649	1.00	0.41	0.96	0.359
Lycaon pictus	MCZ 13233	4.87	5	6	9	17	24	0	0.046	-608	0.87	0.35	0.95	0.837
Macropus robustus	MCZ 63609	5.00	5	7	10	18	22	0	0.028	-615	1.00	0.41	0.97	0.090
Manis temminckii	MCZ 34184	4.00	4	8	16	18	0	0	0.044	-716	1.00	0.41	0.95	0.683
Marmota monax	MCZ 377	5.00	5	7	10	16	19	0	0.025	-584	1.00	0.41	0.97	1.111
Mus musculus	MCZ 59560	5.00	5	7	9	16	21	0	0.049	-685	1.00	0.41	0.94	0.737
Mus musculus	MCZ 59559	5.00	5	7	10	16	18	0	0.022	-595	0.99	0.40	0.97	0.682
Myrmecophaga tridactyla	MCZ 20969	4.96	5	5	7	10	20	0	0.016	-624	0.96	0.39	0.97	0.054
Neotragus moschatus	MCZ 53804	4.07	4	7	11	20	0	0	0.058	-600	0.93	0.38	0.94	0.744
Odocoileus virginianus	MCZ 46590	5.00	5	5	8	17	22	0	0.028	-804	1.00	0.41	0.96	0.676
Ornithorhynchus anatinus	USNM 221110	3.99	4	7	11	21	0	0	0.033	-465	0.99	0.40	0.96	0.344
Orycteropus afer	MCZ 20970	6.00	6	7	11	15	19	25	0.009	-528	1.00	0.41	0.99	0.179
Ovis aries	MCZ 6338	5.00	5	6	8	17	22	0	0.045	-695	1.00	0.41	0.95	0.096
Phascolarctos cinereus	MCZ 58136	4.15	4	6	11	18	0	0	0.037	-456	0.85	0.34	0.96	0.326
Procyon lotor	MCZ 7101	3.91	4	9	18	22	0	0	0.085	-779	0.90	0.36	0.92	0.574
Sus scrofa	MCZ 6246	3.95	4	7	15	20	0	0	0.036	-458	0.95	0.38	0.95	0.760
Tachyglossus aculeatus	MCZ 63621	4.00	4	6	8	19	0	0	0.042	-592	1.00	0.41	0.95	0.776

Tachyglossus aculeatus	MCZ 25458	4.20	4	4	7	18	0	0	0.049	-653	0.80	0.32	0.96	0.863
Tamandua tetradactyla	MCZ 20965	3.97	4	6	9	21	0	0	0.021	-325	0.57	0.22	0.97	0.066
Tapirus bairdii	MCZ 1076	4.60	4	6	10	25	0	0	0.022	-380	0.57	0.23	0.97	1.275
Varecia veregata	MCZ 18740	4.50	5	6	9	17	23	0	0.031	-607	0.50	0.26	0.97	0.358
Vombatus ursinus	MCZ 24974	4.64	5	7	10	18	23	0	0.041	-538	0.74	0.29	0.95	0.076
Zaglossus bruijnii	MCZ 12414	4.38	4	5	7	17	0	0	0.057	-829	0.62	0.27	0.95	0.871

Table S6. Phylogenetic Generalized Least Squares analysis

Phylogenetic generalized least squares of heterogeneity (log variance) against regionalization score. Includes effects of slope (relationship between the variables) and grouping. Interactions were all insignificant and were removed from the model. Coeff.: coefficients; std. error: standard error. Lambda indicates the degree of phylogenetic signal in the model which was corrected for in the analysis.

	Coeff.	**Std. Error**	**t-value**	**p-value**	**AIC**
Uncorrected					
Intercept	-8.78	0.948	-9.263	<0.001	
Slope	1.152	0.214	5.369	<0.001	
Sauropsid-synapsid	2.095	0.361	5.811	<0.001	
Pagels lambda			***lambda=0.835***		**192.87**
Intercept	-4.18	1.441	-2.901	0.0052	
Slope	0.079	0.229	0.343	0.732	
Sauropsid-synapsid	0.288	1.174	0.245	0.807	

Table S7. Evolutionary models

Comparison of evolutionary models for regionalization score and heterogeneity (log variance). BM: Brownian motion; OU1: Ornstein-Uhlenbeck with single optimum; Bayou: OU with shifts based on bayou method; Surface: OU with shifts based on surface method. Estimated parameters and confidence intervals (based on simulation) are provided. σ^2: sigma-squared, evolutionary rate; α: alpha, strength of pull toward the optimum; θi: theta, evolutionary optimum for regime i; AICc: corrected Akaike information criterion. Significance of the best-fitting model (surface) relative to the other models was assessed using simulation of likelihood ratios. $\delta_{Surface}$: likelihood ratio of model versus surface; p-val: p-value of this likelihood ratio based on simulation. For likelihood distributions upon which p-values are based, see Figure S8.

