UNIVERSITY *of* York

This is a repository copy of *Palaeoproteomics resolves sloth relationships*.

1 **Palaeoproteomics resolves sloth relationships**

2

3 Samantha Presslee[1,2,3,¶], Graham J. Slater[4,¶], François Pujos[5], Analía M. Forasiepi[5], Roman

4 Fischer[6], Kelly Molloy[7], Meaghan Mackie[3,8], Jesper V. Olsen[8], Alejandro Kramarz[9], Matías

5 Taglioretti[10], Fernando Scaglia[10], Maximiliano Lezcano[11], José Luis Lanata[11], John Southon[12],

6 Robert Feranec[13], Jonathan Bloch[14], Adam Hajduk[15], Fabiana M. Martin[16], Rodolfo Salas

7 Gismondi[17], Marcelo Reguero[18], Christian de Muizon[19], Alex Greenwood[20, 21], Brian T. Chait[7],

8 Kirsty Penkman[22], Matthew Collins[1,3], Ross D. E. MacPhee[2,*]

9

10 **Authors' Affiliations:**

11 [1]Department of Archaeology and BioArCh, University of York, Wentworth Way, Heslington,

12   York YO10 5NG, UK

13 [2]Department of Mammalogy, American Museum of Natural History, New York, NY 10024 USA

14 [3]Paleoproteomics Group, Natural History Museum of Denmark and University of Copenhagen

15   1353 Copenhagen, Denmark

16 [4]Department of the Geophysical Sciences, University of Chicago, 5734 S. Ellis Avenue,

17   Chicago, IL 60637 USA

18 [5]Instituto Argentino de Nivología, Glaciología y Ciencias Ambientales, CCT-CONICET-

19   Mendoza, Parque Gral San Martin, 5500 Mendoza, Argentina

20 [6]Target Discovery Institute, Nuffield Department of Medicine, University of Oxford, Roosevelt

21   Drive, Oxford OX3 7FZ, UK

22 [7]Chait Laboratory and National Resource for the Mass Spectrometric Analysis of Biological

23   Macromolecules, The Rockefeller University, 1230 York Avenue, New York, NY 10065 USA

24   [8]Novo Nordisk Foundation Center for Protein Research, Blegdamsvej 3B, 2200, Copenhagen,

25   Denmark

26   [9]Sección Paleovertebrados, Museo Argentino de Ciencias Naturales "Bernardino Rivadavia",

27   Av. Angel Gallardo 470, C1405DJR, Buenos Aires, Argentina.

28   [10]Museo Municipal de Ciencias Naturales "Lorenzo Scaglia", Plaza España s/n, 7600 Mar del

29   Plata, Argentina

30   [11]Instituto de Investigaciones en Diversidad Cultural y Procesos de Cambio, CONICET and

31   Universidad Nacional de Río Negro, 8400 Bariloche, Argentina

32   [12]Keck-CCAMS Group, Earth System Science Department, B321 Croul Hall, University of

33   California, Irvine, Irvine, CA 92697-3100, USA

34   [13]New York State Museum, Albany, NY 12230 USA

35   [14]Florida Museum of Natural History, University of Florida, Gainesville, FL 32611-7800 USA

36   [15]Museo de la Patagonia "F. P. Moreno", Bariloche 8400, Argentina

37   [16]Centro de Estudios del Hombre Austral, Instituto de la Patagonia, Universidad de Magallanes,

38   Punta Arenas, Chile

39   [17]BioGeoCiencias Lab, Facultad de Ciencias y Filosofía/CIDIS, Universidad Peruana Cayetano

40   Heredia, Lima, Perú

41   [18]CONICET and División Paleontología de Vertebrados, Museo de La Plata. Facultad de

42   Ciencias Naturales, Universidad Nacional de La Plata, Paseo del Bosque s/n, B1900FWA, La

43   Plata, Argentina

44   [19]Centre de Recherches sur la Paléobiodiversité et les Paléoenvironnements (CR2P), UMR

45   CNRS 7207, Muséum national d'Histoire naturelle, 8, rue Buffon 75231 Paris, France

46  [20]Leibniz Institute for Zoo and Wildlife Research, Alfred-Kowalke-Straße 17, 10315 Berlin,

47  Germany

48  [21]Department of Veterinary Medicine, Freie Universität Berlin, 14163 Berlin, Germany

49  [22]Department of Chemistry, University of York, Wentworth Way, Heslington, York YO10 5DD,

50  UK

51

52  ¶These authors contributed equally to this work.

53

54  *Author for correspondence: macphee@amnh.org

55

56  **Abstract**

57  The living tree sloths *Choloepus* and *Bradypus* are the only remaining members of Folivora, a

58  major xenarthran radiation that occupied a wide range of habitats in many parts of the western

59  hemisphere during the Cenozoic, including both continents and the West Indies. Ancient DNA

60  evidence has played only a minor role in folivoran systematics, as most sloths lived in places not

61  conducive to genomic preservation. Here we utilise collagen sequence information, both

62  separately and in combination with published mtDNA evidence, to assess the relationships of

63  tree sloths and their extinct relatives. Results from phylogenetic analysis of these datasets differ

64  substantially from morphology-based concepts: *Choloepus* groups with Mylodontidae, not

65  Megalonychidae; *Bradypus* and *Megalonyx* pair together as megatherioids, while monophyletic

66  Antillean sloths may be sister to all other folivorans. Divergence estimates are consistent with

67  fossil evidence for mid-Cenozoic presence of sloths in the West Indies and an early Miocene

68  radiation in South America.

69

70

71  The sloths (Xenarthra, Folivora), nowadays a taxonomically narrow (6 species in 2 genera)

72  component of the fauna of South and Central America[1,2] were once a highly successful clade of

73  placental mammals as measured by higher-level diversity (Fig. 1). Diverging sometime in the

74  Palaeogene from their closest relatives, the anteaters (Vermilingua), folivorans greatly expanded

75  their diversity and range, eventually reaching North America as well as the West Indies[3-8].

76  During the late Cenozoic sloth lineage diversity may have expanded and contracted several

77  times[9]. Final collapse occurred in the late Quaternary (end-Pleistocene on the continents, mid-

4

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

124 coded by only 2 genes, COL 1A1 and COL 1A2, only a small fraction of a species' genome can

125 be accessed with this probe. In the context of palaeontology, phylogenetic analyses of type 1

126 collagen have been shown to yield results that are highly congruent with those produced by

127 aDNA, especially at higher taxonomic levels[43,53].

128 One such application is testing morphology-based hypotheses of higher-level relationships where

129 there is a strong possibility that pervasive homoplasy among and between target groups has

130 affected morphological character analysis and therefore classification, as in the case of

131 incorrectly homologized caniniform tooth loci in living tree sloths[54]. Because dental features

132 have always played a large role in folivoran systematics,[7, 10, 12, 13 16, 31] such fundamental

133 reinterpretations are likely to have a significant impact. Clearly, it is desirable to use as many

134 sources of inference as possible in reconstructing phylogeny. Also, molecular data lend

135 themselves well to estimating divergence timing of major clades–another critical problem in

136 folivoran systematics[29, 34, 35].

137

138 **RESULTS**

139 To address some of the questions raised in the previous section, as well as to add to the available

140 molecular database for folivorans, we utilised proteomic data collected from fossil and living

141 sloths in order to focus on three fundamental issues: (1) relationships of tree sloths to each other

142 and to other folivorans; (2) composition of folivoran superfamilies Megatherioidea and

143 Mylodontoidea; and (3) divergence dating of major sloth ingroups. Results were tested against

144 datasets that additionally incorporated published genomic and phenomic information.

