# Palaeoproteomics resolves sloth relationships

Samantha Presslee[1,2,3,¶], Graham J. Slater[4,¶], 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[1,3], Ross D. E. MacPhee[2,*]

**Authors' Affiliations:**

[1]Department of Archaeology and BioArCh, University of York, Wentworth Way, Heslington, York YO10 5NG, UK

[2]Department of Mammalogy, American Museum of Natural History, New York, NY 10024 USA

[3]Paleoproteomics Group, Natural History Museum of Denmark and University of Copenhagen 1353 Copenhagen, Denmark

[4]Department of the Geophysical Sciences, University of Chicago, 5734 S. Ellis Avenue, Chicago, IL 60637 USA

[5]Instituto Argentino de Nivología, Glaciología y Ciencias Ambientales, CCT-CONICET-Mendoza, Parque Gral San Martin, 5500 Mendoza, Argentina

[6]Target Discovery Institute, Nuffield Department of Medicine, University of Oxford, Roosevelt Drive, Oxford OX3 7FZ, UK

[7]Chait Laboratory and National Resource for the Mass Spectrometric Analysis of Biological Macromolecules, The Rockefeller University, 1230 York Avenue, New York, NY 10065 USA

[8]Novo Nordisk Foundation Center for Protein Research, Blegdamsvej 3B, 2200, Copenhagen, Denmark

[9]Sección Paleovertebrados, Museo Argentino de Ciencias Naturales "Bernardino Rivadavia", Av. Angel Gallardo 470, C1405DJR, Buenos Aires, Argentina.

[10]Museo Municipal de Ciencias Naturales "Lorenzo Scaglia", Plaza España s/n, 7600 Mar del Plata, Argentina

[11]Instituto de Investigaciones en Diversidad Cultural y Procesos de Cambio, CONICET and Universidad Nacional de Río Negro, 8400 Bariloche, Argentina

[12]Keck-CCAMS Group, Earth System Science Department, B321 Croul Hall, University of California, Irvine, Irvine, CA 92697-3100, USA

[13]New York State Museum, Albany, NY 12230 USA

[14]Florida Museum of Natural History, University of Florida, Gainesville, FL 32611-7800 USA

[15]Museo de la Patagonia "F. P. Moreno", Bariloche 8400, 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, Perú

[18]CONICET and División Paleontología de Vertebrados, Museo de La Plata. Facultad de Ciencias Naturales, Universidad Nacional de La Plata, Paseo del Bosque s/n, B1900FWA, La Plata, Argentina

[19]Centre de Recherches sur la Paléobiodiversité et les Paléoenvironnements (CR2P), UMR CNRS 7207, Muséum national d'Histoire naturelle, 8, rue Buffon 75231 Paris, France

[20]Leibniz Institute for Zoo and Wildlife Research, Alfred-Kowalke-Straße 17, 10315 Berlin, Germany

[21]Department of Veterinary Medicine, Freie Universität Berlin, 14163 Berlin, Germany

[22]Department of Chemistry, University of York, Wentworth Way, Heslington, York YO10 5DD, UK

¶These authors contributed equally to this work.

*Author for correspondence: macphee@amnh.org

**Abstract**

The living tree sloths *Choloepus* and *Bradypus* are the only remaining members of Folivora, a major xenarthran radiation that occupied a wide range of habitats in many parts of the western hemisphere during the Cenozoic, including both continents and the West Indies. Ancient DNA evidence has played only a minor role in folivoran systematics, as most sloths lived in places not conducive to genomic preservation. Here we utilise collagen sequence information, both separately and in combination with published mtDNA evidence, to assess the relationships of tree sloths and their extinct relatives. Results from phylogenetic analysis of these datasets differ substantially from morphology-based concepts: *Choloepus* groups with Mylodontidae, not Megalonychidae; *Bradypus* and *Megalonyx* pair together as megatherioids, while monophyletic Antillean sloths may be sister to all other folivorans. Divergence estimates are consistent with fossil evidence for mid-Cenozoic presence of sloths in the West Indies and an early Miocene radiation in South America.

The sloths (Xenarthra, Folivora), nowadays a taxonomically narrow (6 species in 2 genera) component of the fauna of South and Central America[1,2] were once a highly successful clade of placental mammals as measured by higher-level diversity (Fig. 1). Diverging sometime in the Palaeogene from their closest relatives, the anteaters (Vermilingua), folivorans greatly expanded their diversity and range, eventually reaching North America as well as the West Indies[3-8]. During the late Cenozoic sloth lineage diversity may have expanded and contracted several times[9]. Final collapse occurred in the late Quaternary (end-Pleistocene on the continents, mid-

78 Holocene in the West Indies), leaving only the lineages that culminated in the extant two-toed
79 (*Choloepus*) and three-toed (*Bradypus*) tree sloths.

80 Radically differing from other sloth taxa in their manifold adaptations for "inverted" suspensory
81 locomotion, tree sloths have an obscure evolutionary history[10]. Despite their overall similarity in
82 body plans, tree sloths probably acquired their remarkable locomotor adaptations separately, one
83 of many indications that the course of folivoran evolution has been marked by detailed
84 convergences among evolutionarily distinct clades[11-19]. The current consensus [8-10,16, 17] in
85 morphology-based phylogenetic treatments is to place the three-toed sloth as sister to all other
86 folivorans (Fig. 1, "eutardigrades"), while *Choloepus* is typically nested within the otherwise
87 extinct family Megalonychidae, either proximate to or actually within the group that radiated in
88 the West Indies[3, 7, 11, 13, 16, 21, 22]. Although this arrangement recognizes the existence of
89 convergence in the origins of arboreality in tree sloths, it has proven difficult to effectively test.
90 Sloth palaeontology is an active field of inquiry (e.g., refs 10, 17, 22-31), but the placement of a
91 number of early Neogene clades is uncertain or disputed[32] (e.g., "unallocated basal
92 megatherioids" in Fig. 1), and the nature of their relationships with the tree sloths is accordingly
93 indeterminate. This has an obvious impact on our ability to make macroevolutionary inferences[14]
94 (e.g., ancestral modes of locomotion) for tree sloth species, which have no known pre-
95 Quaternary fossil record[10].

96 Genomic evidence, now routinely used in mammalian systematic research and phylogenetic
97 reconstruction, has so far been of limited use in evaluating these issues. Mitochondrial and at
98 least some nuclear sequence data are available for most well-defined species of living tree sloths,
99 but published ancient DNA (aDNA) evidence exists for only two late Pleistocene species [33-36].
100 Lack of aDNA evidence is not surprising, given that the vast majority of sloth species lived in

101 temperate or tropical environments not conducive to aDNA preservation. Yet despite these
102 limitations, aDNA analyses have tentatively pointed to a set of relationships between extant
103 sloths and their extinct relatives that are very different from those implied by morphological
104 data: the three-toed sloth is consistently recovered in association with the North American
105 megatherioid *Nothrotheriops shastensis*[34,38,39], a position reflected in some older classifications[13,
106 20, 21] while the two-toed sloth is firmly established as sister to the South American mylodontoid
107 *Mylodon darwinii*[34-40.] This, however, is not enough information to rigorously test, with
108 molecular evidence, cladistic relationships established solely on morphological grounds.
109 There is another potential source of ancient biomolecular evidence: sequence information
110 derived from proteins[41-44]. Because an organism's proteins are coded by its DNA, amino acid
111 sequences in a protein are directly controlled by the gene sequences which specify them.
112 Importantly, proteins—especially structural proteins like collagen and myosin—
113 characteristically degrade at a slower rate than DNA[45-47]. Using tandem mass spectrometry
114 coupled with high-performance liquid chromatography, it has proven possible to recover
115 authentic collagen sequence information from mammalian fossils as old as mid-Pliocene (3.5-3.8
116 Ma)[48], which exceeds the current aDNA record (560–780 kyr BP) by a substantial interval[49, 50].
117 Another advantage is that proteomic data can be potentially recovered from specimens from a
118 wide range of taphonomic contexts, including ones generally inimical to aDNA preservation[51].
119 There are of course limitations. Bones and teeth are typically the only parts of vertebrate bodies
120 that preserve as fossils, which restricts the choice of proteins to ones that occur in significant
121 amounts in such tissues. Type 1 collagen comprises ~90% of the organic fraction of vertebrate
122 bone[52] and is the only bone protein[46] that is well represented in taxonomically extensive libraries
123 such as the National Center for Biotechnology Information (NCBI). Since type 1 collagen is

coded by only 2 genes, COL 1A1 and COL 1A2, only a small fraction of a species' genome can be accessed with this probe. In the context of palaeontology, phylogenetic analyses of type 1 collagen have been shown to yield results that are highly congruent with those produced by aDNA, especially at higher taxonomic levels[43,53].

One such application is testing morphology-based hypotheses of higher-level relationships where there is a strong possibility that pervasive homoplasy among and between target groups has affected morphological character analysis and therefore classification, as in the case of incorrectly homologized caniniform tooth loci in living tree sloths[54]. Because dental features have always played a large role in folivoran systematics,[7, 10, 12, 13 16, 31] such fundamental reinterpretations are likely to have a significant impact. Clearly, it is desirable to use as many sources of inference as possible in reconstructing phylogeny. Also, molecular data lend themselves well to estimating divergence timing of major clades–another critical problem in folivoran systematics[29, 34, 35].

## RESULTS

To address some of the questions raised in the previous section, as well as to add to the available molecular database for folivorans, we utilised proteomic data collected from fossil and living sloths in order to focus on three fundamental issues: (1) relationships of tree sloths to each other and to other folivorans; (2) composition of folivoran superfamilies Megatherioidea and Mylodontoidea; and (3) divergence dating of major sloth ingroups. Results were tested against datasets that additionally incorporated published genomic and phenomic information.

**Samples**. A total of 120 xenarthran samples comprising 24 different genus-level taxa (see Supplementary Information, Table S1 and Fig. S1) were screened for protein survival using both AAR (Amino Acid Racemization) and MALDI-ToF (Matrix-Assisted Laser Desorption/Ionization Time-of-Flight) mass spectrometry. Three additional xenarthran sequences were taken from the literature (see Methods, Proteomic Analysis). Of these, 34 or 28.3% of the total number of samples (including 31.0% of 103 folivoran samples) produced promising results for both AAR and MALDI-ToF MS. From these, the best sample per taxon was selected for LC-MS/MS (Liquid Chromatography-Tandem Mass Spectrometry) analysis to derive protein sequences, with some additions to maximize taxonomic coverage (Fig. 2, Table 1). We resampled the specimen of *Megatherium* previously utilised by ref. 44; the results presented here are *de novo*. The samples of *Neocnus dousman* and *Megalocnus zile* did not pass both MALDI-ToF and AAR screening criteria, but it was decided to analyse them because they were the best representatives of their species. However, because coverage for the *Megalocnus* sample was particularly poor, recovered sequence being mostly contaminants, it was not used in the phylogenetic analyses. To provide modern comparisons, samples of *Bradypus variegatus* (AMNH 20820) and *Choloepus hoffmanni* (AMNH 139772) were also subjected to LC-MS/MS analysis. For further details on all samples, see Supplemental Information, especially Table S1. Relevant procedures for recovering sequence information and estimating phylogenetic relationships are presented in Methods.

Samples ranged in assigned age from late Miocene to mid-Holocene (Supplementary Information Table S1), but the 19 successfully-screened samples are all Quaternary (Table 1). Of these, 15 were selected for radiocarbon dating, and 10 returned finite $^{14}C$ ages (Supplementary Information Table S2). The oldest specimen that yielded sequence information, *Glossotherium*

168 *robustum* MACN-PV 2652, is catalogued as Bonaerian SALMA (South American Land
169 Mammal Age, 128-400 ka[55]), but this age assignment cannot be independently confirmed.