	BM	**OU1**	**Bayou**	**Surface**
Regionalization				
$\boldsymbol{\sigma^2}$	0.020 (0.013-0.029)	0.060 (0.024-0.186)	0.091(0.029-4.524)	3.29(2.45-4.22)
$\boldsymbol{\alpha}$		0.060 (0.024-0.222)	0.131(0.045-7.879)	7.07(6.58-9.16)
$\boldsymbol{\theta_1}$		4.62 (4.39-4.86)	3.42(2.87-3.97)	3.01(2.53-3.44)
$\boldsymbol{\theta_2}$			4.69(4.51-4.88)	4.80(4.63-4.97)
$\boldsymbol{\theta_3}$				3.89(3.62-3.89)
AICc	132.02	104.45	99.28	86.29
$\boldsymbol{\delta_{Surface}}$	61.14	28.91	18.64	-
p-val	<0.001	<0.001	<0.001	-
Heterogeneity				
$\boldsymbol{\sigma^2}$	0.027(0.017-0.038)	0.047(0.027-0.097)	0.048(0.027-0.097)	0.062(0.03-0.137)
$\boldsymbol{\alpha}$		0.016(0.008-0.043)	0.022(0.012-0.051)	0.048(0.025-0.129)
$\boldsymbol{\theta_1}$		-1.20(-1.83--0.58)	-4.18(-6.12--2.48)	-3.94(-4.93--2.98)
$\boldsymbol{\theta_2}$			-0.95(-1.39--0.52)	-1.57(-2.06--1.14)
$\boldsymbol{\theta_3}$				-0.48(-0.82--0.13)
AICc	151.99	146.56	142.38	130.31
$\boldsymbol{\delta_{Surface}}$	33.43	25.82	16.59	-
p-val	<0.001	<0.001	0.001	-

REFERENCES CITED

19. K. D. Angielczyk, B. S. Rubidge, Skeletal morphology, phylogenetic relationships and stratigraphic range of *Eosimops newtoni* Broom, 1921, a pylaecephalid dicynodont (Therapsida, Anomodontia) from the Middle Permian of South Africa. *J Syst Palaeontol* **11**, 191-231 (2013).
20. G. M. King, The functional anatomy of a Permian dicynodont. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.*, 243-322 (1981).
21. V. Fernandez *et al.*, Synchrotron reveals Early Triassic odd couple: Injured amphibian and aestivating therapsid share burrow. *PLoS ONE* **8**, (2013).
22. H. D. Sues, F. A. Jenkins, Jr., in *Amniote Paleobiology: Perspectives on the Evolution of Mammals, Birds and Reptiles*, M. T. Carrano, T. J. Gaudin, R. W. Blob, J. R. Wible, Eds. (University of Chicago Press, Chicago, 2006), pp. 114-152.
23. T. S. Kemp, Primitive cynodont *Procynosuchus*: Structure, function and evolution of the postcranial skeleton. *Philos. Trans. R. Soc. Lond., Ser. B: Biol. Sci.* **288**, 217 (1980).
24. R. W. Hook, N. Hotton III, A new sphenacodontid pelycosaur (Synapsida) from the Wichita Group, Lower Permian of north-central Texas. *J Vert Paleontol* **11**, 37-44 (1991).
25. E. D. Cope, Second contribution to the history of the Vertebrata of the Permian formation of Texas. *Proc. Amer. Philos. Soc.* **19**, 38-58 (1880).
26. T. S. Kemp, The skeleton of a baurioid therocephalian therapsid from the Lower Triassic (*Lystrosaurus* Zone) of South Africa. *J Vert Paleontol* **6**, 215-232 (1986).
27. E. C. Case, *A revision of the Cotylosauria of North America*. (Carnegie Institute of Washington, 1911).
28. A. S. Romer, L. I. Price, Review of the Pelycosauria. *Geol Soc Am Spec Pap* **28**, 1-538 (1940).
29. E. Case, S. Williston, in *Permo− Carboniferous Vertebrates from New Mexico,* E. Case, S. Williston, M. Mehl, Eds. (Carnegie Institution of Washington, 1913), vol. 181, pp. 61-70.
30. L. Boonstra, A contribution to the morphology of the Gorgonopsia. *Ann. S. Afr. Mus.* **31**, 137-174 (1934).
31. F. A. Jenkins, Jr., The Chanares (Argentina) Triassic reptile fauna VII. The postcranial skeleton of the transversodontid *Massetognathus pascuali* (Therapsida, Cynodontia). *Breviora* **352**, 1-28 (1970).
32. S. S. Sumida, Reinterpretation of vertebral structure in the Early Permian pelycosaur Varanosaurus acutirostris (Amniota, Synapsida). *J Vert Paleontol* **9**, 451-458 (1989).
33. G. M. King, The aquatic Lystrosaurus: a palaeontological myth. *Hist Biol* **4**, 285-321 (1990).
34. H. F. Osborn, C. H. Sternberg, A mounted skeleton of Naosaurus, a pelycosaur from the Permian of Texas. Bulletin of the AMNH; v. 23, article 14. (1907).
35. W. S. Rasband. (National Institutes of Health, Bethesda, Maryland, USA, 2004).

