nature ecology & evolution

SUPPLEMENTARY INFORMATION

https://doi.org/10.1038/s41559-019-0909-z

In the format provided by the authors and unedited.

# Palaeoproteomics resolves sloth relationships

**Samantha Presslee[1,2,3,24], Graham J. Slater[4,24], François Pujos[5], Analía M. Forasiepi[5], Roman Fischer[6], Kelly Molloy[7], Meaghan Mackie[3,8], Jesper V. Olsen[8], Alejandro Kramarz[9], Matías Taglioretti[10], Fernando Scaglia[10], Maximiliano Lezcano[11], José Luis Lanata[11], John Southon[12], Robert Feranec[13], Jonathan Bloch[14], Adam Hajduk[15], Fabiana M. Martin[16], Rodolfo Salas Gismondi[17], Marcelo Reguero[18], Christian de Muizon[19], Alex Greenwood[20,21], Brian T. Chait[7], Kirsty Penkman[22], Matthew Collins[3,23] and Ross D. E. MacPhee[2]***

[1]Department of Archaeology and BioArCh, University of York, Heslington, UK. [2]Department of Mammalogy, American Museum of Natural History, New York, NY, USA. [3]Paleoproteomics Group, Natural History Museum of Denmark and University of Copenhagen, Copenhagen, Denmark. [4]Department of the Geophysical Sciences, University of Chicago, Chicago, IL, USA. [5]Instituto Argentino de Nivología, Glaciología y Ciencias Ambientales, CCT-CONICET-Mendoza, Mendoza, Argentina. [6]Target Discovery Institute, Nuffield Department of Medicine, University of Oxford, Oxford, UK. [7]Chait Laboratory and National Resource for the Mass Spectrometric Analysis of Biological Macromolecules, The Rockefeller University, New York, NY, USA. [8]Novo Nordisk Foundation Center for Protein Research, Copenhagen, Denmark. [9]Sección Paleovertebrados, Museo Argentino de Ciencias Naturales 'Bernardino Rivadavia', Buenos Aires, Argentina. [10]Museo Municipal de Ciencias Naturales 'Lorenzo Scaglia', Mar del Plata, Argentina. [11]Instituto de Investigaciones en Diversidad Cultural y Procesos de Cambio, CONICET and Universidad Nacional de Río Negro, Bariloche, Argentina. [12]Keck-CCAMS Group, Earth System Science Department, University of California, Irvine, Irvine, CA, USA. [13]New York State Museum, Albany, NY, USA. [14]Florida Museum of Natural History, University of Florida, Gainesville, FL, USA. [15]Museo de la Patagonia 'F. P. Moreno', Bariloche, Argentina. [16]Centro de Estudios del Hombre Austral, Instituto de la Patagonia, Universidad de Magallanes, Punta Arenas, Chile. [17]BioGeoCiencias Lab, Facultad de Ciencias y Filosofía/CIDIS, Universidad Peruana Cayetano Heredia, Lima, Peru. [18]CONICET and División Paleontología de Vertebrados, Museo de La Plata. Facultad de Ciencias Naturales, Universidad Nacional de La Plata, La Plata, Argentina. [19]Centre de Recherches sur la Paléobiodiversité et les Paléoenvironnements, Muséum national d'Histoire naturelle, Paris, France. [20]Leibniz Institute for Zoo and Wildlife Research, Berlin, Germany. [21]Department of Veterinary Medicine, Freie Universität Berlin, Berlin, Germany. [22]Department of Chemistry, University of York, Heslington, UK. [23]McDonald Institute for Archaeological Research, University of Cambridge, Cambridge, UK. [24]These authors contributed equally: S. Presslee, G. J. Slater. *e-mail: macphee@amnh.org

# SUPPLEMENTARY INFORMATION

## Table of Contents

## Molecular analyses

### Previous palaeogenomic and palaeoproteomic studies

Over the past 25 years there have been a number of efforts to extract and amplify DNA from sloth fossils[1-4], but most have been limited in scope. Results described in ref. 5 were the first to supply evidence of apparent conflicts between molecular and morphological assessments of folivoran relationships. This latter study was limited to mitochondrial DNA (12S and cytochrome b) and included only two other folivorans in addition to tree sloths. In all but one of their phylogenetic analyses, the tree sloths occupied mutually exclusive tree positions, with *Bradypus variegatus* grouping with the nothrotheriid *Nothrotheriops shastensis* while *Choloepus didactylus* grouped with the mylodontid *Mylodon darwinii*. The *Bradypus*-*Nothrotheriops* dyad was not completely unexpected, inasmuch as morphological evidence had at one time been interpreted to mean that that the three-toed sloth is specifically related to traditional Megatherioidea[6,7]. Rather, it is the pairing of the two-toed sloth with *Mylodon* that seems anomalous in this context: if *Choloepus* is a member of Megalonychidae, as has been frequently suggested[7-9], and if the latter family is phylogenetically part of Megatherioidea, as is also generally maintained[9-11], then both kinds of tree sloths should have clustered with *Nothrotheriops* in these early aDNA runs.

Recently, the mylodontoid affinities of *Choloepus* have been strongly affirmed with molecular evidence[12]. In this study, which included mitochondrial and nuclear exon data for species of both kinds of tree sloths as well as one extinct species (*Mylodon darwinii*), the two-toed sloth was unambiguously placed as the latter's sister taxon. Exon-based divergence estimates, 22 ± 4 Ma for *Choloepus* and 28 ± 4 for *Bradypus*[13] are in general agreement with other evaluations (e.g.,

ref. 14). However, because of limitations in taxon representation no direct test could be made of the supposed relationship between *Bradypus* and the megatherioid *Nothrotheriops* detected by ref. 5.

Clearly, to establish the positions of the extant tree sloths more precisely, sequence information for the many unrepresented lineages of folivores will be required, especially for the pivotal family Megalonychidae. To date, no aDNA study has included *Megalonyx* or any of its allegedly close relatives in a simultaneous phylogenetic analysis, leaving open the possibility that Megalonychidae is, in fact, a paraphyletic assemblage which traditionally includes at least one definite mylodontoid, *Choloepus*.

Only one published report[15] presents protein sequence information on fossil and modern sloths within a systematic context. In that investigation, collagen sequence data was collected for four folivorans (extant *Choloepus hoffmanni*, *Bradypus variegatus*, and the extinct taxa *Lestodon armatus* and *Megatherium americanum*), with the armadillo *Dasypus novemcinctus* serving as outgroup. No fossil megalonychid was included. Although in the phylogenetic analyses bootstrap and posterior probability scores were high, results were markedly discordant with previous studies employing similar taxonomic sampling. In all trees, *Choloepus* and *Bradypus* were retrieved as sisters, joining the branch supporting the mylodontid *Lestodon* with high associated bootstrap values (92-99). This conflicts with the aDNA results previously mentioned as well as recent morphological studies[8, 9, 16] in which the extant tree sloths do not associate together. The anomalous results reported by ref. 15 may be partly explained by small sample size and the limited scope of the proteomic data analysed (COL 1A2 sequences only).