170 To keep nomenclature manageable, we make frequent reference to the relatively simple
171 traditional taxonomic scheme presented in Fig. 1, which is in turn based on a large simultaneous
172 analysis of folivoran relationships[8, 16]. Significant departures from traditional frameworks will be
173 denoted where necessary by an asterisk, but only for formal taxonomic names (e.g.,
174 *Mylodontoidea, i.e., clade redefined to include *Choloepus*, not a traditional member).

175 P**hylogenetic reconstruction.** Parsimony and Bayesian topology searches resulted in largely
176 congruent topologies. Bootstrap Support (BS) under parsimony was generally low, as might be
177 expected given few variable sites, while Bayesian Posterior Probabilities (PP), which make full
178 use of the data, resulted in somewhat higher clade support (Fig. 3; see Supplementary
179 Information, Fig. S2). Although Antillean sloth relationships are not meaningfully resolved,
180 other folivorans assort into two reciprocally monophyletic clades (PP = 0.99) that are consistent
181 with aDNA results[34, 35]. The first includes the three-toed sloth and various extinct taxa
182 traditionally considered megatherioid (PP = 0.97). The sister group relationship of *Megatherium*
183 and *Nothrotheriops* (PP = 0.93) is noncontroversial (Fig. 1), but in the Bayesian consensus we
184 unexpectedly recovered a previously unreported and moderately well-supported pairing of
185 *Megalonyx* with *Bradypus* (PP = 0.89) (see Discussion). The second monophyletic clade (BS =
186 73, PP = 1.00) consists of traditional mylodontoids plus *Choloepus*. Because inclusion of
187 *Choloepus* in this group markedly contrasts with results achieved using morphological datasets,
188 we designate this clade as *Mylodontoidea. Here, *Scelidotherium* + *Scelidodon* is the earliest
189 diverging branch and *Choloepus* is recovered as part of a clade (PP = 0.83) consistent with
190 accepted mylodontid interrelationships[16, 31, 55].

To further interrogate the reliability of our proteomic topologies, we concatenated our collagen sequences with previously published mitochondrial genome sequences (hereafter, "proteomic + genomic data") for all extant folivorans (2 species of *Choloepus,* 4 species of *Bradypus*), two extinct folivorans (*Mylodon darwinii* and *Nothrotheriops shastensis*) and the two extant outgroup taxa[34, 35]. Bayesian analysis (Fig. S3) of the combined dataset yielded a nearly identical topology to that recovered using proteomic data alone, but in this instance *Megatherioidea (including *Bradypus*) and *Mylodontoidea (including *Choloepus*) were unambiguously recovered as reciprocally monophyletic clades (PP = 1). Recovery of a paraphyletic *Bradypus* (with respect to *Megalonyx*) is almost certainly due to a long genomic branch and lack of proteomic data for *B. torquatus,* combined with a comparable lack of genomic data for *Megalonyx.* As the monophyly of *Bradypus* has never been questioned and this result is based exclusively on relative branch lengths, we constrained *Bradypus* monophyly for subsequent analyses, though analyses without a constraint were not noticeably different.

**Molecular clock considerations and divergence time estimates**. Incorporating time as an analytical component in analysis of the combined dataset yielded a well-supported and monophyletic Antillean clade (PP > 0.99), although within-clade relationships were not satisfactorily resolved. More unexpectedly in light of traditional taxonomic concepts, BEAST placed the Antillean clade as a well-supported sister to *Megatherioidea plus *Mylodontoidea (PP =0.97) rather than pairing it with the one or the other. Support for megatherioid (PP > 0.99) and mylodontoid (PP > 0.99) monophyly remained strong, but variable for constituent sub-clades.

The relatively permissive constraints employed for calculating divergences make it difficult to draw detailed conclusions regarding the tempo of sloth diversification, although mean ages in the

214 combined analysis are reasonably consistent with inferences based on genomic[34, 35] as well as
215 morphological[29] data (Fig. 4; Table 2). Posterior mean node ages suggest an early Oligocene
216 origin for folivorans, with megatherioids and mylodontoids diverging in the middle to late
217 Oligocene (Deseadan SALMA) and the generally-recognized families originating within the
218 middle Miocene (Colloncuran-Laventan SALMAs). The combined analysis indicates that the last
219 time *Choloepus* and *Bradypus* shared a common ancestor was ~ 26.9 Ma (95% HPD interval,
220 17.2 - 34.4), which is notably earlier than the estimate ~ 22.36 Ma (95% HPD interval, 16.87 -
221 28.64 Ma; Figs. S5, S6) based on proteomic evidence only and more in line with some recent
222 morphological assessments (e.g., ref. 29).
223

224 **DISCUSSION**

225 In most respects, our higher-level results for Folivora are consistent with recently-published
226 morphology-only phylogenies, but the few ways in which they differ are critical because they
227 have profound implications for macroevolutionary and biogeographical inference. Harmonizing
228 morphological and molecular datasets is complicated, as the molecular results imply that
229 traditional clades exhibit a massive amount of unrecognized homoplasy—or equally
230 unrecognized plesiomorphies, incorrectly interpreted as (syn)apomorphies. Molecular analyses
231 are of course subject to the same challenges, especially in contexts like the present in which
232 samples sizes and information content are limited. It is already widely appreciated that genomic
233 information is exceptionally useful for testing phylogenetic hypotheses; so is proteomic
234 information, especially when it can be shown to be highly congruent with genetic indicators of
235 relationshp[53]. Together, as illustrated here, they provide a strong basis for formulating
236 evolutionary hypotheses:

***Choloepus* is a mylodontoid.** That the two-toed sloth may be closer to traditional mylodontids than to megalonychids, a possibility occasionally raised in morphological studies[16, 24, 57], has been consistently found in recent aDNA investigations[34, 35, 37-40]. Due to the limited number of extinct taxa included in those investigations the exact nature of their relationship has remained indeterminate. However, the multiple tests of phylogenetic relationships and broad taxonomic sampling used in the present study substantiates the conclusion that *Choloepus* is indeed a mylodontoid.

Given the recent ages of all of the taxa investigated, coupled with low rates of sequence evolution, it is unsurprising that divergence estimates based on proteins alone suggest an early/middle Miocene origin for Scelidotheriidae + Mylodontidae (including *Choloepus*). Inclusion of genomic data helps to push these estimates back to the earliest Miocene, but it should be noted that a number of mylodontoid sloths of late Oligocene to late Miocene/early Pliocene age do not fit neatly into better-defined clades. In the past, these taxa were occasionally gathered[9, 20, 21] into the probably nonmonophyletic grouping Orophodontidae. It would be interesting to know on the basis of molecular evidence whether the inclusion of a putative orophodontid would affect the placement of *Choloepus*, possibly moving it stemward (Fig. S7) or help refine divergence time estimates at the base of *Mylodontoidea. At present there is no evidence on point; however, the youngest of these ambiguously-placed taxa, *Octodontobradys*, is late Miocene/early Pliocene in age[58]—young enough to stand a chance of coming within the range of proteomic methods as these continue to improve.

**Megalocnid sloths are monophyletic, and are not part of traditional Megalonychidae.** Antillean sloths have had a complex taxonomic history[7]. In the past, this geographical grouping of folivorans was sometimes regarded as diphyletic, with different island taxa having diverged

from different mainland antecedents[3, 7, 20, 58]. Diphyly now seems unlikely on the basis of our molecular clock results (Fig. 4; see also Supplemental Information, Figs. S5, S6) as well as recent morphology-based studies[16, 25]. Although within-clade relationships are poorly resolved (cf. paraphyletic *Neocnus*), the Antillean clade as a whole resolves as strongly monophyletic (PP >0.99). In light of this fact, as well as clade antiquity, it is appropriate to remove Megalocninae from traditional Megalonychidae and raise it to family level (*Megalocnidae).

***Megalonyx* and *Bradypus* are megatherioids.** Although recent morphology-oriented cladistic studies have usually recovered *Bradypus* as sister to all other folivorans[8-10, 16], genomic approaches[34,35, 39] have consistently paired the three-toed sloth with the extinct North American Pleistocene megatherioid *Nothrotheriops*. On this point the proteomic data presented here are fully compliant with the genomic evidence and support rejection of the inference[9, 16] that Bradypodoidea (i.e., *Bradypus*) is sister to traditional Megatherioidea + Mylodontoidea, as tested by both parsimony (13 additional steps) and Bayesian inference (2*lnBayes Factor = 6.72, support = Strong). Equally controversial is the sister group relationship detected between *Bradypus* + *Megalonyx* (PP = 0.89 - 0.98; Fig. 4; see Supplementary Information, Fig. S3). Although well supported in analyses of both collagen-only and combined proteomic + genomic data, this remains a surprising finding, inasmuch as such an association has never been reported in any taxon-rich phylogenetic study emphasizing morphology. While both the three-toed sloth and *Megalonyx* are likely to be megatherioids cladistically, settling their deeper relationships will require substantially more data than is currently available.

That none of the Antillean sloths used in this study showed any proteomic affinity for *Megalonyx* is also surprising, because much of what has been understood to morphologically characterize non-South American Megalonychidae was based on Antillean species, the fossils of which tend

to be far more complete than those of most other taxa conventionally included in this family[12, 16, 17]. To resolve this conflict, additional high-quality data will be required, genomic and proteomic as well as phenomic. The only certainty at present is that, if *Choloepus* is excluded, Megalonychidae must now be relegated to the list of formerly diverse but now completely extinct folivoran families.

**The West Indies may have been colonized early.** An early appearance of megalocnid sloths in the West Indies has been proposed on general palaeobiogeographical grounds[3, 10, 17, 24, 60], but at present the only pre-Quaternary fossil evidence for Antillean folivorans consists of a morphologically inconclusive partial femur from the early Oligocene (~31 Ma) Yauco Formation of Puerto Rico[61] and unassociated remains attributable to a folivoran, *Imagocnus zazae*, from the late early Miocene (~17.5 Ma) Lagunitas Formation of Cuba[5]. Although "megalonychid" affinities have been assumed for both on biogeographical grounds, now no longer applicable, neither has been included in formal phylogenetic analyses and their placement within Folivora remains uncertain.

The presence of sloths in the West Indies at least as early as the early Miocene is congruent with our mean age estimate (31.2 Ma; Fig. 4, Table 2) for the last common ancestor of sloths sampled in this study. This inference is also roughly consistent with the GAARlandia dispersal hypothesis[5, 62], which holds that northwestern South America and the Greater Antilles were briefly in land connection during the Eocene-Oligocene transition. Without going beyond the very slim body of molecular evidence currently available, there is now at least some basis for hypothesizing that *Megalocnidae might represent an *in situ* Antillean radiation that was emplaced on the islands during the earliest phases of the evolution of the folivoran crown-group—much earlier than previously thought and inconsistent with the hypothesis of a

306 Patagonian origin for Folivora as a whole[9]. If it proves possible to acquire genomic information
307 from Greater Antillean sloth taxa known to have survived into the mid-Holocene[63], we may
308 expect more light to be shed on megalocnid origins.