145   **Samples**. A total of 120 xenarthran samples comprising 24 different genus-level taxa (see

146   Supplementary Information, Table S1 and Fig. S1) were screened for protein survival using both

147   AAR (Amino Acid Racemization) and MALDI-ToF (Matrix-Assisted Laser

148   Desorption/Ionization Time-of-Flight) mass spectrometry. Three additional xenarthran sequences

149   were taken from the literature (see Methods, Proteomic Analysis). Of these, 34 or 28.3% of the

150   total number of samples (including 31.0% of 103 folivoran samples) produced promising results

151   for both AAR and MALDI-ToF MS. From these, the best sample per taxon was selected for LC-

152   MS/MS (Liquid Chromatography-Tandem Mass Spectrometry) analysis to derive protein

153   sequences, with some additions to maximize taxonomic coverage (Fig. 2, Table 1). We

154   resampled the specimen of *Megatherium* previously utilised by ref. 44; the results presented here

155   are *de novo*. The samples of *Neocnus dousman* and *Megalocnus zile* did not pass both MALDI-

156   ToF and AAR screening criteria, but it was decided to analyse them because they were the best

157   representatives of their species. However, because coverage for the *Megalocnus* sample was

158   particularly poor, recovered sequence being mostly contaminants, it was not used in the

159   phylogenetic analyses. To provide modern comparisons, samples of *Bradypus variegatus*

160   (AMNH 20820) and *Choloepus hoffmanni* (AMNH 139772) were also subjected to LC-MS/MS

161   analysis. For further details on all samples, see Supplemental Information, especially Table S1.

162   Relevant procedures for recovering sequence information and estimating phylogenetic

163   relationships are presented in Methods.


164   Samples ranged in assigned age from late Miocene to mid-Holocene (Supplementary

165   Information Table S1), but the 19 successfully-screened samples are all Quaternary (Table 1). Of

166   these, 15 were selected for radiocarbon dating, and 10 returned finite $^{14}$C ages (Supplementary

167   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    **Phylogenetic 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].

191    To further interrogate the reliability of our proteomic topologies, we concatenated our collagen

192    sequences with previously published mitochondrial genome sequences (hereafter, "proteomic +

193    genomic data") for all extant folivorans (2 species of *Choloepus,* 4 species of *Bradypus*), two

194    extinct folivorans (*Mylodon darwinii* and *Nothrotheriops shastensis*) and the two extant outgroup

195    taxa[34, 35]. Bayesian analysis (Fig. S3) of the combined dataset yielded a nearly identical topology

196    to that recovered using proteomic data alone, but in this instance *Megatherioidea (including

197    *Bradypus*) and *Mylodontoidea (including *Choloepus*) were unambiguously recovered as

198    reciprocally monophyletic clades (PP = 1). Recovery of a paraphyletic *Bradypus* (with respect to

199    *Megalonyx*) is almost certainly due to a long genomic branch and lack of proteomic data for *B.*

200    *torquatus,* combined with a comparable lack of genomic data for *Megalonyx.* As the monophyly

201    of *Bradypus* has never been questioned and this result is based exclusively on relative branch

202    lengths, we constrained *Bradypus* monophyly for subsequent analyses, though analyses without a

203    constraint were not noticeably different.

204    **Molecular clock considerations and divergence time estimates**. Incorporating time as an

205    analytical component in analysis of the combined dataset yielded a well-supported and

206    monophyletic Antillean clade (PP > 0.99), although within-clade relationships were not

207    satisfactorily resolved. More unexpectedly in light of traditional taxonomic concepts, BEAST

208    placed the Antillean clade as a well-supported sister to *Megatherioidea plus *Mylodontoidea

209    (PP =0.97) rather than pairing it with the one or the other. Support for megatherioid (PP > 0.99)

210    and mylodontoid (PP > 0.99) monophyly remained strong, but variable for constituent sub-

211    clades.

212    The relatively permissive constraints employed for calculating divergences make it difficult to

213    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:

237   *Choloepus* **is a mylodontoid.** That the two-toed sloth may be closer to traditional mylodontids

238   than to megalonychids, a possibility occasionally raised in morphological studies[16, 24, 57], has

239   been consistently found in recent aDNA investigations[34, 35, 37-40]. Due to the limited number of

240   extinct taxa included in those investigations the exact nature of their relationship has remained

241   indeterminate. However, the multiple tests of phylogenetic relationships and broad taxonomic

242   sampling used in the present study substantiates the conclusion that *Choloepus* is indeed a

243   mylodontoid.

244   Given the recent ages of all of the taxa investigated, coupled with low rates of sequence

245   evolution, it is unsurprising that divergence estimates based on proteins alone suggest an

246   early/middle Miocene origin for Scelidotheriidae + Mylodontidae (including *Choloepus*).

247   Inclusion of genomic data helps to push these estimates back to the earliest Miocene, but it

248   should be noted that a number of mylodontoid sloths of late Oligocene to late Miocene/early

249   Pliocene age do not fit neatly into better-defined clades. In the past, these taxa were occasionally

250   gathered[9, 20, 21] into the probably nonmonophyletic grouping Orophodontidae. It would be

251   interesting to know on the basis of molecular evidence whether the inclusion of a putative

252   orophodontid would affect the placement of *Choloepus*, possibly moving it stemward (Fig. S7)

253   or help refine divergence time estimates at the base of *Mylodontoidea. At present there is no

254   evidence on point; however, the youngest of these ambiguously-placed taxa, *Octodontobradys*, is

255   late Miocene/early Pliocene in age[58]—young enough to stand a chance of coming within the

256   range of proteomic methods as these continue to improve.

257   **Megalocnid sloths are monophyletic, and are not part of traditional Megalonychidae.**

258   Antillean sloths have had a complex taxonomic history[7]. In the past, this geographical grouping

259   of folivorans was sometimes regarded as diphyletic, with different island taxa having diverged

260     from different mainland antecedents[3, 7, 20, 58]. Diphyly now seems unlikely on the basis of our

261     molecular clock results (Fig. 4; see also Supplemental Information, Figs. S5, S6) as well as

262     recent morphology-based studies[16, 25]. Although within-clade relationships are poorly resolved

263     (cf. paraphyletic *Neocnus*), the Antillean clade as a whole resolves as strongly monophyletic (PP

264     >0.99). In light of this fact, as well as clade antiquity, it is appropriate to remove Megalocninae

265     from traditional Megalonychidae and raise it to family level (*Megalocnidae).

266     ***Megalonyx* and *Bradypus* are megatherioids.** Although recent morphology-oriented cladistic

267     studies have usually recovered *Bradypus* as sister to all other folivorans[8-10, 16], genomic

268     approaches[34,35, 39] have consistently paired the three-toed sloth with the extinct North American

269     Pleistocene megatherioid *Nothrotheriops*. On this point the proteomic data presented here are

270     fully compliant with the genomic evidence and support rejection of the inference[9, 16] that

271     Bradypodoidea (i.e., *Bradypus*) is sister to traditional Megatherioidea + Mylodontoidea, as tested

272     by both parsimony (13 additional steps) and Bayesian inference (2*lnBayes Factor = 6.72,

273     support = Strong). Equally controversial is the sister group relationship detected between

274     *Bradypus* + *Megalonyx* (PP = 0.89 - 0.98; Fig. 4; see Supplementary Information, Fig. S3).

275     Although well supported in analyses of both collagen-only and combined proteomic + genomic

276     data, this remains a surprising finding, inasmuch as such an association has never been reported

277     in any taxon-rich phylogenetic study emphasizing morphology. While both the three-toed sloth

278     and *Megalonyx* are likely to be megatherioids cladistically, settling their deeper relationships will

279     require substantially more data than is currently available.

280     That none of the Antillean sloths used in this study showed any proteomic affinity for *Megalonyx*

281     is also surprising, because much of what has been understood to morphologically characterize

282     non-South American Megalonychidae was based on Antillean species, the fossils of which tend

283    to be far more complete than those of most other taxa conventionally included in this family[12, 16,

284    17]. To resolve this conflict, additional high-quality data will be required, genomic and proteomic

285    as well as phenomic. The only certainty at present is that, if *Choloepus* is excluded,

286    Megalonychidae must now be relegated to the list of formerly diverse but now completely

287    extinct folivoran families.