## Utilization of database searching vs. *de novo* sequencing

We used PEAKS (v. 7.5) to compile sequences rather that exclusively rely on a MASCOT database search. PEAKS has the advantage of combining *de novo* sequencing with a database search[17, 18]. This allows for discovery of novel amino acid substitutions that may not occur in the collagen database. The most frequent substitution is serine/alanine (S/A)[19], the only difference between these amino acids being the presence of an hydroxyl (OH) group in serine. Since sequences containing hydroxyproline-alanine or proline-serine will be identical in mass, what this means in practical terms is that PEAKS will not be able to distinguish them. Ancient proteins introduce additional analytical problems. For example, serine may lose the OH group due to protein degradation, in which case the resultant peptide sequence will register as alanine. In order to attempt to overcome such difficulties not only for S/A but for other novel substitutions as well as, we established several criteria:

**1. The number of product ion spectra.** A minimum of 2 spectra had to be identified for either alanine or serine.

**2. The presence of b and y ions**. If product ion spectra were found for both possible substitutions, a higher confidence was given to spectra that identified the selected amino acid using both b and y ions.

**3. The location of hydroxyproline.** The locations of hydroxyproline are generally conserved in collagens, therefore if a HyP-A sequence was identified, the chance of this being genuine was checked against other collagen sequences.

If the existence of a possible substitution was uncertain then it was called as missing sequence data (cf. “x” in Table S4). This reduced the level of total sequence variation, and therefore

possible phylogenetic resolution, but it increased confidence in our concatenated sequences. Currently, instrumental limits on proteomic retrieval mean that some portions of the collagen sequence are not well covered by spectra. Although we have tried to address this by the use of two different enzymes, it must be noted that some amino acid calls made on the basis of only 2 product ion spectra may require modification as more sequence information becomes available for xenarthrans.

**Effect of database limitations on sequence concatenation**

Collagen sequence information has been reported[15] for two extinct species, *Megatherium americanum* and *Lestodon armatus*. In that study, sequences were modelled using a database of mammalian collagen sequences that included only one xenarthran, *Choloepus*, and then only for COL 1A2. By contrast, in the present study COL 1A1 and COL 1A2 sequences were recovered from 13 extinct sloth taxa. These were concatenated using a multi-taxon collagen database[20] that included *Dasypus novemcinctus* (nine-banded armadillo), *Cyclopes didactylus* (silky anteater), and *Mylodon darwinii*, as well as previously-published sequence information for *Lestodon armatus*[15]

As noted in the main text, we sampled the same well-preserved megatheriid specimen (*Megatherium americanum* MAPBAR 3965) utilised by ref. 15, but we achieved a much greater level of coverage (81% vs. 57%). This difference can be attributed, at least in part, to an increase in the number of xenarthran sequences in collagen databases that can be used to call amino acids, as the following example shows. There are 9 amino acid differences between the *Megatherium* sequence developed for this study and the one previously reported[15]---a substantial number, in

view of the highly conserved nature of type 1 collagen[21]. These differences required further analysis because of their potential impact on phylogenetic reconstruction. Columns on the left side of Table S4 present amino acid calls for the 9 sites for which ref. 15 (denoted B) and this study (denoted TS) differ. It is noticeable that most of them concern sequence positions for which *Megatherium* B and *Choloepus* B were given the same amino acid call. As *Choloepus* was the only xenarthran in the database search made by ref. 15, its sequence unavoidably influenced the one derived for *Megatherium* B.

Aligning *Megatherium* B sequence with the larger number of xenarthran sequences used in this study and testing the result in PAUP produced 30 MPTs (as opposed to 13 MPTs when *Megatherium* B sequence was not included). Inspection revealed that the increased number of MPTs was due to additional instances of *Megatherium* B grouping with mylodontoids rather than *Megatherium* TS. The fact that the sequences employed for the two-toed and three-toed sloths in ref. 15 were virtually the same introduced other problems in the analysis (e.g., failure to resolve *Megalonyx* and *Bradypus*). This highlights the importance of verified taxonomic representation in databases used to establish collagen sequences, especially in fossil material.

## Phylogenetic analyses

### Divergence time estimation

Apart from the living tree sloths, the only molecular data currently available for divergence estimation within Folivora come from the same narrow set of taxa employed in aDNA studies[5, 12, 13, 22]. Morphological phylogenetic treatments generally agree that stem folivorans differentiated from Vermilingua during the late Palaeogene[11, 23]. So do some molecular studies[14,

[24], although other investigations making different assumptions have claimed that the separation of the major pilosan taxa may have occurred shortly after the Cretaceous/Palaeogene transition[25]. In any case, fossil evidence for putative folivorans of Palaeogene age is very scanty and provisionally includes taxa that differ markedly from all later sloths, such as late Eocene/early Oligocene *Pseudoglyptodon*[26, 27]. The folivore record does not substantially improve until late in the early Miocene during the Santacrucian SALMA (South American Land Mammal Age), 18–16 Ma[11]. Following a major decline thereafter, diversity increased in the late Miocene but was sharply truncated by a major drop at the end of the Huayquerian (late Miocene), as has recently been modelled[10].

**Backbone constraint**

Following procedures outlined in ref. 14, we selected 34 fossil xenarthrans for which relatively well-constrained stratigraphic dates were available. Most of these can be phylogenetically constrained on the basis of formal phylogenetic analyses based on morphological data. For cases where precise dating was unavailable, we used the rank of the South American Land Mammal Age (SALMA) from which the taxon is known. Compared to the topological constraint used previously[14], our backbone (Fig. S3) is less resolved because many of the clades recovered in our initial Parsimony and Bayesian topology searches for the present paper indicated a lack of support for the monophyly of traditionally recognized folivoran clades. Nonetheless, we opted to include all of the folivoran taxa previously included to maximize fossil sampling and improve estimation of the parameters of the fossilized birth-death process[28-30]. For this study we included

eight additional fossil taxa in order to improve divergence time estimates both within and outside of Folivora, as follows:

**Vermilingua:**

**Taxon**: *Protamandua rothi.* **Age:** 17.5 - 11.608 Ma. **Stratigraphic Context:** Santa Cruz Fm, Patagonia.

**Taxon:** *Palaeomyrmidon incomptus* **Age:** 6.8 - 3.0 Ma. **Stratigraphic Context:** Araucano Fm, Argentina

**Taxon:** *Neotamandua conspicua* **Age:** 6.8 - 3.0 Ma. **Stratigraphic Context:** La Venta Fm, Colombia; Araucano Fm, Argentina

**Comment:** See ref. 34 for further information. Because these taxa are constrained to form a monophyletic clade with our single sampled vermilinguan, *Cyclopes*, coarse stratigraphic ranges for these taxa were taken from The Paleobiology Database (https://paleobiodb.org/).