36. J. J. Head, P. D. Polly, Evolution of the snake body form reveals homoplasy in amniote Hox gene function. *Nature* **520**, 86-89 (2015).
37. S. B. Hedges, J. Dudley, S. Kumar, TimeTree: a public knowledge-base of divergence times among organisms. *Bioinformatics* **22**, 2971-2972 (2006).
38. C. A. Sidor, J. A. Hopson, Ghost lineages and "mammalness": assessing the temporal pattern of character acquisition in the Synapsida. *Paleobiology* **24**, 254-273 (1998).
39. L. J. Revell, phytools: an R package for phylogenetic comparative biology (and other things). *Methods Ecol. Evol.* **3**, 217-223 (2012).
40. J. Pinheiro, D. Bates, S. DebRoy, D. Sarkar, nlme: linear and nonlinear mixed effects models. R package version 3.1-117. (2014).
41. E. Paradis, J. Claude, K. Strimmer, APE: analyses of phylogenetics and evolution in R language. *Bioinformatics* **20**, 289-290 (2004).
42. M. A. Butler, A. A. King, Multivariate comparative analysis using OUCH. *Integr. Comp. Biol.* **49**, E24-E24 (2009).
43. J. F. Uyeda, L. J. Harmon, A novel Bayesian method for inferring and interpreting the dynamics of adaptive landscapes from phylogenetic comparative data. *Syst. Biol.*, (2014).
44. T. Ingram, D. L. Mahler, SURFACE: detecting convergent evolution from comparative data by fitting Ornstein-Uhlenbeck models with stepwise Akaike Information Criterion. *Methods Ecol. Evol.* **4**, 416-425 (2013).
45. C. Boettiger, G. Coop, P. Ralph, Is your phylogeny informative? Measuring the power of comparative methods. *Evolution* **66**, 2240-2251 (2012).
46. N. Cooper, G. H. Thomas, C. Venditti, A. Meade, R. P. Freckleton, A cautionary note on the use of Ornstein Uhlenbeck models in macroevolutionary studies. *Biol. J. Linn. Soc.* **118**, 64-77 (2016).
47. E. A. Buchholtz, C. C. Stepien, Anatomical transformation in mammals: developmental origin of aberrant cervical anatomy in tree sloths. *Evol Dev* **11**, 69-79 (2009).
48. K. E. Jones, L. T. Holbrook, The evolution of lateral accessory articulations in the lumbar region of perissodactyls. *J Vert Paleontol*, e1224892 (2016).
49. D. S. Berman, A. C. Henrici, R. A. Kissel, S. S. Sumida, T. Martens, A new diadectid (Diadectomorpha), Orobates pabsti, from the Early Permian of central Germany. *Bull Carnegie Mus Nat Hist*, 1-36 (2004).
50. J. Müller *et al.*, Homeotic effects, somitogenesis and the evolution of vertebral numbers in recent and fossil amniotes. *Proc Natl Acad Sci U S A* **107**, 2118-2123 (2010).