309 Systematic repositioning of *Bradypus, Choloepus* and megalocnid sloths also permits a better
310 understanding of how often "extreme" arboreality arose during folivoran evolution. The living
311 tree sloths are uniquely defined among extant vertebrates by a combination of relatively rigid
312 hooklike hands and feet, marked limb mobility, extremely long arms, and powerful flexion
313 capabilities in proximal limb joints[19]. None of the West Indian sloths possessed all of these
314 osteological traits, but, importantly, some came close—notably the Puerco Rican species
315 *Acratocnus odontrigonus*, which may have been technically capable of hand- and foot-
316 suspension but probably did not perform the "upside-down" form of locomotion characteristic of
317 extant sloths[7,14]. Remains assigned to the early Miocene Patagonian sloth *Eucholoeops*, possibly
318 part of a clade ancestral to the Antillean radiation, also display many features consistent with
319 highly-developed arboreality[14, 18]. Our phylogenetic results suggest that evolutionary
320 experiments connected with life in the trees probably occurred multiple times, and early on, in
321 folivore evolution. If so, it is puzzling that small-bodied sloths with highly mobile limbs and
322 other arboreal adaptations are as yet unknown for the interval between the early Miocene (e.g.,
323 *Eucholoeops*) and the Quaternary (e.g., *Diabolotherium*)[18]. It is possible that their absence is
324 only apparent, if they lived in heavily forested tropical environments that do not favour
325 fossilization (e.g., mid-Cenozoic proto-Amazonia[64, 65]).

326 The advent of molecular resources providing novel information on both extinct and extant
327 species offers new ways of testing hypotheses about relationships that, in the past, were by
328 necessity based on morphological data alone. Thanks to ongoing improvements in

instrumentation and applicable software, the future for palaeoproteomics should be bright if it can continue to make significant contributions to solving difficult questions like the ones explored here.

A new aDNA study[87] of folivoran phylogeny, published as this paper was going to press, reaches conclusions almost identical to ours regarding the evolutionary relationships of living tree sloths and the phylogenetic distinctiveness of the West Indian radiation. Because the taxonomic distribution of sampled species is not identical in the two studies, there are some minor differences in lower-level relationships and estimated divergence times. However, their detailed agreement overall supports the argument that high-quality protein sequence information is a reliable source of evidence for reconstructing phylogenetic relationships.

## METHODS

### Proteomic Analyses

The 5-number codes following taxon names in this section refer to lab sample ID numbers referenced in Table 1.

**AAR.** Samples were prepared using a slightly modified version of the protocol in ref. 66. A small sub-sample of bone (~1 mg) was hydrolysed in 7M HCL (100 µl per mg) under $N_2$ for 18 hours at 110$^{o}$C. After hydrolysis, the samples were dried down overnight before being re-hydrated in 0.01mM L-homo-arginine as an internal standard. The samples were analysed using reversed phase high pressure liquid chromatography (RP-HPLC) following a slightly modified

version of the protocol developed by ref. 67. Amino acid composition and extent of racemization was used to assess promising samples for sequencing.

**Sample preparation for MS.** The majority of samples (see Supplementary Information, Table S1) were prepared using a slightly modified version of the ZooMS protocol for bone reported by ref. 43. Bone samples (15-30 mg) were demineralized in 250 µl 0.6M HCl for a minimum of 3 weeks at -20°C. This allowed for a gentler demineralization and helped to protect any remaining collagen. After demineralization, the samples were rinsed once in 200 µl 0.01M NaOH, and three times in 200 µl 50mM ammonium bicarbonate (Ambic). The samples were gelatinized by being resuspended in 100 µl 50mM Ambic and heated at 65°C for 1 hour before being digested overnight at 37°C; 50 µl of the heated sample was digested using 1 µl of 0.5 µg/µl porcine trypsin in trypsin resuspension buffer (Promega, UK) and the other 50 µl was dried down and resuspended in 50 µl 100mM Tris solution to be digested with elastase (Worthington; USA) at the same concentration in 10% Tris solution. Two different enzymes were used to increase the protein sequence coverage for LC-MS/MS[43, 68]. Digestion was stopped by the addition of trifluoroacetic acid (TFA) at a concentration of 0.5-1% of the total solution. Peptides were desalted using zip-tips[64] and eluted in 100 µl of 50% acetonitrile (ACN)/0.1% TFA (v/v).

**SDS-PAGE.** Selected samples were analysed using SDS PAGE (Table 1). This method was used on certain samples as the standard ZooMS protocol had not yielded positive results on certain samples that were deemed potentially important phylogenetically. Bone samples were crushed to ~1 um sized particles using a Retsch PM100 ball mill cooled with liquid nitrogen. The ball mill was cleaned with distilled water and methanol before and after each sample[69]. Nanoscale crushing allowed for the highest potential retrieval of proteomic information. 50 mg of powdered sample was heated at 70°C for 10 minutes in 200 µl SDS solubilizing buffer (0.5M Tris base, 5%

SDS, 130mM DTT), cysteines were alkylated by the addition of 6 µl 1M IAA at room temperature in the dark for 30 minutes before the addition of 200 µl of dye solution (0.05% bromophenol blue, 5% glycerol). 20 µl of the samples were run on a Bis-Tris gel (NuPAGE) for 10 minutes to concentrate the samples into a gel plug which was briefly washed in a fixing solution (16% methanol, 10% acetic acid), before being washed twice in boiling water. The gel was stained using Coomassie stain.

The gel plug was cut into approximately 1mm sized cubes in a fume hood with a scalpel and the gel cubes for each sample placed in a separate Eppendorf. The gel pieces were washed in a de-staining solution (66% ammonium bicarbonate 33% acetonitrile) until no more dye could be seen before being washed in the following solvents for 10 minutes per solvent; ACN, HPLC grade water, ACN and 50mM ammonium bicarbonate[70]. The samples were digested overnight with 100 µl 3.125 µg/µl trypsin in 50mM ammonium bicarbonate at 37$^{o}$C and then the tryptic digest was pipetted into a cleaned Eppendorf tube. 100 µl of 70%ACN/1.7% formic acid/0.1% TFA was added to the gel pieces and the gel was heated at 37$^{o}$C for 1 hour with the supernatant being collected and added to the tryptic digest. This step was repeated sequentially with 100mM triethyl ammonium bicarbonate (TEAB) and ACN. The extracted peptides were dried down and then resuspended in 5% Formic acid/0.1% TFA desalted and purified on C18 membranes (Empore) before being eluted in 80% ACN/0.5% acetic acid. The purified peptides were spun to dryness ready for LC-MS/MS analysis.

**MALDI-ToF MS.** 1 µl of sample was spotted in triplicate onto an MTP384 Bruker ground steel MALDI target plate. 1 µl of α-cyano-4-hydroxycinnamic acid matrix solution (1% in 50% Acetonitrile/0.1% Trifluoroacetic acid (v/v/v)) was added to each sample spot and mixed with

the sample[43]. All samples were analysed on a Bruker Ultraflex MALDI-ToF mass spectrometer in triplicate.

**LC-MS/MS.** Most samples were analysed at the Discovery Proteomic Facility (DPF) at Oxford (Table 1). *Choloepus* 17009 and *Mylodon* 16222 were analysed at the Novo Nordisk Foundation Centre for Protein Research (NNFCPR), University of Copenhagen. The *Megalonyx* sample (ID 16849) was run at the Laboratory of Mass Spectrometry and Gaseous Ion Chemistry, Rockefeller University.

At DPF, sample batches were analysed on an Orbitrap Fusion Lumos or Q-Exactive with identical front-end separation, employing an Easyspray column (ES803, 500mmx75µm, Thermo) and a gradient of 2%-35% ACN in 0.1% FA/5%DMSO over 60 minutes. On the Fusion Lumos, MS1 resolution was set to 120,000 with an AGC target of 400,000. MS2 spectra were acquired in TopSpeed mode (3 seconds duty cycle) in the linear ion trap (rapid scan mode) for up to 250ms, with an AGC target of 4,000 and fragmentation in CID mode (35% normalized collision energy). The MS1 resolution on the Q-Exactive was set to 70,000 with an AGC target of 3E6. MS2 spectra for up to 15 precursors were acquired with a resolution of 17,500 and an AGC target of 1E5 for up to 128ms and 28% normalized collision energy (higher-energy collision dissociation). On both instruments, precursors were excluded for 27 seconds from re-selection.

At NNFCPR, dried peptides were resuspended in 50µl of 80% ACN and 0.1% formic acid before being transferred to a 96 well plate and placed in a vacuum centrifuge at 40ºC until approximately 3 µL of solution was left. The samples were rehydrated with 5 or 10 µL (*Mylodon* 16222 and *Choloepus* 17009 respectively) of 0.1% TFA, 5% ACN. Samples were separated on a

15 cm column (75 µm inner diameter) in-house laser pulled and packed with 1.9 µm C18 beads (Dr. Maisch, Germany) on an EASY-nLC 1000 (Proxeon, Odense, Denmark) connected to a Q-Exactive HF (Thermo Scientific, Bremen, Germany) on a 77 min gradient. 5 µl of sample was injected. Buffer A was milliQ water. The peptides were separated with increasing buffer B (80% ACN and 0.1% formic acid), going from 5% to 80% over an 80 minute gradient and a flow rate of 250 nL/min. In addition, a wash-blank injecting 2µl 0.1% TFA, 5% ACN was run in-between each sample to hinder cross-contamination.

The Q-Exactive HF was operated in data dependent top 10 mode. Full scan mass spectra (350-1400 m/z) were recorded at a resolution of 120,000 at m/z 200 with a target value of 3e6 and a maximum injection time of 25 ms for *Choloepus* 17009 and 45ms for *Mylodon* 16222. Fragment ions were recorded with a maximum ion injection time set to 108 ms and a target value set to 2e5 and recorded at a resolution of 60,000 for *Choloepus* 17009 and 30,000 for *Mylodon* 16222. Normalized collision energy was set at 28% and the isolation window was 1.2 m/z with the dynamic exclusion set to 20 s.

At Rockefeller University, peptides were resuspended in 20 uL 5% methanol, 0.2% formic acid. 10 uL were loaded onto an EASY-Spray column (Thermo Fisher Scientific ES800: 15 cm × 75 µm ID, PepMap C18, 3 µm) via an EASY-nLC 1200 and separated over a 120 minute gradient of 2-32% Solvent B (Solvent A = 0.1% formic acid in water, Solvent B = 0.1% formic acid, 95% acetonitrile) during online ESI–MS and MS/MS analyses with a Q Exactive Plus mass spectrometer (Thermo Fisher Scientific). MS/MS analyses of the top 25 precursors in each full scan (300 to 1700 m/z) used the following parameters: resolution: 17,500 (at 200 Th); AGC target: $2 \times 10^5$; maximum injection time: 200ms; isolation width: 2.0 m/z; normalized collision energy: 24%.

**Protein sequence analysis.** The LC-MS/MS raw files were converted to MGF files using Proteowizard[71] and searched against a mammal collagen database which included common contaminants (http://www.thegpm.org/crap/) in PEAKS v7.5. Mass tolerances were set at 0.5Da for the fragment ions and 10ppm for precursor ions and up to 3 missed cleavages were permitted. Searches allowed various post translational modifications (PTMs) including oxidation (MHW) and hydroxylation of proline (both +15.99), deamidation (NQ; +0.98) and pyro-glu from E (-18.01) as well as a fixed PTM of carbamidomethylation (+57.02) which occurs as part of the sample preparation. A maximum of 3 PTMs were allowed per peptide. Protein tolerances were set at 0.5% false discovery rate (FDR), >50% average local confidence (ALC; *de novo* only) and -10lgP score ≥ 20.