**Folivora:**

**Taxon**: *Paramylodon sp.* **Age**: 4.7 - 3.6 Ma. **Stratigraphic Context**: Early Blancan NALMA of Guanajuato, Mexico. **Comment:** There is some confusion regarding early North American occurrences of *Paramylodon* due to uncertainties regarding the taxonomy of this genus and the South American taxon *Glossotherium*[31]. North American records of this mylodontid lineage are

first recorded in Guanajuato, Central Mexico by the early Blancan[32-33], providing a minimum age for the genus.

**Cingulata**

**Taxon:** *Propalaeohoplophorus australis.* **Age:** 17.5 - 16.3 Ma. **Stratigraphic Context:** Santa Cruz Fm, Argentina. **Comment:** *Propalaeohoplophorus* is one of the oldest and best known glyptodonts[35]. Due to uncertainty regarding its relationship to Neogene taxa, however, we did not constrain it to fall outside of (*Glyptodon* + *Doedicurus*), our sampled glyptodonts, but rather placed the three taxa in an unresolved trichotomy.

**Taxon:** *Stegotherium tasselatum* **Age:** 17.5 - 16.3 Ma. **Stratigraphic Context:** Santa Cruz Fm, Argentina. **Comment:** *Stegotherium* is robustly resolved as the sister taxon to extant *Dasypus* in ref. 35.

**Taxon:** *Kuntinaru boliviensis.* **Age:** 30 - 23 Ma. **Stratigraphic Context:** Deseaden SALMA (late Oligocene) of Salla, Bolivia. **Comment:** *Kuntinaru* was recovered as a tolypeutine dasypodid in ref. 36. Its chronostratigraphic position thus provides a minimum age for the divergence of *Dasyus* and the two glyptodontines sampled for this study (see above).

**Taxon:** *Riostegotherium yanei*. **Age:** 59.0 - 57.5 Ma. **Stratigraphic Context:** Itaboraian (late Paleocene) of Brazil. **Comments:** *Riostegotherium yanei* is claimed to be the earliest known cingulate[37, 38], and we treat it as such here to calibrate the age of the crown xenarthran ancestor.

## Supplementary Information References

1 Höss, M., Dilling, A., Currant, A. & Pääbo, S. Molecular phylogeny of the extinct ground sloth *Mylodon darwinii*. *Proc. Nat. Acad. Sci. USA* **93** (1), 181–185 (1996). PMID: 8552600

2 Poinar, H.N. *et al*. Molecular coproscopy: Dung and diet of the extinct ground sloth *Nothrotheriops shastensis*. *Science* **281** (5375), 402–406 (1998). PMID: 9665881

3 Hofreiter, M., Betancourt, J. L. , Sbriller, A. P., Markgraf, V., McDonald, H. G. Phylogeny, diet, and habitat of an extinct ground sloth from Cuchillo Cura, Neuquen Province, southwest Argentina. *Quat. Res.* **59** (3), 364–378 (2003). doi: 10.1016/S0033-5894(03)00030-9

4 Clack, A. A., MacPhee, R. D. E. & Poinar, H. N. Case study: ancient sloth DNA recovered from hairs preserved in paleofeces. *Methods Mol. Biol.* **840**, 51–6 (2012). doi: 10.1007/978-1-61779-516-9_7

5 Greenwood, A. D., Castresana, J., Feldmaier-Fuchs, G. & Pääbo, S. A molecular phylogeny of two extinct sloths. *Mol. Phylogen. Evol.* **18** (1), 94–103 (2001). PMID: 11161746

6 Patterson, B. & Pascual, R. Evolution of mammals on southern continents. *Q. Rev. Biol.* **43** (4), 409–451 (1968). https://www.jstor.org/stable/2819014

7 Webb, S. D. in *The Evolution and Ecology of Armadillos, Sloths, and Vermilinguas* (ed Montgomery, G. G.) 105-112 (Smithsonian Institution Press, 1985).

8 White, J. & MacPhee, R. D. E. in *Biogeography of the West Indies: Patterns and Perspectives*, 2nd ed (eds Woods, C. A. & Sergile, F. E.) 201-236 (CRC Press: Boca Raton LA, 2001).

9 Gaudin, T. J. Phylogenetic relationships among sloths (Mammalia, Xenarthra, Tardigrada): the craniodental evidence. *Zool. J. Linn. Soc.* **140**, 255-305 (2004). doi: 10.1111/j.1096-3642.2003.00100.x

10 Varela, L., Tambusso, P. S., McDonald, H. G. & Fariña, R. A. Phylogeny, macroevolutionary trends and historical biogeography of sloths: Insights from a Bayesian morphological clock analysis. *Syst. Biol.* (2018). doi: 10.1093/sysbio/syy058

11 Gaudin T. J. & McDonald, H. G. in *The Biology of Xenarthra* (eds Vizcaíno, S. F. & Loughry, W. J.) 24-36 (University Press of Florida, Gainesville, 2008).

12 Delsuc, F., Kuch, M., Gibb, G. C., Hughes, J., Szpak, P., Southon, J., Enk, J., Duggan, A. T. & Poinar, H. N. Resolving the phylogenetic position of Darwin's extinct ground sloth (*Mylodon darwinii*) using mitogenomic and nuclear exon data. *Proc. Roy. Soc. B* **285**, 20180214 (2018). doi: 10.1098/rspb.2018.0214

13 Delsuc, F. & Douzery, E. J. P. in *The Biology of Xenarthra* (eds Vizcaino, S. F. & Loughry, W. J.) 11-23 (University Press of Florida, Gainesville, 2008).

14 Slater, G., Cui, P., Forasiepi, A.M., Lenz, D., Tsangaras, K., Voirin, B., de Moraes, N., MacPhee R.D.E. & Greenwood, A.D. Evolutionary relationships among extinct and extant sloths: The evidence of mitogenomes and retroviruses. *Genome Biol. Evol.* **8** (3), 607–621 (2016). doi: 10.1093/gbe/evw023

15 Buckley, M., Fariña, R. A., Lawless, C., Tambusso, P. S., Varela, L., Carlini, A. A., Powell, J. E. & Martinez, J. G. Collagen sequence analysis of the extinct giant ground sloths *Lestodon* and *Megatherium*. *PloS One* **10** (12), e0144793 (2015). doi: 10.1371/journal.pone.0144793

16 Pujos, F. *et al.* A peculiar climbing Megalonychidae from the Pleistocene of Peru and its implications for sloth history. *Zoo. J. Linn. Soc.* 149 (2), 179–235 (2007). doi: 10.1111/j.1096-3642.2007.00240.x

17 Ma, B., Zhang, K., Hendrie, C., Liang, C., Li, M., Doherty-Kirby, A. & Lajoie, G. A. PEAKS: powerful software for peptide de novo sequencing by tandem mass spectrometry. *Rapid Comm. Mass Spectrometry* **17** (20), 2337–2342 (2003).