288    **The West Indies may have been colonized early.** An early appearance of megalocnid sloths in

289    the West Indies has been proposed on general palaeobiogeographical grounds[3, 10, 17, 24, 60], but at

290    present the only pre-Quaternary fossil evidence for Antillean folivorans consists of a

291    morphologically inconclusive partial femur from the early Oligocene (~31 Ma) Yauco Formation

292    of Puerto Rico[61] and unassociated remains attributable to a folivoran, *Imagocnus zazae*, from the

293    late early Miocene (~17.5 Ma) Lagunitas Formation of Cuba[5]. Although "megalonychid"

294    affinities have been assumed for both on biogeographical grounds, now no longer applicable,

295    neither has been included in formal phylogenetic analyses and their placement within Folivora

296    remains uncertain.

297    The presence of sloths in the West Indies at least as early as the early Miocene is congruent with

298    our mean age estimate (31.2 Ma; Fig. 4, Table 2) for the last common ancestor of sloths sampled

299    in this study. This inference is also roughly consistent with the GAARlandia dispersal

300    hypothesis[5, 62], which holds that northwestern South America and the Greater Antilles were

301    briefly in land connection during the Eocene-Oligocene transition. Without going beyond the

302    very slim body of molecular evidence currently available, there is now at least some basis for

303    hypothesizing that *Megalocnidae might represent an *in situ* Antillean radiation that was

304    emplaced on the islands during the earliest phases of the evolution of the folivoran crown-

305    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

15

329    instrumentation and applicable software, the future for palaeoproteomics should be bright if it

330    can continue to make significant contributions to solving difficult questions like the ones

331    explored here.

332    A new aDNA study[87] of folivoran phylogeny, published as this paper was going to press, reaches

333    conclusions almost identical to ours regarding the evolutionary relationships of living tree sloths

334    and the phylogenetic distinctiveness of the West Indian radiation. Because the taxonomic

335    distribution of sampled species is not identical in the two studies, there are some minor

336    differences in lower-level relationships and estimated divergence times. However, their detailed

337    agreement overall supports the argument that high-quality protein sequence information is a

338    reliable source of evidence for reconstructing phylogenetic relationships.

339

340

341    **METHODS**

342    **Proteomic Analyses**

343    The 5-number codes following taxon names in this section refer to lab sample ID numbers

344    referenced in Table 1.

345    **AAR.** Samples were prepared using a slightly modified version of the protocol in ref. 66. A

346    small sub-sample of bone (~1 mg) was hydrolysed in 7M HCL (100 µl per mg) under $N_2$ for 18

347    hours at $110^{o}$C. After hydrolysis, the samples were dried down overnight before being re-

348    hydrated in 0.01mM L-homo-arginine as an internal standard. The samples were analysed using

349    reversed phase high pressure liquid chromatography (RP-HPLC) following a slightly modified

16

350    version of the protocol developed by ref. 67. Amino acid composition and extent of racemization

351    was used to assess promising samples for sequencing.

352    **Sample preparation for MS.** The majority of samples (see Supplementary Information, Table

353    S1) were prepared using a slightly modified version of the ZooMS protocol for bone reported by

354    ref. 43. Bone samples (15-30 mg) were demineralized in 250 µl 0.6M HCl for a minimum of 3

355    weeks at -20°C. This allowed for a gentler demineralization and helped to protect any remaining

356    collagen. After demineralization, the samples were rinsed once in 200 µl 0.01M NaOH, and

357    three times in 200 µl 50mM ammonium bicarbonate (Ambic). The samples were gelatinized by

358    being resuspended in 100 µl 50mM Ambic and heated at 65°C for 1 hour before being digested

359    overnight at 37°C; 50 µl of the heated sample was digested using 1 µl of 0.5 µg/µl porcine

360    trypsin in trypsin resuspension buffer (Promega, UK) and the other 50 µl was dried down and

361    resuspended in 50 µl 100mM Tris solution to be digested with elastase (Worthington; USA) at

362    the same concentration in 10% Tris solution. Two different enzymes were used to increase the

363    protein sequence coverage for LC-MS/MS[43, 68]. Digestion was stopped by the addition of

364    trifluoroacetic acid (TFA) at a concentration of 0.5-1% of the total solution. Peptides were

365    desalted using zip-tips[64] and eluted in 100 µl of 50% acetonitrile (ACN)/0.1% TFA (v/v).

366    **SDS-PAGE.** Selected samples were analysed using SDS PAGE (Table 1). This method was used

367    on certain samples as the standard ZooMS protocol had not yielded positive results on certain

368    samples that were deemed potentially important phylogenetically. Bone samples were crushed to

369    ~1 um sized particles using a Retsch PM100 ball mill cooled with liquid nitrogen. The ball mill

370    was cleaned with distilled water and methanol before and after each sample[69]. Nanoscale

371    crushing allowed for the highest potential retrieval of proteomic information. 50 mg of powdered

372    sample was heated at 70°C for 10 minutes in 200 µl SDS solubilizing buffer (0.5M Tris base, 5%

373     SDS, 130mM DTT), cysteines were alkylated by the addition of 6 µl 1M IAA at room

374     temperature in the dark for 30 minutes before the addition of 200 µl of dye solution (0.05%

375     bromophenol blue, 5% glycerol). 20 µl of the samples were run on a Bis-Tris gel (NuPAGE) for

376     10 minutes to concentrate the samples into a gel plug which was briefly washed in a fixing

377     solution (16% methanol, 10% acetic acid), before being washed twice in boiling water. The gel

378     was stained using Coomassie stain.

379     The gel plug was cut into approximately 1mm sized cubes in a fume hood with a scalpel and the

380     gel cubes for each sample placed in a separate Eppendorf. The gel pieces were washed in a de-

381     staining solution (66% ammonium bicarbonate 33% acetonitrile) until no more dye could be seen

382     before being washed in the following solvents for 10 minutes per solvent; ACN, HPLC grade

383     water, ACN and 50mM ammonium bicarbonate[70]. The samples were digested overnight with 100

384     µl 3.125 µg/µl trypsin in 50mM ammonium bicarbonate at 37$^{\circ}$C and then the tryptic digest was

385     pipetted into a cleaned Eppendorf tube. 100 µl of 70%ACN/1.7% formic acid/0.1% TFA was

386     added to the gel pieces and the gel was heated at 37$^{\circ}$C for 1 hour with the supernatant being

387     collected and added to the tryptic digest. This step was repeated sequentially with 100mM

388     triethyl ammonium bicarbonate (TEAB) and ACN. The extracted peptides were dried down and

389     then resuspended in 5% Formic acid/0.1% TFA desalted and purified on C18 membranes

390     (Empore) before being eluted in 80% ACN/0.5% acetic acid. The purified peptides were spun to

391     dryness ready for LC-MS/MS analysis.

392     **MALDI-ToF MS.** 1 µl of sample was spotted in triplicate onto an MTP384 Bruker ground steel

393     MALDI target plate. 1 µl of α-cyano-4-hydroxycinnamic acid matrix solution (1% in 50%

394     Acetonitrile/0.1% Trifluoroacetic acid (v/v/v)) was added to each sample spot and mixed with

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

396    in triplicate.

397    **LC-MS/MS.** Most samples were analysed at the Discovery Proteomic Facility (DPF) at Oxford

398    (Table 1). *Choloepus* 17009 and *Mylodon* 16222 were analysed at the Novo Nordisk Foundation

399    Centre for Protein Research (NNFCPR), University of Copenhagen. The *Megalonyx* sample (ID

400    16849) was run at the Laboratory of Mass Spectrometry and Gaseous Ion Chemistry, Rockefeller

401    University.

402    At DPF, sample batches were analysed on an Orbitrap Fusion Lumos or Q-Exactive with

403    identical front-end separation, employing an Easyspray column (ES803, 500mmx75µm,

404    Thermo) and a gradient of 2%-35% ACN in 0.1% FA/5%DMSO over 60 minutes. On the Fusion

405    Lumos, MS1 resolution was set to 120,000 with an AGC target of 400,000. MS2 spectra were

406    acquired in TopSpeed mode (3 seconds duty cycle) in the linear ion trap (rapid scan mode) for up

407    to 250ms, with an AGC target of 4,000 and fragmentation in CID mode (35% normalized

408    collision energy). The MS1 resolution on the Q-Exactive was set to 70,000 with an AGC target

409    of 3E6. MS2 spectra for up to 15 precursors were acquired with a resolution of 17,500 and an

410    AGC target of 1E5 for up to 128ms and 28% normalized collision energy (higher-energy

411    collision dissociation). On both instruments, precursors were excluded for 27 seconds from re-

412    selection.