Sequences of both COL 1A1 and COL 1A2 were concatenated using previously published mammal collagen consensus sequences taken from NCBI, including sequences for the xenarthrans *Dasypus novemcinctus* (nine-banded armadillo; GenBank: XP_004470764), *Cyclopes didactylus* (silky anteater; Uniprot: COHJP1/COHJP2), and *Lestodon armatus* (extinct mylodontoid sloth, ref. 44). Telopeptides very rarely survive in fossil samples and so these were removed from all sequences. Isoleucine and leucine cannot be differentiated using low energy tandem mass spectrometry and *de novo* sequencing as both amino acids are isobaric. Therefore, the identification of leucine/isoleucine was consistent throughout the sequence analyses concatenated in this study. Our approach is in line with previous phylogenetic studies using collagen as probe[43], under the assumption that MS/MS sequence variation was not interpreted as significant phylogenetic change (see below, Phylogenetic Analyses).

Once a potential collagen sequence was compiled for a given sloth taxon, the sequence was added to the collagen database and the sample was re-run through PEAKS to check for coverage

463 and sequence substitutions. Any differences noted in either the consensus sequences or between
464 different species of sloths were inspected manually. In order for a difference to be considered
465 authentic, it had to occur in more than 1 product ion spectrum and be covered by both b and y
466 ions. For additional discussion, see Supplemental Information and Table S4.

467

468 **Phylogenetic Analyses**

469 Sequences developed from the MS/MS analyses were aligned in Geneious v. 9.1.7[72] using the
470 MUSCLE algorithm[73] with default settings and then checked by eye. Mitochondrial sequence
471 data for extant folivorans and *Mylodon darwinii* were obtained from ref. 35 and supplemented
472 with protein coding sequences for *Nothrotheriops shastensis* from ref. 34. Because the order of
473 genes differs between these two alignments, we extracted and aligned genes for *Nothrotheriops*
474 individually using MUSCLE in Geneious, checking each by eye to ensure accuracy. Of the 2096
475 amino acids in our alignment of the type 1 collagen molecule, 134 (6.4%) were variable and 76
476 (56 % of variable sites, 3.6% of total) were parsimony informative for the taxa represented.

477 We conducted three sets of phylogenetic analyses on the resulting protein alignment (see
478 Results). We first performed a Strict Parsimony (SP) analysis using PAUP v. 4.0a (build 157)[74].
479 We employed a branch and bound search with all sites treated as unordered and equally
480 weighted. To assess clade support, we performed 10,000 bootstrap replicates using full heuristic
481 tree searches and generated a weighted 50% majority rule (MR) consensus tree from the
482 resulting sample of most-parsimonious bootstrapped trees.

483 We performed two forms of model-based phylogenetic analyses, both in a Bayesian framework.
484 We used PartitionFinder v. 2.1.1[75, 76] to determine the most appropriate model(s) of amino-acid

485 substitution and partitioning scheme for our concatenated alignment, resulting in selection of
486 separate Dayhoff models [77] with gamma-distributed rates for COL 1A1 and COL 1A2. The first
487 set of Bayesian phylogenetic analyses used MrBayes v 3.2.5[78]. We performed two Markov Chain
488 Monte Carlo (MCMC) runs, each of four chains (one cold, three heated), for 10,000,000
489 generations, sampling from the chain every 5000 generations. After checking for convergence of
490 the two chains based on Gelman-Rubin statistics and ensuring that effective sample sizes for all
491 parameters were sufficient (> 200), we discarded the first 50% of each chain as burn-in,
492 combined the remaining posterior samples and summarized them as a 50% majority rule
493 consensus tree, with clade frequencies interpreted as posterior probabilities for a given clade. To
494 determine whether our unconstrained topology provided a better explanation of the data than a
495 previously proposed morphological topology[16] in which *Bradypus* is the sister lineage to all other
496 folivorans and *Choloepus,* Megalocnidae, and *Megalonyx* form a monophyletic Megalonychidae
497 (including other taxa not referenced here), we estimated the marginal likelihood of the data on
498 unconstrained and constrained topologies using the stepping stone algorithm in MrBayes. We
499 performed two runs, each with four chains (three heated, one cold) for 10,000,000 generations
500 over 50 steps, with default settings for the Alpha parameter of the Beta distribution (0.4) and
501 burn-in (-1). We calculated $2*Ln(lnLk_{unconstrained} - lnLk_{constrained})$ from the resulting estimates and
502 assessed support using the scale in ref. 79.

503 The fact that we cannot differentiate between isoleucine and leucine using low energy tandem
504 mass spectrometry creates a unique problem for model-based phylogenetic inference procedures.
505 The standard approach in ancient protein studies[43] is to designate all sites with a molecular mass
506 of 131.17 g/mol as leucine, but this has the potential to bias estimates of the instantaneous rate
507 matrix, branch lengths and, possibly, topology by entirely excluding one amino acid. We

508 investigated this by replacing all peptides coded as leucine with ambiguous codings {IL} and
509 repeating Bayesian estimation of topology and branch lengths using MrBayes. The resulting 50%
510 majority rule consensus tree was identical across coding schemes, and comparison of branch
511 length estimates among analyses show no significant deviation from 1:1 (branch $length_{[Leucine]}$ = -
512 0.00009 + branch $length_{[ambiguous]}$*0.96, $R^2 = 0.995$, $p << 0.001$), indicating that the use of leucine
513 is appropriate. We repeated Bayesian analyses of the combined proteomic + genomic dataset
514 using the same settings but with partitioning schemes and substitution models for genetic data
515 following ref. 35.

516 We attempted to integrate our combined molecular dataset with a large, recently-published
517 morphological dataset (ref. 9). The resulting majority rule consensus tree (Fig S7) is congruent in
518 some respects with our molecular topologies (e.g., C*holoepus* was recovered as a mylodontoid
519 and *Bradypus* as a megatherioid) but other results repeatedly found in molecular analyses were
520 not obtained. In particular, we recovered a strong (PP = 1.0) traditional Megalonychidae nested
521 within Megatherioidea that included Antillean sloths minus *Choloepus*. Although the Antillean
522 species were represented in the total dataset by proteomic sequences, genomic data were
523 unavailable. This result suggests that the large number of morphological characters, some known
524 to be highly homoplastic[54], were able to swamp the signal arising from the smaller proteomic
525 dataset. While combined analysis of morphological and molecular data will ultimately be
526 necessary to fully resolve folivoran phylogeny, this exercise suggests that it is premature to
527 consider such simultaneous analyses reliable at this point in time.

528
529 Our MrBayes analyses sample tree topologies with branch lengths in units of substitutions per
530 site and so ignore temporal information inherent in phylogenetic analysis of non-
531 contemporaneous tips or external information about relative branch lengths that can be provided

by the fossil record. We therefore also performed a series of Bayesian tree searches assuming a molecular clock under the fossilized birth-death framework[80-82], as implemented in BEAST v2.5.1[83]. Briefly, this framework allowed us to sample from the posterior distribution of time-scaled trees for taxa in our proteomic dataset, inferred using their sequences and stratigraphic ages, while using phylogenetically constrained fossil taxa that lack amino acid data to provide additional information on relative branch lengths and divergence times. Our choice of fossil taxa and topological constraints broadly followed the approach undertaken in ref. 34 for sloth mitogenomes. However, our proteomic topologies raise questions about the phylogenetic positioning of some fossil folivorans that have previously been considered on morphological grounds as early representatives of Pleistocene and Holocene families. For example, some extinct folivorans, such as the Huayquerian nothrotheriid *Mionothropus*[84] can be plausibly assigned to a specific terminal branch in our proteomic topology. Others, however, are customarily assigned to clades that we failed to recover. This applies to the Santacrucian taxon *Eucholoeops,* usually interpreted as a basal megalonychid[24, 85, 86] and therefore as a member of a clade not found to be monophyletic in our analyses. Such issues inevitably affect efforts to calibrate the proteomic + genomic data clock and to infer divergence times. Acknowledging this, we employed a minimal set of constraints (see Supplemental information, Fig. S4) on the positioning of fossil folivorans in our Bayesian estimation of topology and divergence times, integrating over all possible placements of phylogenetically uncertain fossils using stratigraphic context alone when necessary. We performed analyses with and without a monophyly constrain on *Bradypus* and results did not differ at unaffected nodes.

The use of a Bayesian approach requires the specification of prior probabilities on model parameters. We used default priors on substitution model parameters but specified the following:

net diversification ~ Exp(1), yielding a broad, vague prior; turnover ~ beta(2,1), yielding high prior weight on extinction ≅ speciation; sampling probability ~ beta(2,2) yielding a humped distribution that placed most prior weight on sampling probabilities of 0.5; origin ~ U(61.5, 150) yielding a flat prior on ages older than 61.5 Ma up to 150 Ma. In addition, the analysis was conditioned on the number of extant taxa sampled ($\rho = 0.129$ in the xenarthran proteomic analyses, $\rho = 0.333$ in the folivoran proteomic analyses, $\rho = 0.266$ in the combined analyses). Based on comparisons of marginal likelihoods computed via Path Sampling (see Supplementary Information, Table S3), we employed a relaxed uncorrelated clock with log-normally distributed rates for proteomic and combined analyses, with an exponential prior (mean=0.1) placed on the mean of log-normal distribution and the default gamma $\Gamma(0.5396, 0.3819)$ on the standard deviation. Two MCMC analysis were run for 10 million generations each, sampling every 1000 generations, after which fossils without data were pruned from the trees, the first 20% of the retained samples were discarded as burn-in, the samples combined, and maximum clade credibility trees constructed using the tree annotator software accompanying the BEAST suite. Runs from the prior using a fixed topology (the maximum clade credibility tree based on the pre-pruning sample) were used to confirm that divergence time estimates were not simply returning the prior.

**Data availability** Mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE partner repository with the dataset identifier PXD012859. Collagen sequences are available on the Uniprot website (https://www.uniprot.org/); the complete list can be found in Supplemental Information, Table S5. Phylogenetic datasets have been deposited on DataDryad (doi:10.5061/dryad.7dd64gs).

## Supplementary Information

[URL tk]


## Acknowledgements

We thank the curatorial staffs of the following museums and private collections for permission to sample specimens in their care: AMNH-M, American Museum of Natural History (Mammalogy), New York, USA; AMNH-P, American Museum of Natural History (Paleontology), New York, USA; CIV, Iota Quatro faunal collection, courtesy of Lazaro Vinola; El Trebol faunal collection, Bariloche, Argentina; FR, Forest Reserve (Trinidad) faunal collection currently housed in Department of Mammalogy, AMNH, New York, USA; IANIGLA-PV, Instituto Argentino de Nivología, Glaciología y Ciencias Ambientales, CCT-CONICET-Mendoza, Mendoza, Argentina; MACN-PV, Museo Argentino de Ciencias Naturales "Bernardino Rivadavia" (Sección Paleovertebrados), Buenos Aires, Argentina; MAPBAR, Museo de la Asociación Paleontológica Bariloche (APB), prov. Río Negro, Argentina; MMP, Museo Municipal de Ciencias Naturales "Lorenzo Scaglia" Mar del Plata, prov. Buenos Aires, Argentina; MNHN SAO, Muséum national d'Histoire naturelle, Paris, France; MPS, Museo Paleontológico "Fray Manuel de Torres", San Pedro, prov. Buenos Aires, Argentina; MUSM, Museo de Historia Natural de la Universidad Nacional Mayor de San Marcos, Lima, Peru; NYSM VP, New York State Museum (Vertebrate Paleontology), Albany, New York, USA; RM, Cuban faunal collection currently housed in Department of Mammalogy, AMNH, New York, USA; UF, University of Florida, Natural History Museum of Florida (Vertebrate Paleontology), Gainesville, USA; UMAG ah, Instituto de La Patagonia, Universidad de Magallanes, Punta

Arenas, Chile; USNM, United States National Museum of Natural History (Paleobiology), Washington DC, USA. SP would like to thank Beatrice Demarchi for useful discussion and support. The authors thank the National Science Foundation for grants OPP 0636639 to RDEM and DEB 1547414 to RDEM, MC, and KP.