18 Zhang, J., Xin, L., Shan, B., Chen, W., Xie, M., Yuen, D., Zhang, W., Zhang, Z., Lajoie, G.A. & Ma, B. PEAKS DB: de novo sequencing assisted database search for sensitive and accurate peptide identification. *Mol. Cell. Proteomics* **11** (4), M111.010587 (2012).

19 Hendy, J., Welker, F., Demarchi, B., Speller, C., Warinner, C. & Collins, M.J. A guide to ancient protein studies. *Nature Ecology & Evolution* **2**, 791–799 (2018). doi: 10.1038/s41559-018-0510-x

20 Welker, F. *et al.* Ancient proteins resolve the evolutionary history of Darwin's South-American ungulates. *Nature* **522**, 81-84 (2015). doi: 10.1038/nature14249.

21 Welker, F. Elucidation of cross-species proteomic effects in human and hominin bone proteome identification through a bioinformatics experiment. *BMC Evol. Biol.* **18** (1), 23 (2018).

22 Moraes-Barros, N., Silva, J. A. & Morgante, J. S. Morphology, molecular phylogeny, and taxonomic inconsistencies in the study of *Bradypus* sloths (Pilosa: Bradypodidae). *J. Mammal.* **92**, 86–100 (2011). doi: 10.1644/10-MAMM-A-086.1

23 O'Leary, M. A. *et al.* The placental mammal ancestor and the post–K-Pg radiation of placentals. *Science* **339,** 662–667 (2013). doi: 10.1126/science.1229237

24 Murphy, W. J., Pringle, T. H., Crider, T. A., Springer, M. S. & Miller, W. Using genomic data to unravel the root of the placental mammal phylogeny. *Genome Res.***17**, 413–421 (2007). doi: 10.1101/gr.5918807

25 Meredith, R.W. *et al.* Impacts of the Cretaceous terrestrial revolution and KPg extinction on mammal diversification. *Science* **334**, 521–524 (2011). doi: 10.1126/science.1211028.

280

281 26 McKenna, M.C., Wyss, A. R. & Flynn, J.J. Paleogene pseudoglyptodont xenarthrans from
282 central Chile and Argentine Patagonia. *Amer. Mus. Novitates* **3536**, 1-18 (2006).
283 http://hdl.handle.net/2246/5814

284

285 27 Gaudin, T. J. & Croft, D. Paleogene Xenarthra and the evolution of South American
286 mammals. *J. Mamm.* 96 (4) 622–634 (2015). doi: 10.1093/jmammal/gyv073.

287

288 28 Heath, T. A, Huelsenbeck, J. P. & Stadler, T. The fossilized birth-death process for coherent
289 calibration of divergence-time estimates. *Proc. Nat. Acad. Sci. USA.* **111 (**29), E2957--E2966
290 (2014). doi: 10.1073/pnas.1319091111

291

292 29 Gavryushkina, A., Welch, D., Stadler, T. & Drummond, A. J. Bayesian inference of sampled
293 ancestor trees for epidemiology and fossil calibration. *PLoS Comput. Biol.* **10** (12), e1003919
294 (2014). doi: 10.1371/journal.pcbi.1003919

295

296 30 Gavryushkina, A., Heath, T. A., Ksepka, D. T., Stadler, T., Welch, D. & Drummond, A. J.
297 Bayesian total-evidence dating reveals the recent crown radiation of penguins. *Syst. Biol.* **66** (1),
298 57-73 (2017). doi: 10.1093/sysbio/syw060

299

300 31 McAfee, R. K. Feeding mechanics and dietary implications in the fossil sloth *Neocnus*
301 (Mammalia: Xenarthra: Megalonychidae) from Haiti. *J. Morphol.* **272**, 1204-1216 (2011). doi:
302 10.1002/jmor.10976.

32 Miller, W. E. & Carranza-Castaneda, O. Late Tertiary canids from central Mexico. *J. Paleontol.* **72** (3), 546-556 (1998).

33 Flynn, J.J., Kowallis, B. J., Nuñez, C., Carranza-Castañeda, O., Miller, W. E., Swisher, C. C. & Lindsay, E. Geochronology of Hemphillian-Blancan aged strata, Guanajuato, Mexico, and implications for timing of the Great American Biotic Interchange. *J. Geol.* **113** (3), 287-307 2005.

34 Gaudin, T. J. & Branham, D. G. The phylogeny of the Myrmecophagidae (Mammalia, Xenarthra, Vermilingua) and the relationship of *Eurotamandua* to the Vermilingua. *J. Mamm. Evol.* **5** (3), 237-265 (1998).

35 Gaudin, T. J. & Wible, J. R. The phylogeny of living and extinct armadillos (Mammalia, Xenarthra, Cingulata): a craniodental analysis. In *Amniote Paleobiology: Perspectives on the Evolution of Mammals, Birds, and Reptiles* (eds Carrano, M. T., Gaudin, T. J., Blob, R. W. & Wible, J. R.) 153-198 (Univ. Chicago Press, 2006).

36 Billet, G., Hautier, L., de Muizon, C. & Valentin, X. Oldest cingulate skulls provide congruence between morphological and molecular scenarios of armadillo evolution. *Proc. Roy. Soc. B* **278**, 2791-2797 (2011) doi:10.1098/rspb.2010.2443

37 Oliveira, É. V. & Bergqvist, L. P. A new Paleocene armadillo (Mammalia, Dasypodoidea) from the Itaboraí Basin, Brazil. *Publicación Electrónica de la Asociación Paleontológica Argentina* **5** (1) (1998).

38 Bergqvist, L. P., Abrantes, E. A. L. & Avilla, L.S. The Xenarthra (Mammalia) of São José de Itaboraí Basin (Upper Paleocene, Itaboraian), Río de Janeiro, Brazil. *Geodiversitas* **26** (2), 323-337 (2004).

## LEGENDS FOR SUPPLEMENTARY FIGURES AND TABLES

**Figure S1**: Strict consensus of 192 most parsimonious trees (Length = 214, CI = 0.706, RI = 0.707). Values above nodes represent bootstrap support derived from 10,000 bootstrap replicates.

**Figure S2:** 50% majority rule consensus tree from Bayesian analysis of collagen sequences plus published mitochondrial genomes. Values below nodes are posterior probabilities for the descendant clade. Note that *Bradypus* is rendered paraphyletic with respect to *Megalonyx jeffersoni* but this is likely due to a lack of overlapping data (*Megalonyx* is represented by proteomic data only, while all *Bradypus* species except *B. variegatus* are represented by genomic data only).