413    At NNFCPR, dried peptides were resuspended in 50µl of 80% ACN and 0.1% formic acid before

414    being transferred to a 96 well plate and placed in a vacuum centrifuge at 40ºC until

415    approximately 3 µL of solution was left. The samples were rehydrated with 5 or 10 µL (*Mylodon*

416    16222 and *Choloepus* 17009 respectively) of 0.1% TFA, 5% ACN. Samples were separated on a

417   15 cm column (75 μm inner diameter) in-house laser pulled and packed with 1.9 μm C18 beads

418   (Dr. Maisch, Germany) on an EASY-nLC 1000 (Proxeon, Odense, Denmark) connected to a Q-

419   Exactive HF (Thermo Scientific, Bremen, Germany) on a 77 min gradient. 5 μl of sample was

420   injected. Buffer A was milliQ water. The peptides were separated with increasing buffer B (80%

421   ACN and 0.1% formic acid), going from 5% to 80% over an 80 minute gradient and a flow rate

422   of 250 nL/min. In addition, a wash-blank injecting 2μl 0.1% TFA, 5% ACN was run in-between

423   each sample to hinder cross-contamination.

424   The Q-Exactive HF was operated in data dependent top 10 mode. Full scan mass spectra (350-

425   1400 m/z) were recorded at a resolution of 120,000 at m/z 200 with a target value of 3e6 and a

426   maximum injection time of 25 ms for *Choloepus* 17009 and 45ms for *Mylodon* 16222. Fragment

427   ions were recorded with a maximum ion injection time set to 108 ms and a target value set to 2e5

428   and recorded at a resolution of 60,000 for *Choloepus* 17009 and 30,000 for *Mylodon* 16222.

429   Normalized collision energy was set at 28% and the isolation window was 1.2 m/z with the

430   dynamic exclusion set to 20 s.

431   At Rockefeller University, peptides were resuspended in 20 uL 5% methanol, 0.2% formic acid.

432   10 uL were loaded onto an EASY-Spray column (Thermo Fisher Scientific ES800: 15 cm × 75

433   μm ID, PepMap C18, 3 μm) via an EASY-nLC 1200 and separated over a 120 minute gradient of

434   2-32% Solvent B (Solvent A = 0.1% formic acid in water, Solvent B = 0.1% formic acid, 95%

435   acetonitrile) during online ESI–MS and MS/MS analyses with a Q Exactive Plus mass

436   spectrometer (Thermo Fisher Scientific). MS/MS analyses of the top 25 precursors in each full

437   scan (300 to 1700 m/z) used the following parameters: resolution: 17,500 (at 200 Th); AGC

438   target: $2 \times 10^5$; maximum injection time: 200ms; isolation width: 2.0 m/z; normalized collision

439   energy: 24%.

440 **Protein sequence analysis.** The LC-MS/MS raw files were converted to MGF files using

441 Proteowizard[71] and searched against a mammal collagen database which included common

442 contaminants (http://www.thegpm.org/crap/) in PEAKS v7.5. Mass tolerances were set at 0.5Da

443 for the fragment ions and 10ppm for precursor ions and up to 3 missed cleavages were permitted.

444 Searches allowed various post translational modifications (PTMs) including oxidation (MHW)

445 and hydroxylation of proline (both +15.99), deamidation (NQ; +0.98) and pyro-glu from E (-

446 18.01) as well as a fixed PTM of carbamidomethylation (+57.02) which occurs as part of the

447 sample preparation. A maximum of 3 PTMs were allowed per peptide. Protein tolerances were

448 set at 0.5% false discovery rate (FDR), >50% average local confidence (ALC; *de novo* only) and

449 -10lgP score $\geq$ 20.

450 Sequences of both COL 1A1 and COL 1A2 were concatenated using previously published

451 mammal collagen consensus sequences taken from NCBI, including sequences for the

452 xenarthrans *Dasypus novemcinctus* (nine-banded armadillo; GenBank: XP_004470764),

453 *Cyclopes didactylus* (silky anteater; Uniprot: COHJP1/COHJP2), and *Lestodon armatus* (extinct

454 mylodontoid sloth, ref. 44). Telopeptides very rarely survive in fossil samples and so these were

455 removed from all sequences. Isoleucine and leucine cannot be differentiated using low energy

456 tandem mass spectrometry and *de novo* sequencing as both amino acids are isobaric. Therefore,

457 the identification of leucine/isoleucine was consistent throughout the sequence analyses

458 concatenated in this study. Our approach is in line with previous phylogenetic studies using

459 collagen as probe[43], under the assumption that MS/MS sequence variation was not interpreted as

460 significant phylogenetic change (see below, Phylogenetic Analyses).

461 Once a potential collagen sequence was compiled for a given sloth taxon, the sequence was

462 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

23

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, $p \ll 0.001$), 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., *Choloepus* 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

532    by the fossil record. We therefore also performed a series of Bayesian tree searches assuming a

533    molecular clock under the fossilized birth-death framework[80-82], as implemented in BEAST

534    v2.5.1[83]. Briefly, this framework allowed us to sample from the posterior distribution of time-

535    scaled trees for taxa in our proteomic dataset, inferred using their sequences and stratigraphic

536    ages, while using phylogenetically constrained fossil taxa that lack amino acid data to provide

537    additional information on relative branch lengths and divergence times. Our choice of fossil taxa

538    and topological constraints broadly followed the approach undertaken in ref. 34 for sloth

539    mitogenomes. However, our proteomic topologies raise questions about the phylogenetic

540    positioning of some fossil folivorans that have previously been considered on morphological

541    grounds as early representatives of Pleistocene and Holocene families. For example, some

542    extinct folivorans, such as the Huayquerian nothrotheriid *Mionothropus*[84] can be plausibly

543    assigned to a specific terminal branch in our proteomic topology. Others, however, are

544    customarily assigned to clades that we failed to recover. This applies to the Santacrucian taxon

545    *Eucholoeops,* usually interpreted as a basal megalonychid[24, 85, 86] and therefore as a member of a

546    clade not found to be monophyletic in our analyses. Such issues inevitably affect efforts to

547    calibrate the proteomic + genomic data clock and to infer divergence times. Acknowledging this,

548    we employed a minimal set of constraints (see Supplemental information, Fig. S4) on the

549    positioning of fossil folivorans in our Bayesian estimation of topology and divergence times,

550    integrating over all possible placements of phylogenetically uncertain fossils using stratigraphic

551    context alone when necessary. We performed analyses with and without a monophyly constrain

552    on *Bradypus* and results did not differ at unaffected nodes.

553    The use of a Bayesian approach requires the specification of prior probabilities on model

554    parameters. We used default priors on substitution model parameters but specified the following:

555 net diversification ~ Exp(1), yielding a broad, vague prior; turnover ~ beta(2,1), yielding high

556 prior weight on extinction ≅ speciation; sampling probability ~ beta(2,2) yielding a humped

557 distribution that placed most prior weight on sampling probabilities of 0.5; origin ~ U(61.5, 150)

558 yielding a flat prior on ages older than 61.5 Ma up to 150 Ma. In addition, the analysis was

559 conditioned on the number of extant taxa sampled ($\rho = 0.129$ in the xenarthran proteomic

560 analyses, $\rho = 0.333$ in the folivoran proteomic analyses, $\rho = 0.266$ in the combined analyses).

561 Based on comparisons of marginal likelihoods computed via Path Sampling (see Supplementary

562 Information, Table S3), we employed a relaxed uncorrelated clock with log-normally distributed

563 rates for proteomic and combined analyses, with an exponential prior (mean=0.1) placed on the

564 mean of log-normal distribution and the default gamma $\Gamma(0.5396, 0.3819)$ on the standard

565 deviation. Two MCMC analysis were run for 10 million generations each, sampling every 1000

566 generations, after which fossils without data were pruned from the trees, the first 20% of the

567 retained samples were discarded as burn-in, the samples combined, and maximum clade

568 credibility trees constructed using the tree annotator software accompanying the BEAST suite.