## Author contributions

RDEM, MC, and SP conceived the project. SP undertook AAR and proteomic analysis and concatenated collagen sequences, with laboratory and technical assistance from RF, JO, KM, MM, MC, KP, and BC. GJS conducted phylogenetic analyses. FP and AMF supplied paleontological information. AK, MT, FS, ML, AH, AI, RF, JB, JLL, FMM, RSG, MR, AG, CdM, and GB supplied fossil samples, locality information, species identifications, and commentary on the manuscript. RDEM, SP, and GJS wrote the manuscript, with input from all authors.

**Competing interests.** The authors declare no competing interests.

**Materials and Correspondence.** Address inquiries to macphee@amnh.org

617 **REFERENCES CITED**


618 [1] Gardner, A. L. in *Mammals of South America*, **1** (ed Gardner, A. L.) 157-176 (University of
619 Chicago Press, 2007).
620
621 [2] Nowak, R. *Walker's Mammals of the World: Monotremes, Marsupials, Afrotherians,*
622 *Xenarthrans, and Sundatherians* (Johns Hopkins, 2018).
623
624 [3] Kraglievich, L. Descripción de dos cráneos y otros restos del género ''*Pliomorphus*''
625 Ameghino, procedentes de la formación entrerriana de las barrancas del río Paraná. *Anal. Mus.*
626 *Nac. Hist. Nat. Buenos Aires* **33**, 1-56 (1923).
627
628 [4] Hoffstetter, R. in *Traité de Paléontologie*, **6.2** (ed Piveteau, J.) 535-636 (Masson, 1958).
629
630 [5] MacPhee, R. D. E. & Iturralde-Vinent, M. A. Origin of the Greater Antillean land mammal
631 fauna 1: New Tertiary land mammals from Cuba and Puerto Rico. *Amer. Mus. Novitates* **3141**, 1-
632 31 (1995). http://hdl.handle.net/2246/3657
633
634 [6] Iturralde-Vinent, M. A. & MacPhee, R. D. E. Paleogeography of the Caribbean region:
635 implications for Cenozoic biogeography. *Bull. Amer. Mus. Nat. Hist.* **238**, 1-95 (1999).
636 http://hdl.handle.net/2246/1642
637
638 [7] White, J. & MacPhee, R.D.E. in *Biogeography of the West Indies: Patterns and Perspectives*,
639 2nd ed (eds Woods, C.A. & Sergile, F. E.) 201-236 (CRC Press, 2001).

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

[9] 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

[10] Pujos, F., De Iuliis, G. & Cartelle, C. A paleogeographic overview of tropical forest sloths: Towards an understanding of the origin of extant suspensory sloths? *J. Mammal. Evol.* **24** (1), 19-38 (2017). doi:10.1007/s10914-016-9330-4

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

[12] Engelmann, G. F. in *The Evolution and Ecology of Armadillos, Sloths, and Vermilinguas* (ed Montgomery, G. G.) 195-203 (Academic Press, 1985).

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

[14] White, J. Indicators of locomotor habits in xenarthrans: evidence of locomotor heterogeneity among fossil sloths. *J. Vert. Paleontol.* **13** (2), 230-242 (1993). https://www.jstor.org/stable/4523502

[15] Delsuc, F., Catzeflis, F. M., Stanhope, M. J. & Douzery, E. J. P. 2001. The evolution of armadillos, anteaters and sloths depicted by nuclear and mitochondrial phylogenies: implications for the status of the enigmatic fossil *Eurotamandua*. *Proc. Roy. Soc. B* **268**, 1605-1615 (2001). doi: 10.1098/rspb.2001.1702

[16] 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

[17] McDonald, H. G. & De Iuliis, G. in *The Biology of Xenarthra* (eds Vizcaino, S. F. & Loughry, W. J.) 39-55 (University Press of Florida, 2008).

[18] Pujos, F., Gaudin, T. J., De Iuliis, G. & Cartelle, C. Recent advances on variability, morpho-functional adaptations, dental terminology, and evolution of sloths. *J. Mamm. Evol.* **19**, 159-169 (2012). doi: 10.1007/s10914-012-9189-y

[19] Nyakatura, J. A. The convergent evolution of suspensory posture and locomotion in tree sloths. *J. Mamm. Evol.* **19** (3), 225-234 (2012). doi: 10.1007/s10914-011-9174-x

684 [20] McKenna, M. C. & Bell, S. K. *Classification of Mammals above the Species Level* (Columbia
685 University Press, 1997).
686
687 [21] Patterson, B., Turnbull, W. D., Segall, W. & Gaudin, T. J. The ear region in xenarthrans (=
688 Edentata: Mammalia). Part II. Pilosa (sloths, anteaters), palaeanodonts, and a miscellany.
689 *Fieldiana Geol.* **24**, 1-79 (1992). doi: 10.5962/bhl.title.3466
690
691 [22] Pujos, F. *Megatherium celendinense* sp. nov. from the Pleistocene of Peruvian Andes and the
692 megatheriine phylogenetic relationship. *Palaeontology* **49** (2): 285-306 (2006). doi:
693 10.1111/j.1475-4983.2006.00522.x
694
695 [23] Pujos, F., De Iuliis, G. & Mamani Quispe, B. *Hiskatherium saintandrei*, gen. et sp. nov.: An
696 unusual sloth from the Santacrucian of Quebrada Honda (Bolivia) and an overview of Middle
697 Miocene, small megatherioids. *J. Vert. Paleontol.* **31** (5), 1131-1149 (2011). doi:
698 10.1080/02724634.2011.599463
699
700 [24] McDonald, H. G., Rincón, A. D. & Gaudin, T. J. A new genus of megalonychid sloth
701 (Mammalia, Xenarthra) from the late Pleistocene (Lujanian) of Sierra De Perija, Zulia State,
702 Venezuela. *J. Vert. Paleontol.* **33** (5), 1226-1238 (2013). doi: 10.1080/02724634.2013.764883
703
704 [25] McDonald, H. G. & Carranza-Castaneda, O. Increased xenarthran diversity of the Great
705 American Biotic Interchange: a new genus and species of ground sloth (Mammalia, Xenarthra,

706 Megalonychidae) from the Hemphillian (Late Miocene) of Jalisco, Mexico. *J. Paleontol.* **91** (5),
707 1-14 (2017). doi: 10.1017/jpa.2017.45
708
709 [26] Brandoni, D. A new genus of Megalonychidae (Mammalia, Xenarthra) from the Late Miocene
710 of Argentina. *Rev. Bras. Paleontol.* **17** (1), 33-42 (2014). doi: 10.4072/rbp.2014.1.04
711
712 [27] Brandoni, D. The Megalonychidae (Xenarthra, Tardigrada) from the late Miocene of Entre
713 Ríos Province, Argentina, with remarks on their systematics and biogeography. *Geobios* **44**, 33-
714 44 (2011). doi: 10.1016/j.geobios.2010.06.005
715
716 [28] De Iuliis, G., Gaudin, T. J. & Vicars, M. J. A new genus and species of nothrotheriid sloth
717 (Xenarthra, Tardigrada, Nothrotheriidae) from the Late Miocene (Huayquerian) of Peru.
718 *Palaeontology* **54** (1), 171-205 (2011). doi: 10.1111/j.1475-4983.2010.01001.x
719
720 [29] Gaudin, T. J. & Croft, D. Paleogene Xenarthra and the evolution of South American mammals.
721 *J. Mamm.* **96** (4), 622-634 (2015) doi: 10.1093/jmammal/gyv073
722
723 [30] Rincón, A. D., Solórzano, A., McDonald, H. G. & Montellano-Ballesteros, M. Two new
724 megalonychid sloths (Mammalia: Xenarthra) from the Urumaco Formation (Late Miocene), and
725 their phylogenetic affinities. *J. Syst. Palaeontol.* **17** (5), 409-421 (2019).
726 doi:10.1080/14772019.2018.1427639
727

[31] Boscaini, A., Gaudin, T. J., Mamani Quispe, B., Antoine, P.-O. & Pujos, F. New well-preserved craniodental remains of *Simomylodon uccasamamensis* (Xenarthra, Mylodontidae) from the Pliocene of the Bolivian Altiplano: phylogenetic, chronostratigraphic and paleobiogeographic implications. *Zool. J. Linnean Soc*. **185** (2), 459-486 (2019). doi: 10.1093/zoolinnean/zly075

[32] McDonald, H. G. & De Iuliis, G. in *The Biology of Xenarthra* (eds Vizcaino, S. F. & Loughry, W. J.) 39-55 (University Press of Florida, 2008).

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

[34] 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

[35] 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

[36] 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** (1), 86-100 (2011). doi: 10.1644/10-MAMM-A-086.1

[37] 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

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

[39] 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). doi: 10.1006/mpev.2000.0860

[40] 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

[41] Welker, F., Smith, G. M., Hutson, J. M., Kindler, L., Garcia-Moreno, A., Villaluenga, A., Turner, E. & Gaudzinski-Windheuser, S. Middle Pleistocene protein sequences from the rhinoceros genus *Stephanorhinus* and the phylogeny of extant and extinct Middle/Late Pleistocene Rhinocerotidae. *PeerJ* **5**, e3033 (2017). doi: 10.7717/peerj.3033

[42] Welker, F., *et al.* Palaeoproteomic evidence identifies archaic hominins associated with the Châtelperronian at the Grotte du Renne. *Proc. Nat. Acad. Sci. USA* **113** (40), 11162-11167 (2016). doi:org/10.1073/pnas.1605834113

[43] 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

[44] 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

[45] Dobberstein, R. C., Collins, M. J., Craig, O. E., Taylor, G., Penkman, K. E. H. & Ritz-Timme, S. Archaeological collagen: why worry about collagen diagenesis? *Archaeol. Anthropol. Sci.* **1** (1), 31-42 (2009). doi: 10.1007/s12520-009-0002-7

[46] Buckley, M. & Collins, M. J. Collagen survival and its use for species identification in Holocene-Lower Pleistocene bone fragments from British archaeological and palaeontological sites. *Antiqua* **1** (1), e1 (2011). doi: 10.4081/antiqua.2011.e1

[47] Buckley, M. & Wadsworth, C. Proteome degradation in ancient bone: diagenesis and phylogenetic potential. *Palaeogeog. Palaeoclimat. Palaeoecol.* **416**, 69-79 (2014). doi: 10.1016/j.palaeo.2014.06.026

[48] Rybczynski, N., Gosse, J. C., Harington, C. R., Wogelius, R. A., Hidy, A. J. & Buckley, M. Mid-Pliocene warm-period deposits in the High Arctic yield insight into camel evolution. *Nat. Comm.* **4**, 1550 (2013). doi:10.1038/ncomms2516

[49] Allentoft, M. E., *et al.* The half-life of DNA in bone: measuring decay kinetics in 158 dated fossils. *Proc. Roy. Soc. B* **279** (1748), 4724–4733 (2012). doi: 10.1098/rspb.2012.1745

[50] Orlando, L., *et al.* Recalibrating *Equus* evolution using the genome sequence of an early Middle Pleistocene horse. *Nature* **499** (7456), 74-78 (2013). doi: 10.1038/nature12323

[51] Presslee, S., *et al*. Radiocarbon dating and proteomic analysis of highly purified bone collagen derived from Rancho La Brea mammal fossils. *J. Vert. Paleont.* Program & Abstracts 2016, 208 (2016).