**Figure S3:** Backbone topology constraints employed for BEAST analyses under the fossilized birth death process. Taxa in large font are those successfully sequenced for collagen; those in smaller font lack such data but are included to assist in divergence time estimation. Uncertainty regarding the higher-level relationships of Oligo-Miocene taxa that emerge from our topology analyses mean that few constraints can be used in comparison with earlier studies[14]. For further information, see SI, Backbone constraint).

353 **Figure S4:** Time scaled maximum clade credibility tree from BEAST analysis of collagen
354 sequences of 16 extant and extinct folivorans plus four non-folivoran outgroups. As in Fig. 4 in
355 the main text, branch lengths are the mean values from the retained posterior sample, while blue
356 bars represent 95% highest posterior density intervals. Values at nodes are posterior
357 probabilities. Vertical shaded bars correspond to South American land mammal ages (SALMAs),
358 two of which are emphasized: Deseadan (**), 29–21 Ma, and Santacrucian (*), 17.5–16.3 Ma.

359

360 **Figure S5:** Time scaled maximum clade credibility tree from BEAST analysis of 16 extant and
361 extinct folivoran collagen sequences; non-folivoran xenarthran sequences used in Fig. S4 are
362 excluded, and position of root determined by molecular clock. Other conventions as in Fig. S4.

363

364 **Figure S6**: 50% majority rule consensus tree derived from Bayesian analysis of combined
365 proteomic, genomic, and phenomic data. Taxa sampled for molecular data are in bold font. Node
366 labels are posterior probabilities. *Bradypus* and *Choloepus* are recovered in positions similar to
367 ones found in molecular-only analyses, with strong support. By contrast, Antillean sloths and
368 *Megalonyx*, for which no genomic data are available, are recovered here as a monophyletic clade
369 within Megatherioidea.

370

371

**Table S1:** Information on all samples investigated for this study. Blank entries under Age indicate no data.

**Table S2:** Radiocarbon dates for specimens successfully screened for MS/MS (see Table S1)[a,b].

**Table S3:** Marginal Likelihoods estimated for three clock models for proteomic data alone and proteomic + genomic data using the path sampling algorithm in BEAST 2.5.1. Marginal likelihoods were estimated based on 8 steps of 1 million generations, the first 50% discarded as burn-in and an alpha of 0.3. Also given are pairwise model comparisons, where positive numbers should be read as support for the row model over the column model. UCLD: Uncorrelated Lognormal Distribution; RLC: Random Local Clock; Strict: Strict Clock.

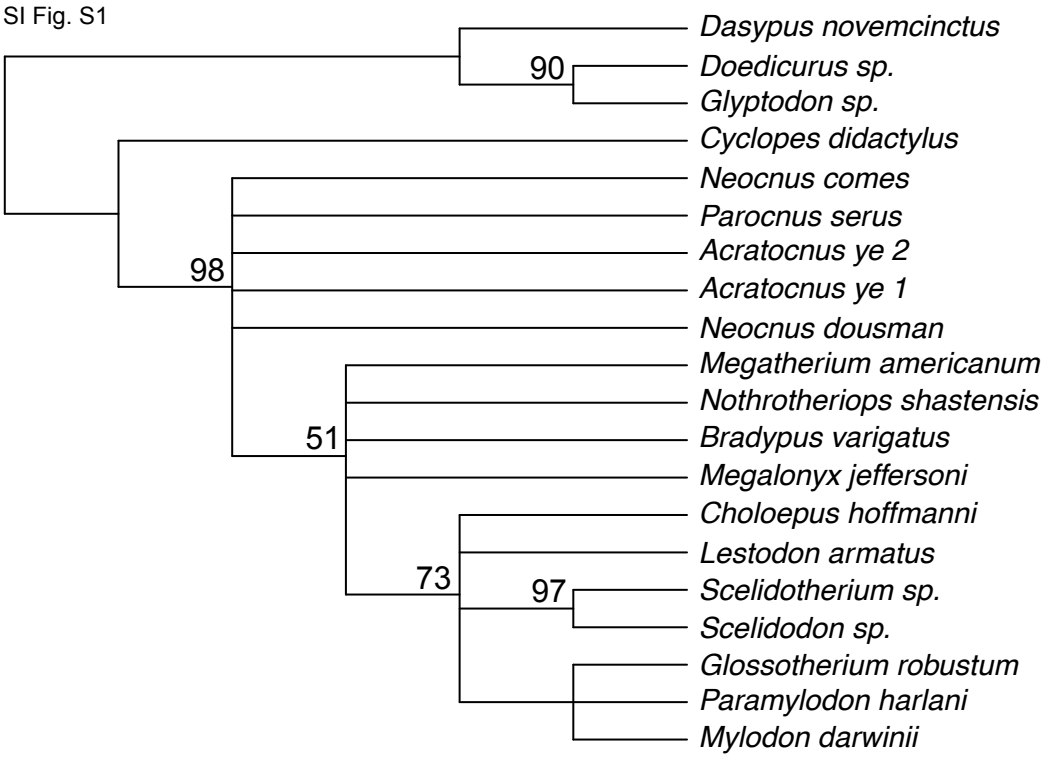
**Table S4:** Detected amino acid differences between *Megatherium* sequence reported in ref. 15 (denoted by B) and the sequence concatenated in this study (TS). All South American sloth taxa analysed in the present study have been included for comparison, as has the data for *Choloepus* B from the same publication[15]. Taxa are color-coded: *Megatherioidea (orange); *Mylodontoidea (blue); Cingulata (green). Empty (x) cells indicate missing sequence information. Red letters represent amino acid (AA) matches among taxa for indicated positions. Absence of any matches between *Megatherium* B and *Megatherium* TS is probably due to low sequence coverage and other limitations of the earlier study[15]. For additional interpretation, see SI, Effect of database limitations on sequence concatenation.