569 Runs from the prior using a fixed topology (the maximum clade credibility tree based on the pre-

570 pruning sample) were used to confirm that divergence time estimates were not simply returning

571 the prior.

572 **Data availability** Mass spectrometry proteomics data have been deposited to the

573 ProteomeXchange Consortium via the PRIDE partner repository with the dataset

574 identifier PXD012859. Collagen sequences are available on the Uniprot website

575 (https://www.uniprot.org/); the complete list can be found in Supplemental Information, Table

576 S5. Phylogenetic datasets have been deposited on DataDryad (doi:10.5061/dryad.7dd64gs).

577

578    **Supplementary Information**

579    [URL tk]

580

581    **Acknowledgements**

604

605    **Author contributions**

606    RDEM, MC, and SP conceived the project. SP undertook AAR and proteomic analysis and

607    concatenated collagen sequences, with laboratory and technical assistance from RF, JO, KM,

608    MM, MC, KP, and BC. GJS conducted phylogenetic analyses. FP and AMF supplied

609    paleontological information. AK, MT, FS, ML, AH, AI, RF, JB, JLL, FMM, RSG, MR, AG,

610    CdM, and GB supplied fossil samples, locality information, species identifications, and

611    commentary on the manuscript. RDEM, SP, and GJS wrote the manuscript, with input from all

612    authors.

613

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

615

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

## REFERENCES CITED

[1] Gardner, A. L. in *Mammals of South America*, **1** (ed Gardner, A. L.) 157-176 (University of Chicago Press, 2007).

[2] Nowak, R. *Walker's Mammals of the World: Monotremes, Marsupials, Afrotherians, Xenarthrans, and Sundatherians* (Johns Hopkins, 2018).

[3] Kraglievich, L. Descripción de dos cráneos y otros restos del género ''*Pliomorphus*'' Ameghino, procedentes de la formación entrerriana de las barrancas del río Paraná. *Anal. Mus. Nac. Hist. Nat. Buenos Aires* **33**, 1-56 (1923).

[4] Hoffstetter, R. in *Traité de Paléontologie*, **6.2** (ed Piveteau, J.) 535-636 (Masson, 1958).

[5] MacPhee, R. D. E. & Iturralde-Vinent, M. A. Origin of the Greater Antillean land mammal fauna 1: New Tertiary land mammals from Cuba and Puerto Rico. *Amer. Mus. Novitates* **3141**, 1-31 (1995). http://hdl.handle.net/2246/3657

[6] Iturralde-Vinent, M. A. & MacPhee, R. D. E. Paleogeography of the Caribbean region: implications for Cenozoic biogeography. *Bull. Amer. Mus. Nat. Hist.* **238**, 1-95 (1999). http://hdl.handle.net/2246/1642

[7] 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, 2001).

640

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).

643

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

647

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

651

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

654

12 Engelmann, G. F. in *The Evolution and Ecology of Armadillos, Sloths, and Vermilinguas* (ed

Montgomery, G. G.) 195-203 (Academic Press, 1985).

657

13 Webb, S. D. in *The Evolution and Ecology of Armadillos, Sloths, and Vermilinguas* (ed

Montgomery, G. G.) 105-112 (Smithsonian Institution Press, 1985).

660

661     [14] White, J. Indicators of locomotor habits in xenarthrans: evidence of locomotor heterogeneity

662     among fossil sloths. *J. Vert. Paleontol.* **13** (2), 230-242 (1993).

663     https://www.jstor.org/stable/4523502

664

665     [15] Delsuc, F., Catzeflis, F. M., Stanhope, M. J. & Douzery, E. J. P. 2001. The evolution of

666     armadillos, anteaters and sloths depicted by nuclear and mitochondrial phylogenies: implications

667     for the status of the enigmatic fossil *Eurotamandua*. *Proc. Roy. Soc. B* **268**, 1605-1615 (2001).

668     doi: 10.1098/rspb.2001.1702

669

670     [16] Gaudin, T. J. Phylogenetic relationships among sloths (Mammalia, Xenarthra, Tardigrada): the

671     craniodental evidence. *Zool. J. Linn. Soc.* **140**, 255-305 (2004). doi: 10.1111/j.1096-

672     3642.2003.00100.x

673

674     [17] McDonald, H. G. & De Iuliis, G. in *The Biology of Xenarthra* (eds Vizcaino, S. F. & Loughry,

675     W. J.) 39-55 (University Press of Florida, 2008).

676

677     [18] Pujos, F., Gaudin, T. J., De Iuliis, G. & Cartelle, C. Recent advances on variability, morpho-

678     functional adaptations, dental terminology, and evolution of sloths. *J. Mamm. Evol.* **19**, 159-169

679     (2012). doi: 10.1007/s10914-012-9189-y

680

681     [19] Nyakatura, J. A. The convergent evolution of suspensory posture and locomotion in tree sloths.

682     *J. Mamm. Evol.* **19** (3), 225-234 (2012). doi: 10.1007/s10914-011-9174-x

683

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

728    [31] Boscaini, A., Gaudin, T. J., Mamani Quispe, B., Antoine, P.-O. & Pujos, F. New well-

729    preserved craniodental remains of *Simomylodon uccasamamensis* (Xenarthra, Mylodontidae)

730    from the Pliocene of the Bolivian Altiplano: phylogenetic, chronostratigraphic and

731    paleobiogeographic implications. *Zool. J. Linnean Soc*. **185** (2), 459-486 (2019). doi:

732    10.1093/zoolinnean/zly075

733

734    [32] McDonald, H. G. & De Iuliis, G. in *The Biology of Xenarthra* (eds Vizcaino, S. F. & Loughry,

735    W. J.) 39-55 (University Press of Florida, 2008).

736

737    [33] Delsuc, F. & Douzery, E. J. P. in *The Biology of Xenarthra* (eds Vizcaino, S. F. & Loughry,

738    W. J.) 11-23 (University Press of Florida, 2008).

739

740    [34] Slater, G., Cui, P., Forasiepi, A. M., Lenz, D., Tsangaras, K., Voirin, B., de Moraes, N.,

741    MacPhee R. D. E. & Greenwood, A. D. Evolutionary relationships among extinct and extant

742    sloths: the evidence of mitogenomes and retroviruses. *Genome Biol. Evol.* **8** (3), 607-621 (2016).

743    doi: 10.1093/gbe/evw023

744

745    [35] Delsuc, F., Kuch, M., Gibb, G. C., Hughes, J., Szpak, P., Southon, J., Enk, J., Duggan, A. T. &

746    Poinar, H. N. Resolving the phylogenetic position of Darwin's extinct ground sloth (*Mylodon*

747    *darwinii*) using mitogenomic and nuclear exon data. *Proc. Roy. Soc. B* **285**, 20180214 (2018).

748    doi: 10.1098/rspb.2018.0214

749

750 [36] Moraes-Barros, N., Silva, J. A. & Morgante, J. S. Morphology, molecular phylogeny, and

751 taxonomic inconsistencies in the study of *Bradypus* sloths (Pilosa: Bradypodidae). *J. Mammal.*

752 **92** (1), 86-100 (2011). doi: 10.1644/10-MAMM-A-086.1

753

754 [37] Höss, M., Dilling, A., Currant, A. & Pääbo, S. Molecular phylogeny of the extinct ground

755 sloth *Mylodon darwinii. Proc. Nat. Acad. Sci. USA* **93** (1), 181-185 (1996). PMID: 8552600

756

757 [38] Poinar, H. N., *et al.* (1998). Molecular coproscopy: Dung and diet of the extinct ground sloth

758 *Nothrotheriops shastensis. Science* **281** (5375), 402-406. PMID: 9665881

759

760 [39] Greenwood, A. D., Castresana, J., Feldmaier-Fuchs, G. & Pääbo, S. A molecular phylogeny of

761 two extinct sloths. *Mol. Phylogen. Evol.* **18** (1), 94-103 (2001). doi: 10.1006/mpev.2000.0860

762

763 [40] Hofreiter, M., Betancourt, J.L., Sbriller A. P., Markgraf, V. & McDonald, H.G. Phylogeny,