[52] Tuross, N. & Stathoplos, L. in *Methods in Enzymology* **224** (ed Zimmer A., White, T. J., Cann, R. L. & Wilson, A. C.) 121-129 (Academic Press, 1993).

[53] Westbury, M., *et al.* A mitogenomic timetree for Darwin's enigmatic "transitional" South American mammal, *Macrauchenia patachonica*. *Nat. Commun.* **8**, 15951 (2017) doi: 10.1038/ncomms15951.

816 [54] Hautier L., Gomes Rodrigues H., Billet G. & Asher R. J. The hidden teeth of sloths:
817 evolutionary vestiges and the development of a simplified dentition. *Sci. Rep.* **6**, 27763 (2016).
818 doi:10.1038/srep27763
819
820 [55] Cione, A. L. & Tonni, E. P. in *Quaternary of South America Antarctic Península* (eds Tonni,
821 E.P. & Cione, A.L.) 23-51 (Balkema,1999).
822
823 [56] Cartelle, C., De Iuliis, G. & Ferreira, R. L. Systematic revision of tropical Brazilian
824 scelidotheriines sloths (Xenarthra, Mylodontoidea). *J. Vertebr. Paleontol.* **29** (2), 555–566
825 (2009). doi:org/10.1671/039.029.0231
826
827 [57] Guth, C. *La Région Temporale des Edentés* (Imprimerie Jeanne d'Arc Le Puy, 1961).
828
829 [58] Guilherme, E., Bocquentin, J. & Porto, A. S. A new specimen of the genus *Octodontobradys*
830 (Orophodontidae, Octodontobradyinae) from the Late Miocene-Pliocene of the southwestern
831 Amazon Basin, Brazil. *Anuár. Inst. Geociências* **34** (1), 64-71 (2011). www.anuario.igeo.ufrj.br
832
833 [59] Varona, L. *Catálogo de los Mamíferos Vivientes y Extinguidos de las Antillas* (Academia de
834 Ciencias de Cuba, 1974).
835
836 [60] Webb, S. D. & Perrigo, S. in *The Evolution and Ecology of Armadillos, Sloths, and*
837 *Vermilinguas* (ed Montgomery, G. G.) 113-120 (Smithsonian Institution Press, 1985).
838

[61] MacPhee, R. D. E., Iturralde-Vinent, M. A. & Gaffney, E. S. Domo de Zaza: An Early Miocene vertebrate locality in south-central Cuba, with notes on the tectonic evolution of Puerto Rico and Mona Passage. *Amer. Mus. Novitates* **3394**, 1-42 (2003). http://hdl.handle.net/2246/2820

[62] Tong, Y. F., Binford, G., Rheims, C. A., Kuntner, M., Liu, J., Agnarsson, I. Huntsmen of the Caribbean: Multiple tests of the GAARlandia hypothesis. *Mol. Phylogenet. Evol.* **130**, 259-268 (2019). doi:10.1016/j.ympev.2018.09.017

[63] Steadman, D. W., Martin, P. S., MacPhee, R. D. E., Jull, A. J. T., McDonald, H. G., Woods, C. A., Iturralde-Vinent, M. A. & Hodgins, G. W. L. Asynchronous extinction of late Quaternary sloths on continents and islands. *Proc. Nat. Acad. Sci. USA* **102** (33), 11763-11768 (2005). doi: 10.1073/pnas.0502777102

[64] Hoorn, C., *et al.* Amazonia through time: Andean uplift, climate change, landscape evolution, and biodiversity. *Science* **330** (6006), 927-931 (2010). doi: 10.1126/science.1194585

[65] Tejada-Lara, J. V., *et al.* Life in proto-Amazonia: Middle Miocene mammals from the Fitzcarrald Arch (Peruvian Amazonia). *Palaeontology* **58** (2), 341-378 (2015). doi: 10.1111/pala.12147

[66] Penkman, K., Kaufman, D. S., Maddy, D. & Collins, M. J. Closed-system behaviour of the intra-crystalline fraction of amino acids in mollusc shells. *Quat. Geochronol.* **3** (1-2), 2-25 (2008). doi: 10.1016/j.quageo.2007.07.001.

[67] Kaufman, D. S. & Manley, W. F. A new procedure for determining DL amino acid ratios in fossils using reverse phase liquid chromatography. *Quat. Sci. Rev.* **17** (11), 987-1000 (1998). doi: 10.1016/S0277-3791(97)00086-3

[68] Demarchi, B., *et al.* Protein sequences bound to mineral surfaces persist into deep time. *eLife* **5,** e17092 (2016). doi: 10.7554/eLife.17092

[69] Kontopoulos, I., Presslee, S., Penkman, K. & Collins, M. J. Preparation of bone powder for FTIR-ATR analysis: the particle size effect. *Vibrational Spectroscopy* **99**, 167–177 (2018). doi: 10.1016/j.vibspec.2018.09.004

[70] Van Doorn, N. L., Hollund, H. & Collins, M. J. A novel and non-destructive approach for ZooMS analysis: Ammonium bicarbonate buffer extraction. *Archaeol. Anthropol. Sci.* **3** (3), 281-289 (2011). doi: 10.1007/s12520-011-0067-y

[71] Chambers, M. C., *et al.* A cross-platform toolkit for mass spectrometry and proteomics. *Nature Biotechnol.* **30**, 918-920 (2012). doi: 10.1038/nbt.2377

882 [72] Kearse, M., *et al.* Geneious Basic: an integrated and extendable desktop software platform for
883 the organization and analysis of sequence data. *Bioinformatics* **28** (12), 1647-1649 (2012). doi:
884 10.1093/bioinformatics/bts199
885
886 [73] Edgar, R. C. MUSCLE: multiple sequence alignment with high accuracy and high throughput.
887 *Nucl. Acids Res.* **32** (5), 1792-1797 (2004). doi: 10.1093/nar/gkh340
888
889 [74] Swofford, D. L. PAUP* *Phylogenetic Analysis Using Parsimony (and Other Methods).*
890 Version 4 (Sinauer Associates, 2002). doi: 10.1002/0471650129.dob0522
891
892 [75] Lanfear, R., Calcott, B., Ho, S. Y. W. & Guindon, S. PartitionFinder: combined selection of
893 partitioning schemes and substitution models for phylogenetic analyses. *Mol. Biol. Evol.* **29** (6),
894 1695-1701 (2012). doi: 10.1093/molbev/mss020
895
896 [76] Lanfear, R., Frandsen, P. B., Wright, A. M., Senfeld, T. & Calcott B. PartitionFinder 2: New
897 methods for selecting partitioned models of evolution for molecular and morphological
898 phylogenetic analyses. *Mol. Biol. Evol.* **34** (3), 772-773 (2017). doi: 10.1093/molbev/msw260
899
900 [77] Dayhoff, M. O., Schwartz, R. M. & Orcutt, B. C. in *Atlas of Protein Sequence and Structure* **5**
901 (ed Dayhoff, M. O.) 345-352 (National Biomedical Research Foundation, 1978).
902
903 [78] Ronquist, F., Teslenko, M., Van Der Mark, P., Ayres, D.L., Darling, A., Höhna, S., Larget, B.,
904 Liu, L., Suchard, M.A. & Huelsenbeck, J.P. MrBayes 3.2: Efficient Bayesian phylogenetic

905 inference and model choice across a large model space. *Sys. Biol.* **61** (3), 539-542 (2012). doi:
906 10.1093/sysbio/sys029
907
908 [79] Kass, R.E. & Raftery, A. E. Bayes Factors. *J. Amer. Stat. Assoc.* **90** (430), 773-795 (1995).
909
910 [80] Heath, T. A, Huelsenbeck, J. P. & Stadler, T. The fossilized birth-death process for coherent
911 calibration of divergence-time estimates. *Proc. Nat. Acad. Sci. USA.* **111** (29), 2957-2966 (2014).
912 doi: 10.1073/pnas.1319091111
913
914 [81] Gavryushkina, A., Welch, D., Stadler, T. & Drummond, A. J. Bayesian inference of sampled
915 ancestor trees for epidemiology and fossil calibration. *PLoS Comput. Biol.* **10** (12), e1003919
916 (2014). doi: 10.1371/journal.pcbi.1003919
917
918 [82] Gavryushkina, A., Heath, T. A., Ksepka, D. T., Stadler, T., Welch, D. & Drummond, A. J.
919 Bayesian total-evidence dating reveals the recent crown radiation of penguins. *Syst. Biol.* **66** (1),
920 57-73 (2017). doi: 10.1093/sysbio/syw060
921
922 [83] Bouckaert, R., Heled, J., Kühnert, D., Vaughan, T., Wu, C.-H., Xie, D., Suchard, M. A.,
923 Rambaut, A. & Drummond, A. J. BEAST 2: a software platform for Bayesian evolutionary
924 analysis. *PLoS Comput. Biol.* **10** (4) e1003537 (2014). doi: 10.1371/journal.pcbi.1003537
925
926 [84] De Iuliis, G., Gaudin, T. J. & Vicars, M. J. A new genus and species of nothrotheriid sloth
927 (Xenarthra, Tardigrada, Nothrotheriidae) from the Late Miocene (Huayquerian) of Peru.
928 *Palaeontology* **54**, 171-205 (2011). doi: 10.1111/j.1475-4983.2010.01001.x

[85] De Iuliis, G., Pujos, F., Toledo, N., Bargo, M. S. & Vizcaíno, S. F. *Eucholoeops* Ameghino, 1887 (Xenarthra, Tardigrada, Megalonychidae) from the Santa Cruz Formation, Argentine Patagonia: implications for the systematics of Santacrucian sloths. *Geodiversitas* **36** (2), 209-255 (2014). doi: 10.5252/g2014n2a2

[86] Hirschfeld, S. E. & Webb, S. D. Plio-Pleistocene megalonychid sloths of North America. *Bull. Florida Mus. Nat. Hist. Biol. Sci.* **12** (5), 213-294 (1968).

[87] Delsuc, F. *et al.* Ancient mitogenomics rewrites the evolutionary history and biogeography of sloths. *Current Biology*. [tk]

## FIGURE AND TABLE LEGENDS (MAIN TEXT)

**Figure 1:** Phylogenetic relationships among major folivoran taxa based on morphological evidence (mostly after ref. 8, 16), with existence of unallocated taxa acknowledged. In this framework, the three-toed tree sloth *Bradypus* is sister to other sloths (grouped here as Eutardigrada), while the two-toed tree sloth *Choloepus* is included within Megalonychidae.

**Figure 2:** Geographical locations of sequenced samples. Sequences for *Cyclopes* and *Lestodon* (**in bold**) taken from the literature; others, this paper (Table 1 and Supplementary Information, Fig. S1).

951 **Figure 3:** 50% majority rule consensus tree from Bayesian analysis of the proteomic data
952 without temporal information, as performed in MrBayes. Values below nodes are posterior
953 probabilities for the descendant clade (see Results). Values above nodes are bootstrap support
954 derived from 10,000 bootstrap replicates. A dash (–) indicates that a node was not represented in
955 the 50% majority rule bootstrap consensus. Extant *Dasypus* and extinct *Doedicurus* and
956 *Glyptodon* are members of the order Cingulata; extant *Cyclopes* is a representative of
957 Vermilingua, which together with Folivora comprise order Pilosa. Cingulates and pilosans
958 together comprise superorder Xenarthra (see also Fig. 4).