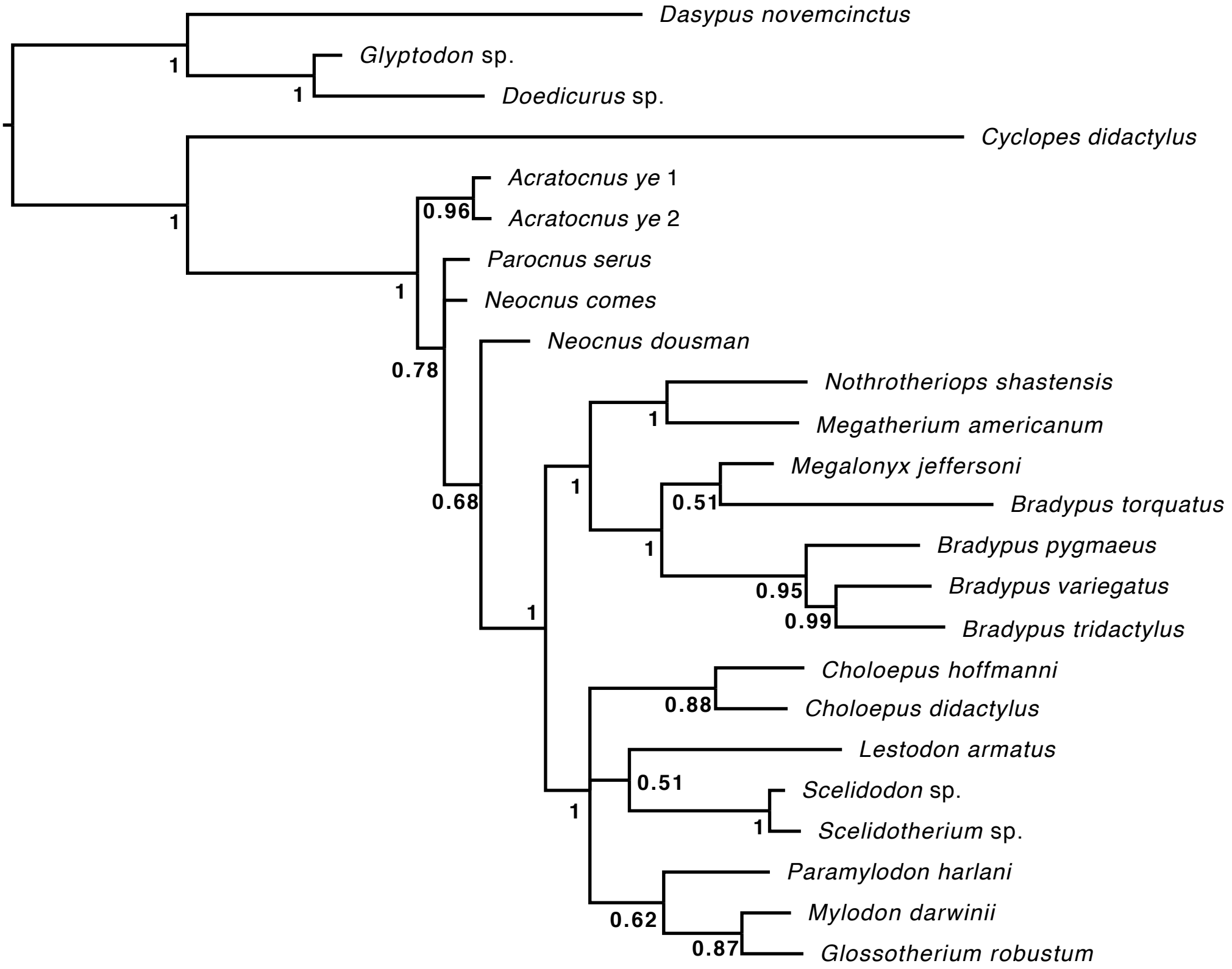
393

394 **Table S5:** Accession numbers of collagen sequences used in this study (available on the Uniprot

395 website, https://www.uniprot.org/)

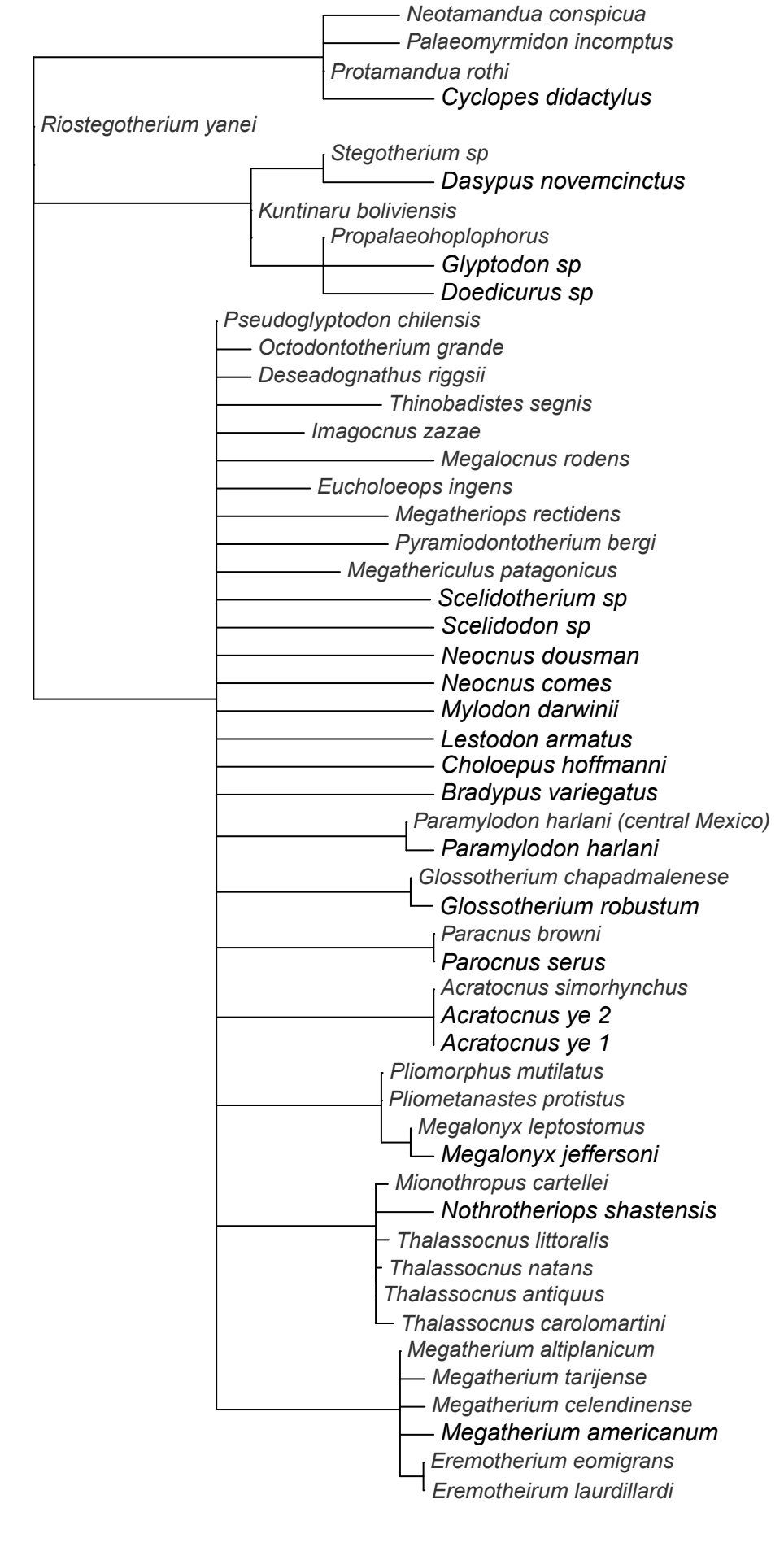
SI Fig. S1


Dasypus novemcinctus
90
Doedicurus sp.
Glyptodon sp.
Cyclopes didactylus
Neocnus comes
Parocnus serus
Acratocnus ye 2
98
Acratocnus ye 1
Neocnus dousman
Megatherium americanum
Nothrotheriops shastensis
51
Bradypus varigatus
Megalonyx jeffersoni
Choloepus hoffmanni
Lestodon armatus
73
97
Scelidotherium sp.
Scelidodon sp.
Glossotherium robustum
Paramylodon harlani
Mylodon darwinii