764 diet, and habitat of an extinct ground sloth from Cuchillo Cura, Neuquen Province, southwest

765 Argentina. *Quat. Res.* **59** (3): 364-378 (2003). doi: 10.1016/S0033-5894(03)00030-9

766

767 [41] Welker, F., Smith, G. M., Hutson, J. M., Kindler, L., Garcia-Moreno, A., Villaluenga, A.,

768 Turner, E. & Gaudzinski-Windheuser, S. Middle Pleistocene protein sequences from the

769 rhinoceros genus *Stephanorhinus* and the phylogeny of extant and extinct Middle/Late

770 Pleistocene Rhinocerotidae. *PeerJ* **5**, e3033 (2017). doi: 10.7717/peerj.3033

771

772 [42] Welker, F., *et al.* Palaeoproteomic evidence identifies archaic hominins associated with the

773 Châtelperronian at the Grotte du Renne. *Proc. Nat. Acad. Sci. USA* **113** (40), 11162-11167

774 (2016). doi:org/10.1073/pnas.1605834113

775

776 [43] Welker, F., *et al.* Ancient proteins resolve the evolutionary history of Darwin's South-

777 American ungulates. *Nature* **522**, 81-84 (2015). doi: 10.1038/nature14249

778

779 [44] Buckley, M., Fariña, R. A., Lawless, C., Tambusso, P. S., Varela, L., Carlini, A. A., Powell, J.

780 E. & Martinez, J. G. Collagen sequence analysis of the extinct giant ground sloths *Lestodon* and

781 *Megatherium. PloS One* **10** (12), e0144793 (2015). doi: 10.1371/journal.pone.0144793

782

783 [45] Dobberstein, R. C., Collins, M. J., Craig, O. E., Taylor, G., Penkman, K. E. H. & Ritz-Timme,

784 S. Archaeological collagen: why worry about collagen diagenesis? *Archaeol. Anthropol. Sci.* **1**

785 (1), 31-42 (2009). doi: 10.1007/s12520-009-0002-7

786

787 [46] Buckley, M. & Collins, M. J. Collagen survival and its use for species identification in

788 Holocene-Lower Pleistocene bone fragments from British archaeological and palaeontological

789 sites. *Antiqua* **1** (1), e1 (2011). doi: 10.4081/antiqua.2011.e1

790

791 [47] Buckley, M. & Wadsworth, C. Proteome degradation in ancient bone: diagenesis and

792 phylogenetic potential. *Palaeogeog. Palaeoclimat. Palaeoecol.* **416**, 69-79 (2014). doi:

793 10.1016/j.palaeo.2014.06.026

794

795 [48] Rybczynski, N., Gosse, J. C., Harington, C. R., Wogelius, R. A., Hidy, A. J. & Buckley, M.

796 Mid-Pliocene warm-period deposits in the High Arctic yield insight into camel evolution. *Nat.*

797 *Comm.* **4**, 1550 (2013). doi:10.1038/ncomms2516

798

799 [49] Allentoft, M. E., *et al.* The half-life of DNA in bone: measuring decay kinetics in 158 dated

800 fossils. *Proc. Roy. Soc. B* **279** (1748), 4724–4733 (2012). doi: 10.1098/rspb.2012.1745

801

802 [50] Orlando, L., *et al.* Recalibrating *Equus* evolution using the genome sequence of an early

803 Middle Pleistocene horse. *Nature* **499** (7456), 74-78 (2013). doi: 10.1038/nature12323

804

805 [51] Presslee, S., *et al.* Radiocarbon dating and proteomic analysis of highly purified bone collagen

806 derived from Rancho La Brea mammal fossils. *J. Vert. Paleont.* Program & Abstracts 2016, 208

807 (2016).

808

809 [52] Tuross, N. & Stathoplos, L. in *Methods in Enzymology* **224** (ed Zimmer A., White, T. J., Cann,

810 R. L. & Wilson, A. C.) 121-129 (Academic Press, 1993).

811

812 [53] Westbury, M., *et al.* A mitogenomic timetree for Darwin's enigmatic "transitional" South

813 American mammal, *Macrauchenia patachonica*. *Nat. Commun.* **8**, 15951 (2017) doi:

814 10.1038/ncomms15951.

815

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

839  [61] MacPhee, R. D. E., Iturralde-Vinent, M. A. & Gaffney, E. S. Domo de Zaza: An Early

840  Miocene vertebrate locality in south-central Cuba, with notes on the tectonic evolution of Puerto

841  Rico and Mona Passage. *Amer. Mus. Novitates* **3394**, 1-42 (2003).

842  http://hdl.handle.net/2246/2820

843

844  [62] Tong, Y. F., Binford, G., Rheims, C. A., Kuntner, M., Liu, J., Agnarsson, I. Huntsmen of the

845  Caribbean: Multiple tests of the GAARlandia hypothesis. *Mol. Phylogenet. Evol.* **130**, 259-268

846  (2019). doi:10.1016/j.ympev.2018.09.017

847

848  [63] Steadman, D. W., Martin, P. S., MacPhee, R. D. E., Jull, A. J. T., McDonald, H. G., Woods, C.

849  A., Iturralde-Vinent, M. A. & Hodgins, G. W. L. Asynchronous extinction of late Quaternary

850  sloths on continents and islands. *Proc. Nat. Acad. Sci. USA* **102** (33), 11763-11768 (2005). doi:

851  10.1073/pnas.0502777102

852

853  [64] Hoorn, C., *et al.* Amazonia through time: Andean uplift, climate change, landscape evolution,

854  and biodiversity. *Science* **330** (6006), 927-931 (2010). doi: 10.1126/science.1194585

855

856  [65] Tejada-Lara, J. V., *et al.* Life in proto-Amazonia: Middle Miocene mammals from the

857  Fitzcarrald Arch (Peruvian Amazonia). *Palaeontology* **58** (2), 341-378 (2015). doi:

858  10.1111/pala.12147

859

860    [66] Penkman, K., Kaufman, D. S., Maddy, D. & Collins, M. J. Closed-system behaviour of the

861    intra-crystalline fraction of amino acids in mollusc shells. *Quat. Geochronol.* **3** (1-2), 2-25

862    (2008). doi: 10.1016/j.quageo.2007.07.001.

863

864    [67] Kaufman, D. S. & Manley, W. F. A new procedure for determining DL amino acid ratios in

865    fossils using reverse phase liquid chromatography. *Quat. Sci. Rev.* **17** (11), 987-1000 (1998). doi:

866    10.1016/S0277-3791(97)00086-3

867

868    [68] Demarchi, B., *et al.* Protein sequences bound to mineral surfaces persist into deep time. *eLife*

869    **5,** e17092 (2016). doi: 10.7554/eLife.17092

870

871    [69] Kontopoulos, I., Presslee, S., Penkman, K. & Collins, M. J. Preparation of bone powder for

872    FTIR-ATR analysis: the particle size effect. *Vibrational Spectroscopy* **99**, 167–177 (2018). doi:

873    10.1016/j.vibspec.2018.09.004

874

875    [70] Van Doorn, N. L., Hollund, H. & Collins, M. J. A novel and non-destructive approach for

876    ZooMS analysis: Ammonium bicarbonate buffer extraction. *Archaeol. Anthropol. Sci.* **3** (3), 281-

877    289 (2011). doi: 10.1007/s12520-011-0067-y

878

879    [71] Chambers, M. C., *et al.* A cross-platform toolkit for mass spectrometry and proteomics. *Nature*

880    *Biotechnol.* **30**, 918-920 (2012). doi: 10.1038/nbt.2377

881

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

929

930 [85] De Iuliis, G., Pujos, F., Toledo, N., Bargo, M. S. & Vizcaíno, S. F. *Eucholoeops* Ameghino,

931 1887 (Xenarthra, Tardigrada, Megalonychidae) from the Santa Cruz Formation, Argentine

932 Patagonia: implications for the systematics of Santacrucian sloths. *Geodiversitas* **36** (2), 209-255

933 (2014). doi: 10.5252/g2014n2a2

934
935 [86] Hirschfeld, S. E. & Webb, S. D. Plio-Pleistocene megalonychid sloths of North America. *Bull.*

936 *Florida Mus. Nat. Hist. Biol. Sci.* **12** (5), 213-294 (1968).