959

960 **Figure 4:** Time scaled maximum clade credibility tree from BEAST analysis of 24 extant and
961 extinct xenarthran collagen sequences plus published mitochondrial genomes (see text). Branch
962 lengths are the mean values from the retained posterior sample, while blue bars represent 95%
963 highest posterior density intervals. Values at nodes are posterior probabilities (note that the
964 monophyly of *Bradypus* is constrained here). Vertical shaded bars correspond to South American
965 land mammal ages (SALMAs), two of which are emphasized: Deseadan (**), 29–21 Ma, during
966 which the first generally-accepted representatives of traditional Megatherioidea and
967 Mylodontoidea appear paleontologically; and the Santacrucian (*), 17.5–16.3 Ma, the SALMA
968 during which mylodontids maintained substantial taxonomic diversity but megalonychids and
969 megatheriids declined[9]. On the right (grey boxes), folivoran species used in analyses are
970 associated with their traditional family names, but with superfamily contents organized
971 according to phylogenetic conclusions in text. Megalocnidae is placed outside traditional
972 superfamily structure in its own (unnamed) box. The tree implies that the fundamental split
973 within Folivora is not between Megatherioidea and Mylodontoidea vs. Bradypodoidea as

974 classically understood, but instead between redefined *Megatherioidea and *Mylodontoidea vs.
975 Megalocnidae.
976

977 **Table 1:** Collagen peptides and per cent coverage of the sequenced ancient and modern samples.
978

979 **Table 2:** Selected divergence time estimates from BEAST analyses using different combinations
980 of taxa and data (see Results and Supplementary Information). Note that, although consistently
981 recovered as monophyletic, the position of Megalocnidae shifted among analyses, falling
982 alternately as sister to all other Folivora (Xenarthra) or Megatherioidea (Folivora).
983

Fig. 1

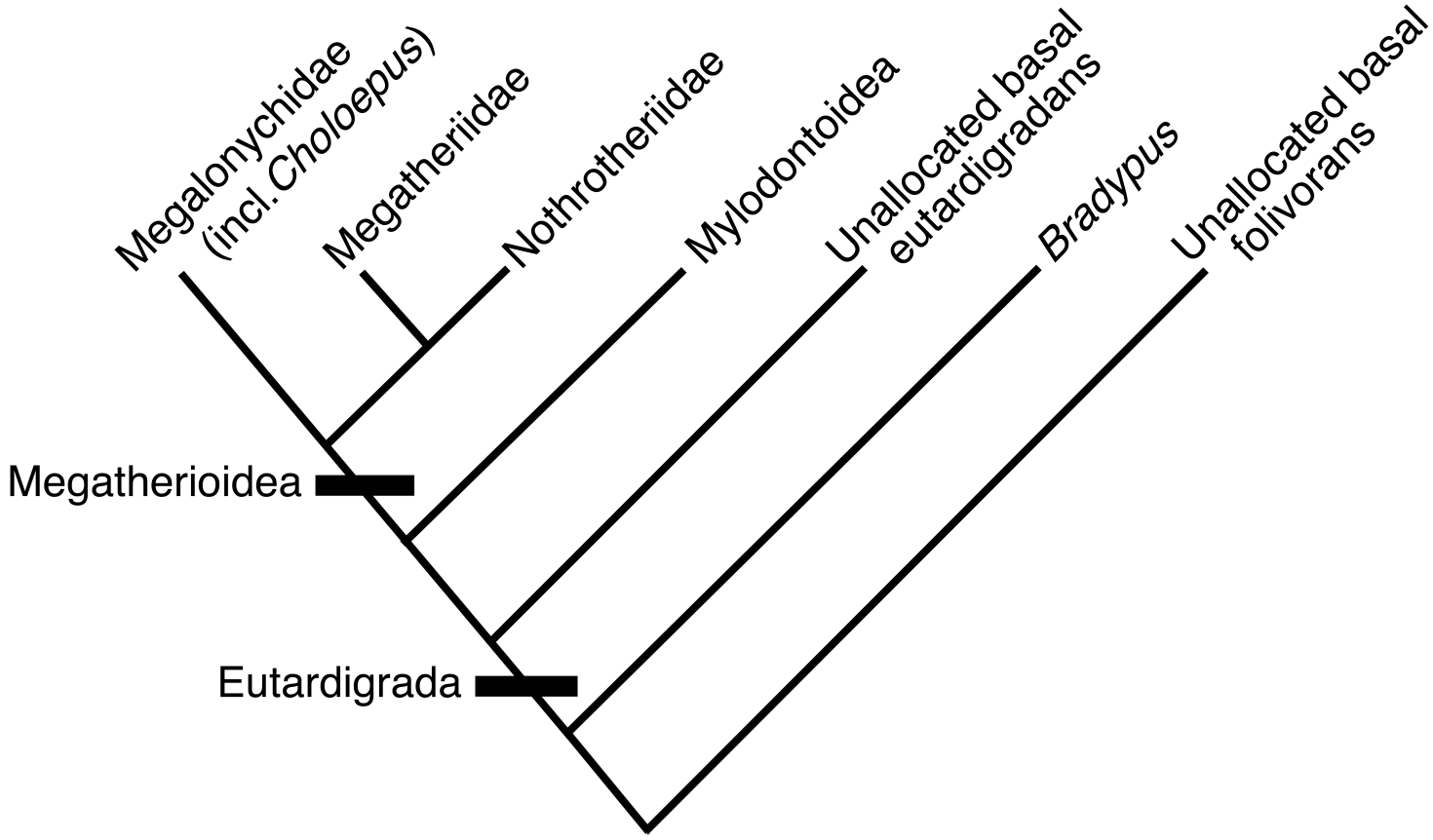
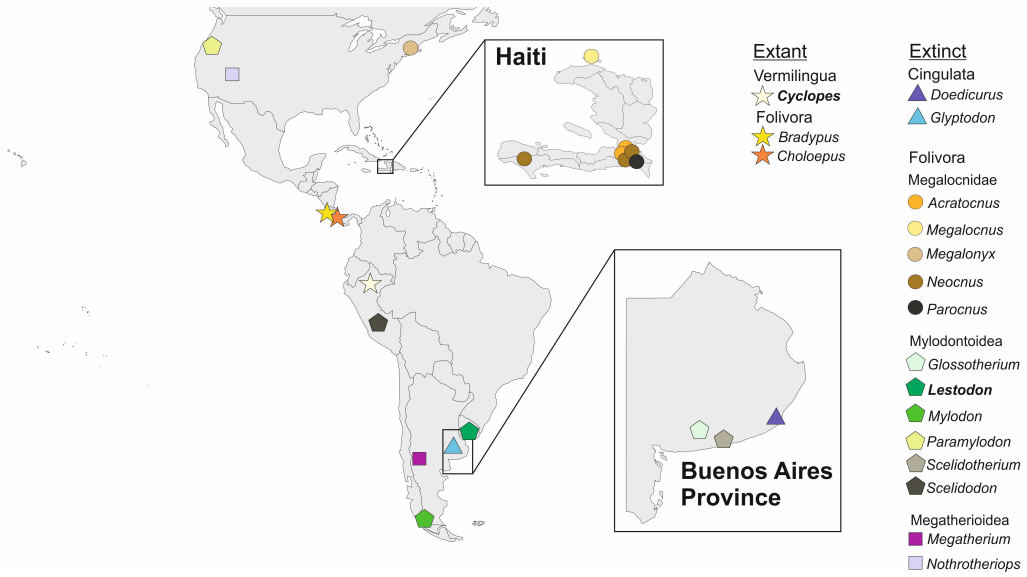

Megalonychidae
(incl. *Choloepus*)
Megatheriidae
Nothrotheriidae
Mylodontoidea
Unallocated basal
eutardigradans
*Bradypus*
Unallocated basal
folivorans
Megatherioidea
Eutardigrada

**Fig. 2**


Haiti
Buenos Aires Province
Extant
Vermilingua
Cyclopes
Folivora
Bradypus
Choloepus
Extinct
Cingulata
Doedicurus
Glyptodon
Folivora
Megalocnidae
Acratocnus
Megalocnus
Megalonyx
Neocnus
Parocnus
Mylodontoidea
Glossotherium
Lestodon
Mylodon
Paramylodon
Scelidotherium
Scelidodon
Megatherioidea
Megatherium
Nothrotheriops

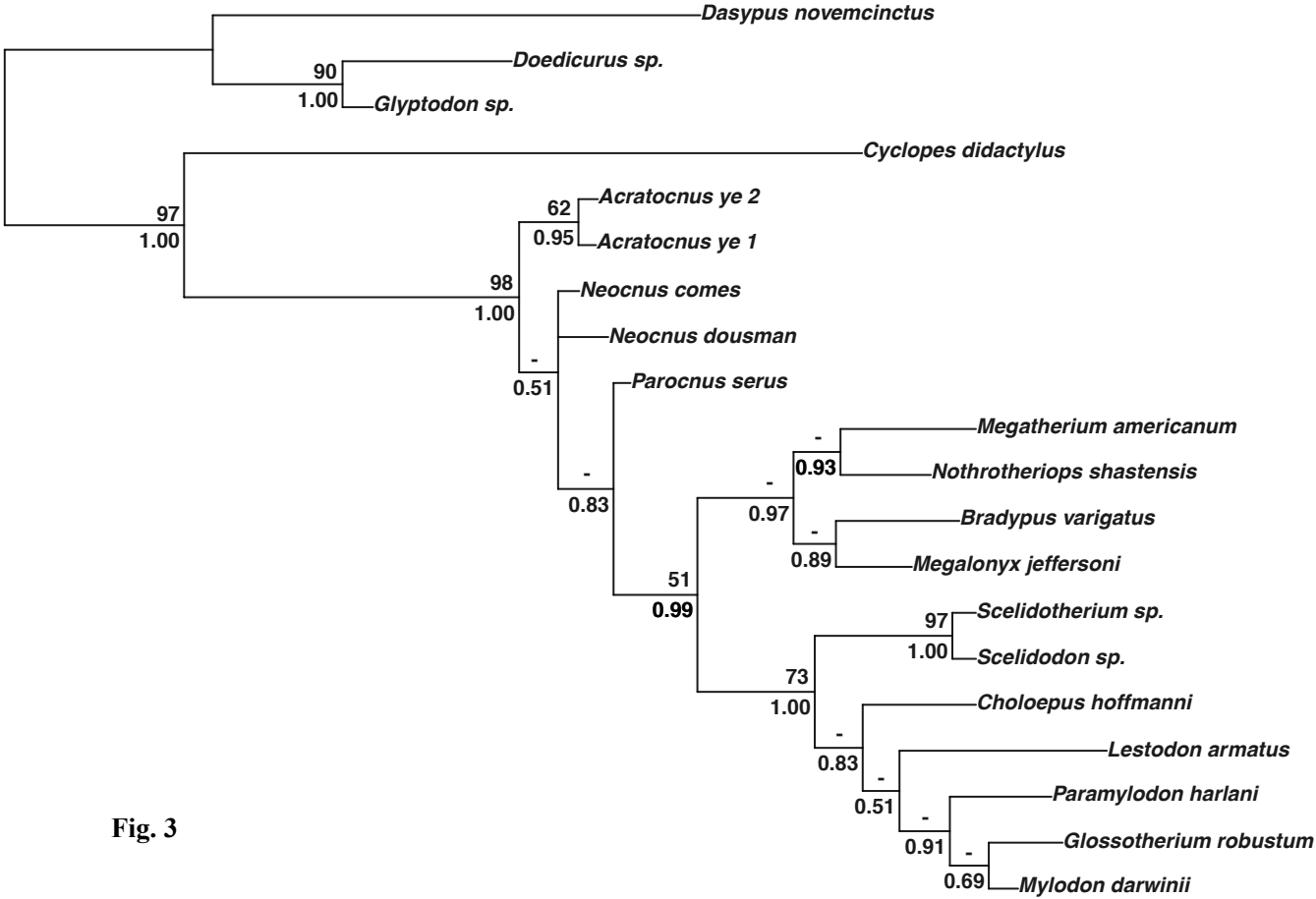

Dasypus novemcinctus
Doedicurus sp.
90
1.00
Glyptodon sp.
Cyclopes didactylus
97
1.00
Acratocnus ye 2
62
0.95
Acratocnus ye 1
98
1.00
Neocnus comes
Neocnus dousman
-
0.51
Parocnus serus
Megatherium americanum
-
0.93
Nothrotheriops shastensis
-
0.97
Bradypus varigatus
-
0.89
Megalonyx jeffersoni
51
0.99
Scelidotherium sp.
97
1.00
Scelidodon sp.
73
1.00
Choloepus hoffmanni
-
0.83
Lestodon armatus
-
0.51
Paramylodon harlani
-
0.91
Glossotherium robustum
-
0.69
Mylodon darwinii