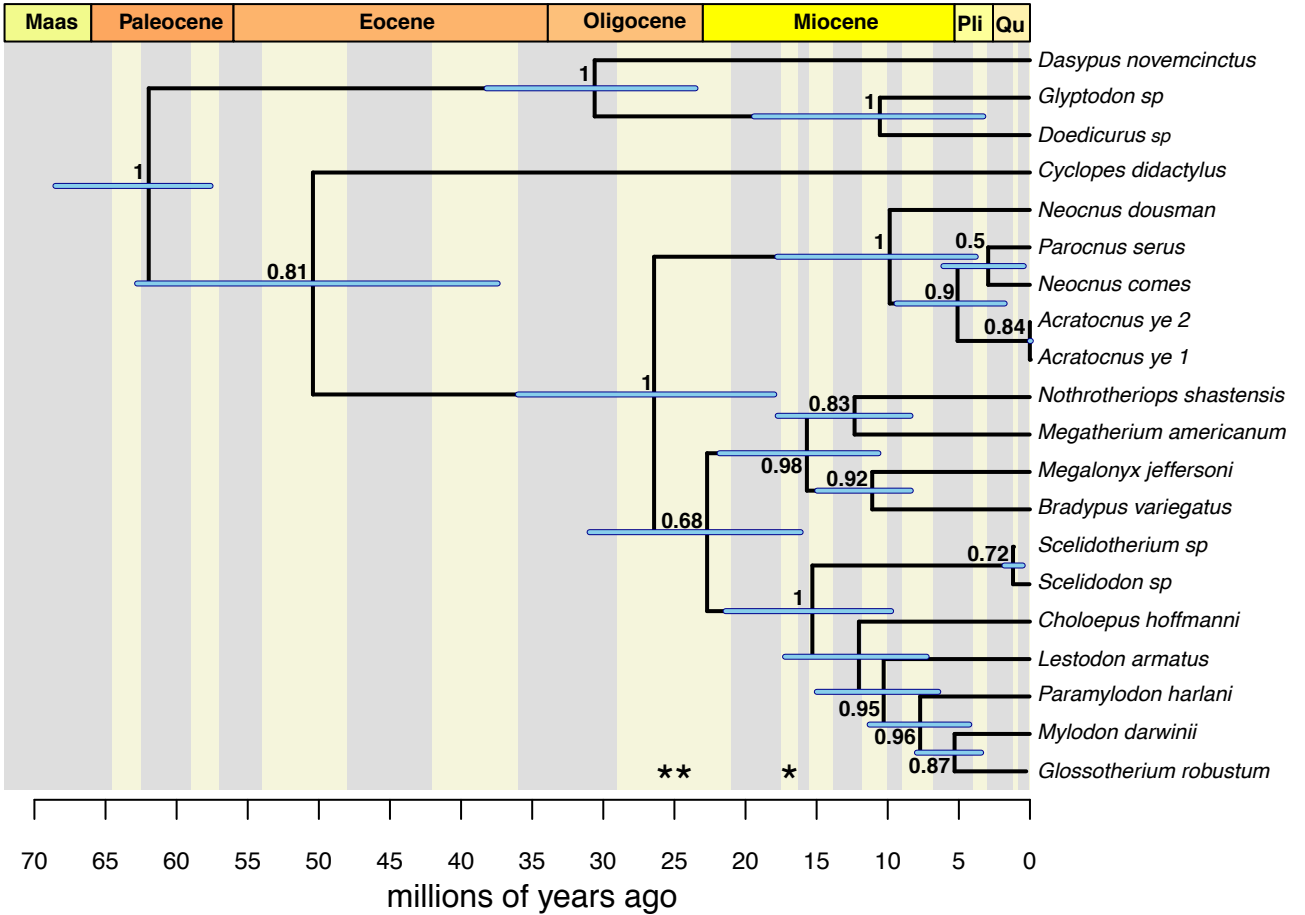

Dasypus novemcinctus
Glyptodon sp.
Doedicurus sp.
Cyclopes didactylus
Acratocnus ye 1
Acratocnus ye 2
Parocnus serus
Neocnus comes
Neocnus dousman
Nothrotheriops shastensis
Megatherium americanum
Megalonyx jeffersoni
Bradypus torquatus
Bradypus pygmaeus
Bradypus variegatus
Bradypus tridactylus
Choloepus hoffmanni
Choloepus didactylus
Lestodon armatus
Scelidodon sp.
Scelidotherium sp.
Paramylodon harlani
Mylodon darwinii
Glossotherium robustum
1
1
0.96
1
1
0.78
0.68
1
1
0.51
1
0.95
0.99
1
0.88
0.51
1
1
0.62
0.87
0.3