937 [87] Delsuc, F. *et al.* Ancient mitogenomics rewrites the evolutionary history and biogeography of

938 sloths. *Current Biology*. [tk]

939


940
941 **FIGURE AND TABLE LEGENDS (MAIN TEXT)**

942 **Figure 1:** Phylogenetic relationships among major folivoran taxa based on morphological

943 evidence (mostly after ref. 8, 16), with existence of unallocated taxa acknowledged. In this

944 framework, the three-toed tree sloth *Bradypus* is sister to other sloths (grouped here as

945 Eutardigrada), while the two-toed tree sloth *Choloepus* is included within Megalonychidae.

946

947 **Figure 2:** Geographical locations of sequenced samples. Sequences for *Cyclopes* and *Lestodon*

948 (**in bold**) taken from the literature; others, this paper (Table 1 and Supplementary Information,

949 Fig. S1).

950

**Figure 3:** 50% majority rule consensus tree from Bayesian analysis of the proteomic data

without temporal information, as performed in MrBayes. Values below nodes are posterior

probabilities for the descendant clade (see Results). Values above nodes are bootstrap support

derived from 10,000 bootstrap replicates. A dash (–) indicates that a node was not represented in

the 50% majority rule bootstrap consensus. Extant *Dasypus* and extinct *Doedicurus* and

*Glyptodon* are members of the order Cingulata; extant *Cyclopes* is a representative of

Vermilingua, which together with Folivora comprise order Pilosa. Cingulates and pilosans

together comprise superorder Xenarthra (see also Fig. 4).


**Figure 4:** Time scaled maximum clade credibility tree from BEAST analysis of 24 extant and

extinct xenarthran collagen sequences plus published mitochondrial genomes (see text). Branch

lengths are the mean values from the retained posterior sample, while blue bars represent 95%

highest posterior density intervals. Values at nodes are posterior probabilities (note that the

monophyly of *Bradypus* is constrained here). Vertical shaded bars correspond to South American

land mammal ages (SALMAs), two of which are emphasized: Deseadan (\*\*), 29–21 Ma, during

which the first generally-accepted representatives of traditional Megatherioidea and

Mylodontoidea appear paleontologically; and the Santacrucian (\*), 17.5–16.3 Ma, the SALMA

during which mylodontids maintained substantial taxonomic diversity but megalonychids and

megatheriids declined[9]. On the right (grey boxes), folivoran species used in analyses are

associated with their traditional family names, but with superfamily contents organized

according to phylogenetic conclusions in text. Megalocnidae is placed outside traditional

superfamily structure in its own (unnamed) box. The tree implies that the fundamental split

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).
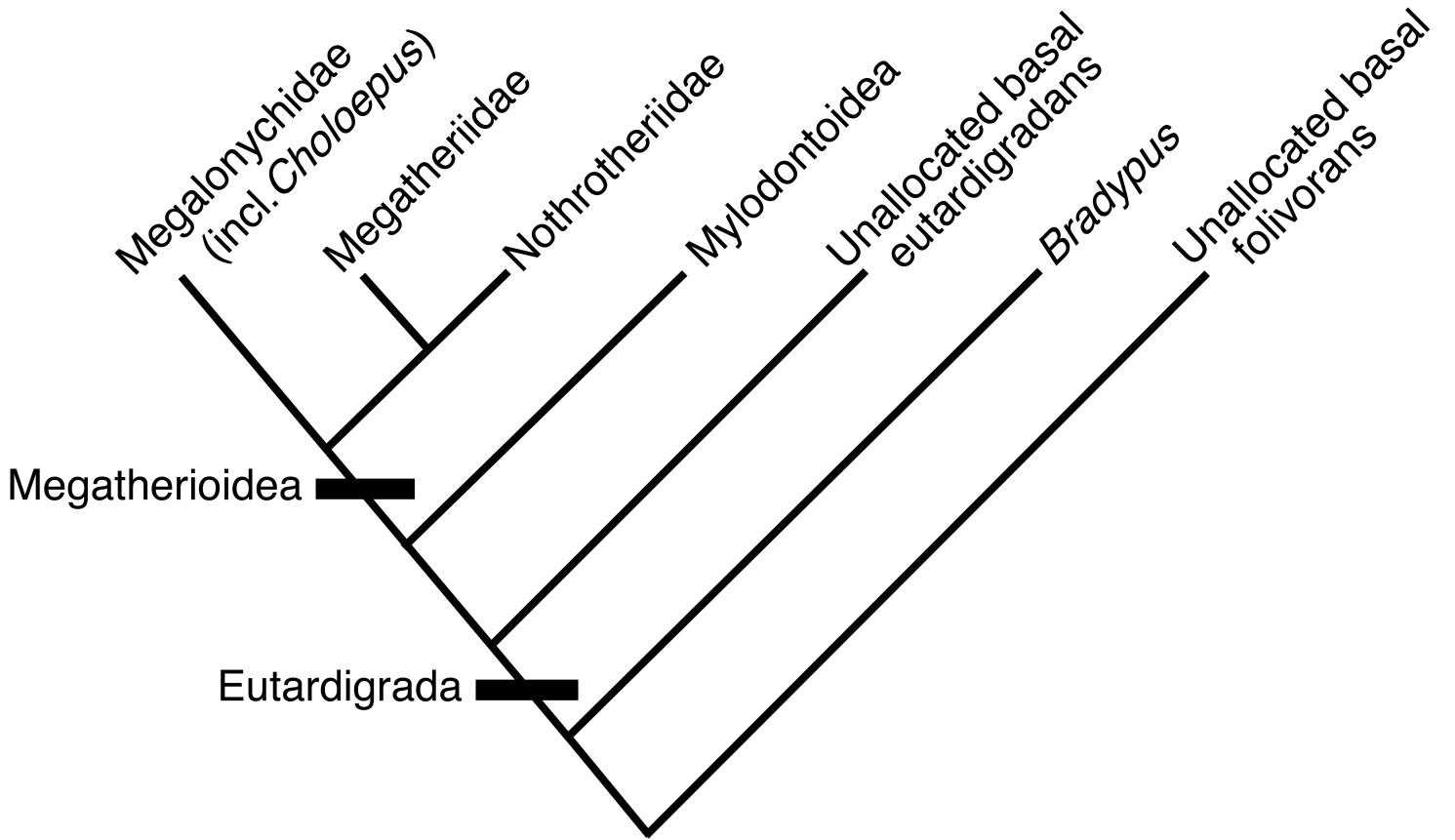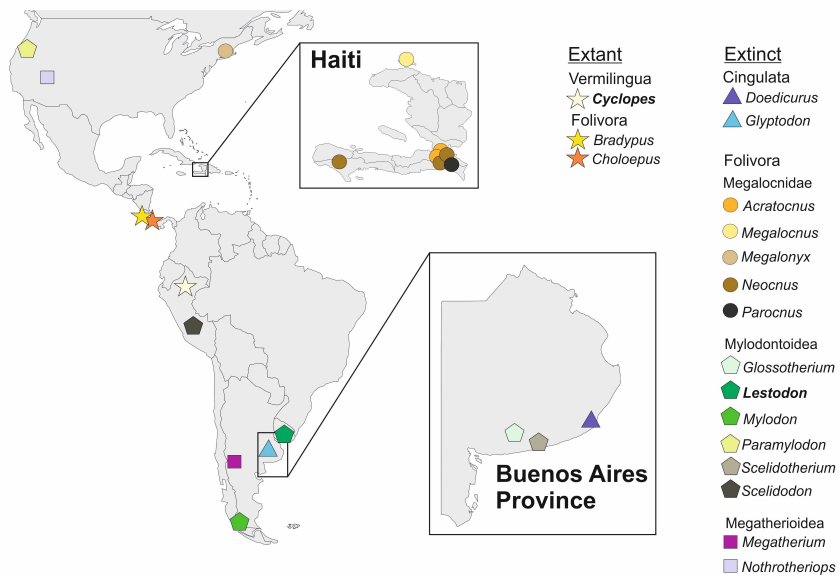
983