**Fig. 3**

Fig. 4

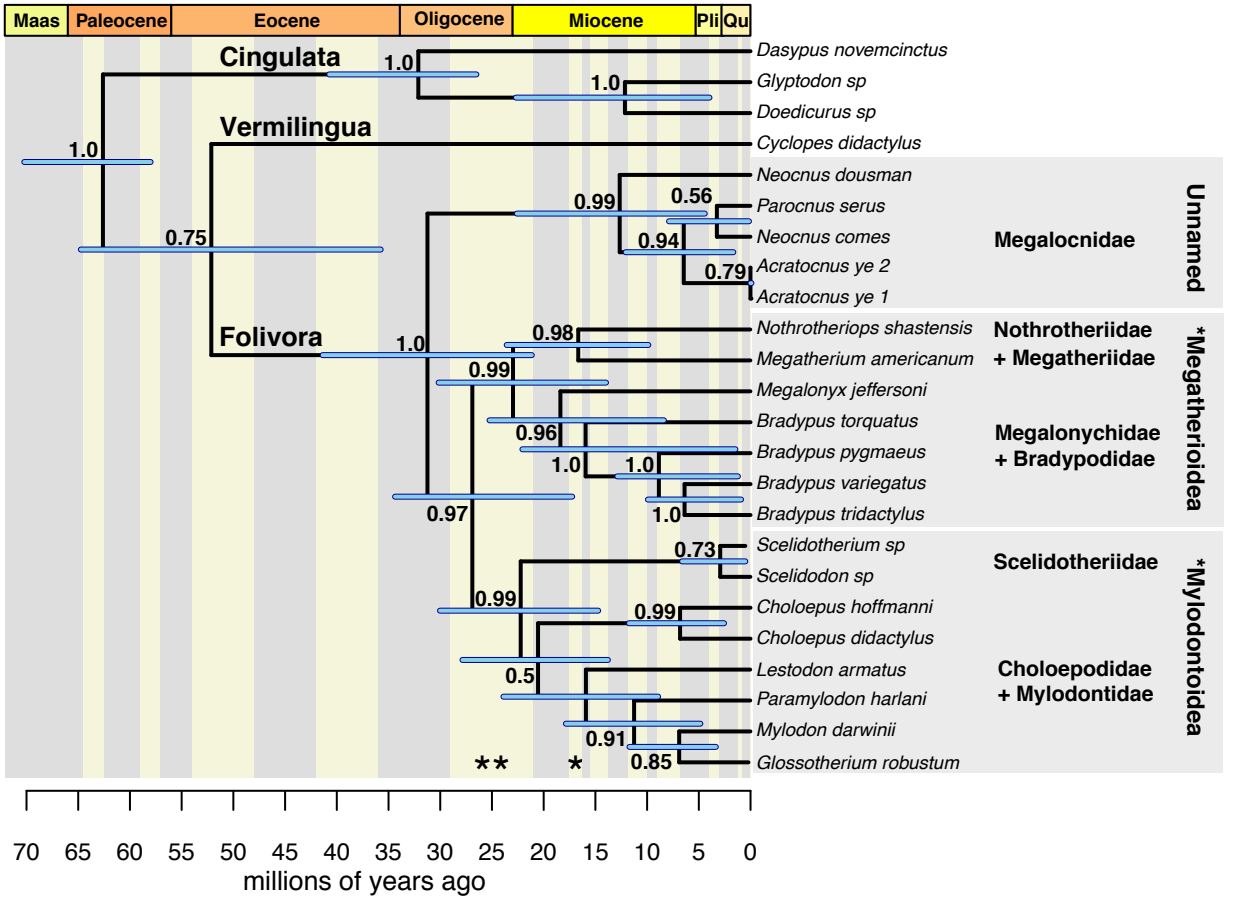

Maas
Paleocene
Eocene
Oligocene
Miocene
Pli
Qu
Cingulata
Vermilingua
Folivora
1.0
1.0
1.0
0.75
1.0
0.99
0.56
0.94
0.79
0.98
0.99
0.96
1.0
1.0
0.97
1.0
0.73
0.99
0.99
0.5
0.91
0.85
**
*
Dasypus novemcinctus
Glyptodon sp
Doedicurus sp
Cyclopes didactylus
Neocnus dousman
Parocnus serus
Neocnus comes
Acratocnus ye 2
Acratocnus ye 1
Nothrotheriops shastensis
Megatherium americanum
Megalonyx jeffersoni
Bradypus torquatus
Bradypus pygmaeus
Bradypus variegatus
Bradypus tridactylus
Scelidotherium sp
Scelidodon sp
Choloepus hoffmanni
Choloepus didactylus
Lestodon armatus
Paramylodon harlani
Mylodon darwinii
Glossotherium robustum
Megalocnidae
Nothrotheriidae + Megatheriidae
Megalonychidae + Bradypodidae
Scelidotheriidae
Choloepodidae + Mylodontidae
Unnamed
*Megatherioidea
*Mylodontoidea
70 65 60 55 50 45 40 35 30 25 20 15 10 5 0
millions of years ago

Table 1 Collagen peptides and per cent coverage

| Museum reference[1] | ID | Species | # Collagen Peptides | % Coverage |
|---|---|---|---|---|
| MMP 5672 | 15191 | *Doedicurus sp.* | 867 | 90 |
| MACN-PV 7 | 15194 | *Glyptodon sp.* | 731 | 84 |
| UF 76796 | 15559 | *Acratocnus ye* | 696 | 86 |
| UF 76385 | 15565 | *Acratocnus ye* | 629 | 87 |
| AMNH 20820 | 16265 | *Bradypus variegatus* | 793 | 88 |
| AMNH 139772 | 17009 | *Choloepus hoffmanni* | 1109 | 94 |
| MACN-PV 2652 | 15216 | *Glossotherium robustum* | 837 | 88 |
| UF 169931 | 15564 | *Megalocnus zile*[2] | 6 | 6 |
| NYSM VP-46 | 16849 | *Megalonyx jeffersonii*[3] | 874 | 85 |
| MAPBAR 3965 | 15225 | *Megatherium americanum* | 520 | 81 |
| UMAG ah 5854 | 16222 | *Mylodon darwinii* | 1371 | 96 |
| UF 171347 | 15548 | *Neocnus comes* | 699 | 84 |
| UF 170210 | 15780 | *Neocnus comes* | 591 | 84 |
| UF 75469 | 15781 | *Neocnus dousman* | 614 | 74 |
| USNM 244372 | 14723 | *Nothrotheriops shastensis* | 528 | 79 |
| USNM 3000 | 14715 | *Paramylodon harlani* | 642 | 87 |
| UF 75526 | 15556 | *Parocnus serus* | 575 | 82 |
| MUSM 1386 | 17480 | *Scelidodon sp.* | 1324 | 92 |
| MACN-PV 1791 | 15202 | *Scelidotherium sp.* | 475 | 76 |

[1]Institutional acronyms:

AMNH-M, American Museum of Natural History (Mammalogy), New York, USA

MACN-PV, Museo Argentino de Ciencias Naturales "Bernardino Rivadavia", Buenos Aires, Argentina

MAPBAR, Museo de la Asociación Paleontológica Bariloche, Bariloche, Argentina

MMP, Museo Municipal de Ciencias Naturales "Lorenzo Scaglia" Mar del Plata, Buenos Aires, Argentina

MUSM, Museo de Historia Natural, Universidad Nacional Mayor de San Marcos, Lima, Peru

NYSM VP, New York State Museum (Vertebrate Paleontology), Albany, New York, USA

UF, University of Florida, Natural History Museum of Florida, Gainesville, USA

UMAG ah, Instituto de La Patagonia, Universidad de Magallanes, Punta Arenas, Chile

USNM, United States National Museum of Natural History (Paleobiology), Washington DC, USA

[2]Mainly contaminants; not sequenced.

[3]SDS/PAGE protein extraction

Table 2: Selected divergence time estimates from BEAST analyses using different combinations of taxa and data (see Results and Supplementary Information). Note that, though consistently recovered as monophyletic, the position of Megalocnidae shifted among analyses, falling alternately as sister to all other Folivora (Xenarthra) or Megatherioidea (Folivora).

| Clade | Protein only | | mtDNA + Protein |
|---|---|---|---|
| | Xenarthra | Folivora | Xenarthra |
| Crown Xenarthra | 62.0 (57.6 - 62.8) | - | 62.6 (58.0 - 70.2) |
| Pilosa | 50.4 (37.4 - 62.8) | - | 52.1 (35.8 - 64.8) |
| Folivora | 26.4 (18.0 - 36.0) | 23.4 (14.9 - 33.9) | 31.2 (21.1-41.4) |
| Megalocnidae | 9.9 (3.8 - 17.8) | 7.7 (3.4 - 13.0 ) | 12.7 (4.4-22.6) |
| Megatherioidea + Megalocnidae | - | 19.4 (12.8 - 27.8) | - |
| Megatherioidea + Mylodontoidea | 22.7 (16.1 - 31.0) | - | 26.9 (17.2 - 34.4) |
| Megatherioidea | 15.7 (10.7 - 21.8) | 13.9 (9.4 - 19.4) | 23.0 (14.0 - 30.1 |
| *Megalonyx* + *Bradypus* | 11.1 (8.4 - 15.0) | 10.5 (8.4 - 14.1) | 18.4 (8.4 - 25.2) |
| *Bradypus* spp. | - | - | 16.0 (1.5 - 22.1) |
| *Megatherium* + *Nothrotheriops* | 12.3 (8.4 17.7) | 10.9 (7.8 - 15.1) | 16.7 (9.9 - 23.6) |
| Mylodontoidea | 15.3 (9.8 - 21.4) | 15.4 (8.9 - 23.4) | 22.2 (14.7 - 30.0) |
| *Choloepus* + Mylodontidae | 12.03 (7.3 - 17.2) | 10.5 (6.2 - 15.9) | 20.5 (13.8 - 27.9) |
| *Choloepus* spp. | - | - | 6.8 (2.6 - 11.8) |