SI Fig. S2

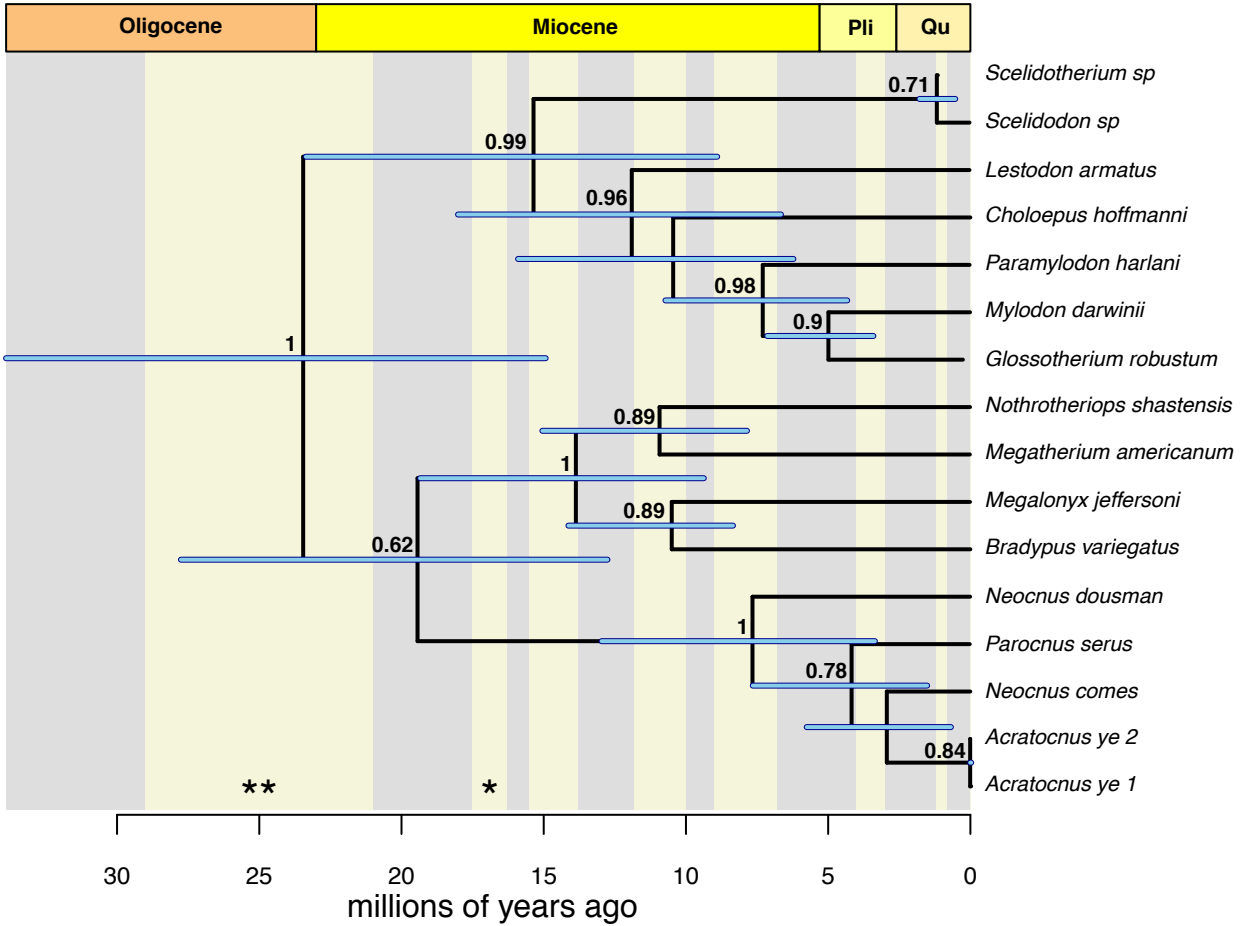
SI Fig. S3


Neotamandua conspicua
Palaeomyrmidon incomptus
Protamandua rothi
Cyclopes didactylus
Riostegotherium yanei
Stegotherium sp
Dasypus novemcinctus
Kuntinaru boliviensis
Propalaeohoplophorus
Glyptodon sp
Doedicurus sp
Pseudoglyptodon chilensis
Octodontotherium grande
Deseadognathus riggsii
Thinobadistes segnis
Imagocnus zazae
Megalocnus rodens
Eucholoeops ingens
Megatheriops rectidens
Pyramiodontotherium bergi
Megathericulus patagonicus
Scelidotherium sp
Scelidodon sp
Neocnus dousman
Neocnus comes
Mylodon darwinii
Lestodon armatus
Choloepus hoffmanni
Bradypus variegatus
Paramylodon harlani (central Mexico)
Paramylodon harlani
Glossotherium chapadmalenese
Glossotherium robustum
Paracnus browni
Parocnus serus
Acratocnus simorhynchus
Acratocnus ye 2
Acratocnus ye 1
Pliomorphus mutilatus
Pliometanastes protistus
Megalonyx leptostomus
Megalonyx jeffersoni
Mionothropus cartellei
Nothrotheriops shastensis
Thalassocnus littoralis
Thalassocnus natans
Thalassocnus antiquus
Thalassocnus carolomartini
Megatherium altiplanicum
Megatherium tarijense
Megatherium celendinense
Megatherium americanum
Eremotherium eomigrans
Eremotheirum laurdillardi
60
50
40
30
20
10
0
millions of years ago

SI Fig. S4


Maas
Paleocene
Eocene
Oligocene
Miocene
Pli
Qu
Dasypus novemcinctus
Glyptodon sp
Doedicurus sp
Cyclopes didactylus
Neocnus dousman
Parocnus serus
Neocnus comes
Acratocnus ye 2
Acratocnus ye 1
Nothrotheriops shastensis
Megatherium americanum
Megalonyx jeffersoni
Bradypus variegatus
Scelidotherium sp
Scelidodon sp
Choloepus hoffmanni
Lestodon armatus
Paramylodon harlani
Mylodon darwinii
Glossotherium robustum
1
1
1
1
0.5
0.9
0.84
0.81
1
0.83
0.98
0.92
0.68
0.72
1
0.95
0.96
0.87
**
*
70
65
60
55
50
45
40
35
30
25
20
15
10
5
0
millions of years ago

SI Fig. S5


Oligocene
Miocene
Pli
Qu
Scelidotherium sp
Scelidodon sp
Lestodon armatus
Choloepus hoffmanni
Paramylodon harlani
Mylodon darwinii
Glossotherium robustum
Nothrotheriops shastensis
Megatherium americanum
Megalonyx jeffersoni
Bradypus variegatus
Neocnus dousman
Parocnus serus
Neocnus comes
Acratocnus ye 2
Acratocnus ye 1
0.71
0.99
0.96
0.98
0.9
1
0.89
1
0.89
0.62
1
0.78
0.84
**
*
30
25
20
15
10
5
0
millions of years ago


Dasypus_novemcinctus
Glyptodon
Doedicurus_sp
Holmesina
TOLYPEUTES
Prozaedyus
EUPHRACTUS
Pseudoglyptodon
Palaeomyrmidon
Protamandua
Cyclopes_didactylus
TAMANDUA
MYRMECOPHAGA
Choloepus_hoffmanni
Choloepus_didactylus
Brievabradys
Octodontotherium
Paroctodontotherium
Octomylodon
Urumacotherium
Octodontobradys
Baraguatherium
Mylodon_darwinii
Paramylodon_harlani
Glossotherium_robustum
Simomylodon
Pleurolestodon
Thinobadistes
Lestodon_armatus
Lestobradys
Bolivartherium
Pseudoprepotherium
Nematherium
Analcitherium
Neonematherium
Sibyllotherium
Proscelidodon
Catonyx
Scelidodon
Scelidotherium
Valgipes
Bradypus_torquatus
Bradypus_pygmaeus
Bradypus_variegatus
Bradypus_tridactylus
Schismotherium
Pelecyodon
Hapalops
Hyperleptus
Eucholoeops
Pseudortotherium
Megalonychotherium
Pliometanastes
Megalonyx_jeffersoni
Australonyx
Ahytherium
Megistonyx
Pliomorphus
Parocnus_serus
Megalocnus
Neocnus_dousman
Neocnus_comes
Acratocnus_ye_1
Acratocnus_ye_2
Analcimorphus
Thalassocnus
Lakukullus
Prepoplanops
Planops
Prepotherium
Nothrotheriops_shastensis
Nothrotherium
Mionothropus
Pronothrotherium
Xyophorus
Aymaratherium
Diabolotherium
Megathericulus
Anisodontherium
Pyramiodontherium
Proeremotherium
Megatherium_americanum
Eremotherium
0.4