Fig. 1

**Fig. 2**



**Haiti**

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*

**Buenos Aires Province**
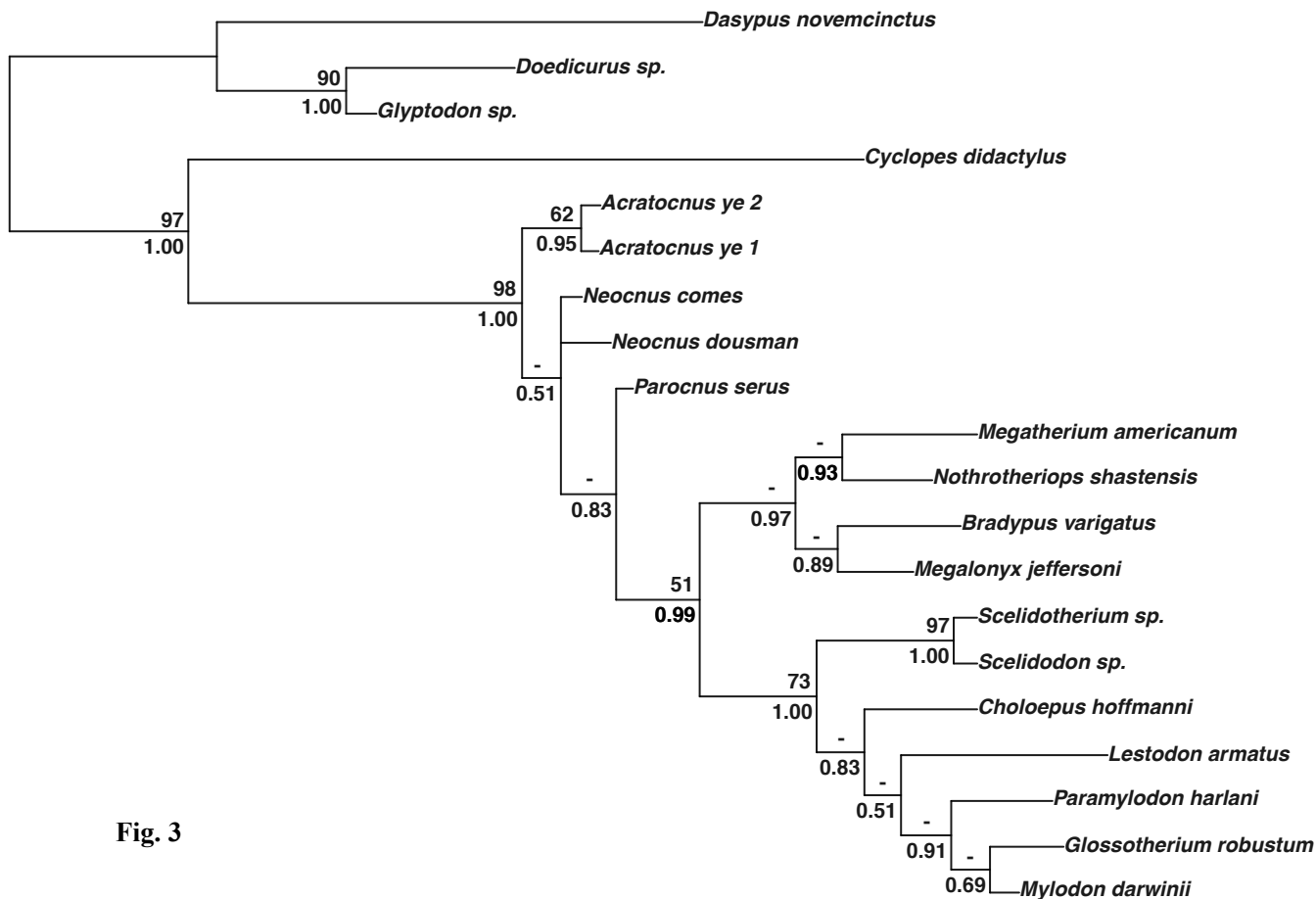
**Fig. 3**

**Fig. 4**
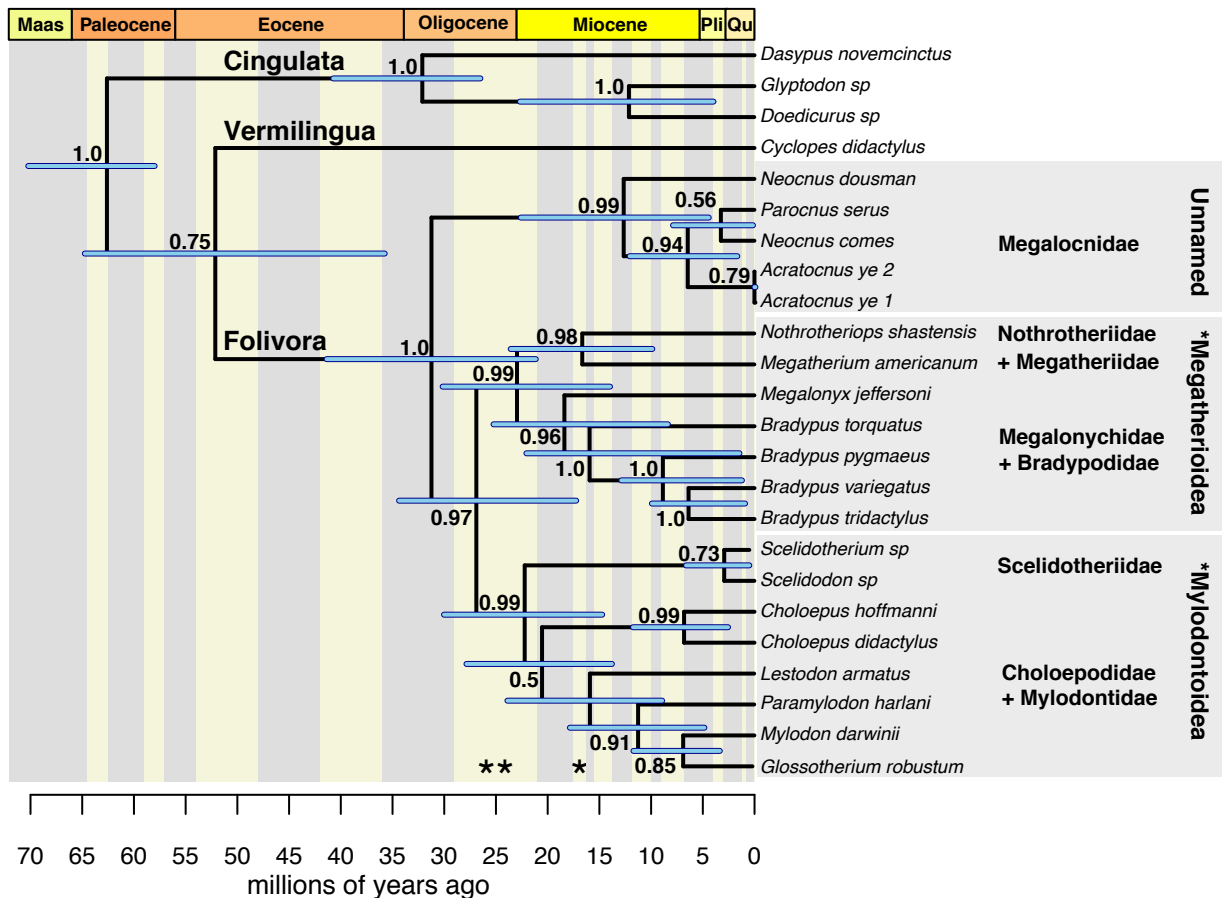
| Museum reference[1] | ID | Species | # Collagen Peptides | % Coverage |
|---|---|---|---|---|
| MMP 5672 | 15191 | *Doedicurus sp.* | 867 | 90 |
| MACN-PV 7 | 15194 | *Glyptodon sp.* | 731 | 84 |
| UF 76796 | 15559 | *Acratocnus ye* | 696 | 86 |
| UF 76385 | 15565 | *Acratocnus ye* | 629 | 87 |
| AMNH 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 |

Table 1 Collagen peptides and per cent coverage

[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) |