Pseudogenes Document Protracted Parallel Regression of Oral Anatomy in Myrmecophagous Mammals

Christopher A. Emerling[1,2,*], Sophie Teullet[2], Rémi Allio[2,3], John Gatesy[4], Mark S. Springer[5] and Frédéric Delsuc[2,*]

[1] Biology Department, Reedley College, Reedley, CA, USA
[2] ISEM, Univ. Montpellier, CNRS, IRD, Montpellier, France
[3] CBGP, INRAE, CIRAD, IRD, Montpellier SupAgro, Univ. Montpellier, Montpellier, France
[4] Division of Vertebrate Zoology, American Museum of Natural History, New York, NY, USA
[5] Department of Evolution, Ecology, and Organismal Biology, University of California, Riverside, CA, USA

***Correspondence to** christopher.emerling@reedleycollege.edu and frederic.delsuc@umontpellier.fr

ABSTRACT

Adaptation to ant and/or termite consumption (myrmecophagy) in mammals constitutes a textbook example of convergent evolution, being independently derived in several mammalian lineages. Myrmecophagous species are characterized by striking convergent morphological adaptations such as skull elongation, enlargement of salivary glands, and long claws to dig into ant and termite nests. These evolutionary modifications also include anatomical regression, such as dental simplification or loss, reduction of masticatory muscles, and possessing a reduced set of taste buds. To gain insights into the molecular changes underlying the regression of these morpho-anatomical traits, we investigated the functionality of candidate genes related to dentition, gustation, and mastication in nine convergent myrmecophagous mammalian lineages. We examined these genes in a comparative phylogenetic context, paired with molecular evolutionary analyses, to estimate the relative timing of loss of gene function over the evolutionary history of each convergent lineage. We found that gustatory reduction and pseudogenization of masticatory myosin often were associated with the regression of dental genes. Evidence of pseudogenization events linked to oral anatomy dates to as early as the Cretaceous/Paleogene boundary, and is an ongoing process including examples of incipient gene inactivations. Whereas we found evidence for gene inactivations across all three functional

categories occurring during distinct temporal intervals, there was variation in the sets of genes lost and the relative timing of inactivation events. The combined evidence suggests that the convergent evolution of myrmecophagy has occurred as a protracted process with distinct phases of anatomical evolution, over timescales as long as 60 Myr.

INTRODUCTION

Convergent evolution is a widespread phenomenon in which distantly-related organisms independently derive highly similar phenotypic traits. Such a pattern-based definition of evolutionary convergence does not make any assumptions on the underlying mechanisms and allows one to test if similar traits evolved as a result of similar selection pressures (Stayton 2015). Deriving similar anatomy, physiology, and behavior can indeed result from modifications to identical, unique or partially overlapping sets of genes, depending on the trait and its underlying molecular basis. The frequency with which convergent evolution is demonstrably repeatable at the genetic level, particularly among highly divergent taxa, is a topic of ongoing interest (Blount et al. 2018; Cerca 2023).

One prominent example of convergent evolution can be found in myrmecophagous mammals, species that specialize in eating ants and/or termites (termitivory), which constitute more than 90% of their diet. Mammalian myrmecophages are phylogenetically widespread (Redford 1987; Reiss 2001), with some of the most specialized myrmecophages being found among the xenarthran anteaters (Vermilingua), pangolins (Pholidota), the aardvark (Tubulidentata, *Orycteropus afer*), tolypeutine armadillos (Cingulata, Tolypeutinae: three-banded [*Tolypeutes* spp.], giant [*Priodontes maximus*], and naked-tailed [*Cabassous* spp.] armadillos), the aardwolves (Carnivora: *Proteles* spp.), the numbat (Dasyuromorphia, *Myrmecobius fasciatus*), and the short-beaked echidna (Monotremata, *Tachyglossus aculeatus*). Other mammals with at least partially myrmecophagous diets include include fairy (Chlamyphorinae), hairy (Euphractinae) and long-nosed (Dasypodidae) armadillos, as well as some carnivorans, such as the bat-eared fox (Canidae*, Otocyon megalotis*) and sloth bear (Ursidae, *Melursus ursinus*) (Redford 1987; Reiss 2001).

As these lineages have independently adapted to consuming social insects, there has been a shift from a more typical ancestral mammalian masticatory apparatus (specialized multicuspid teeth, robust jaws and associated muscles) to a characteristically derived myrmecophagous oral apparatus: dental simplification and/or loss, thinning of the mandible, a reduced or absent zygomatic arch, an elongated tongue that can rapidly probe ant and termite

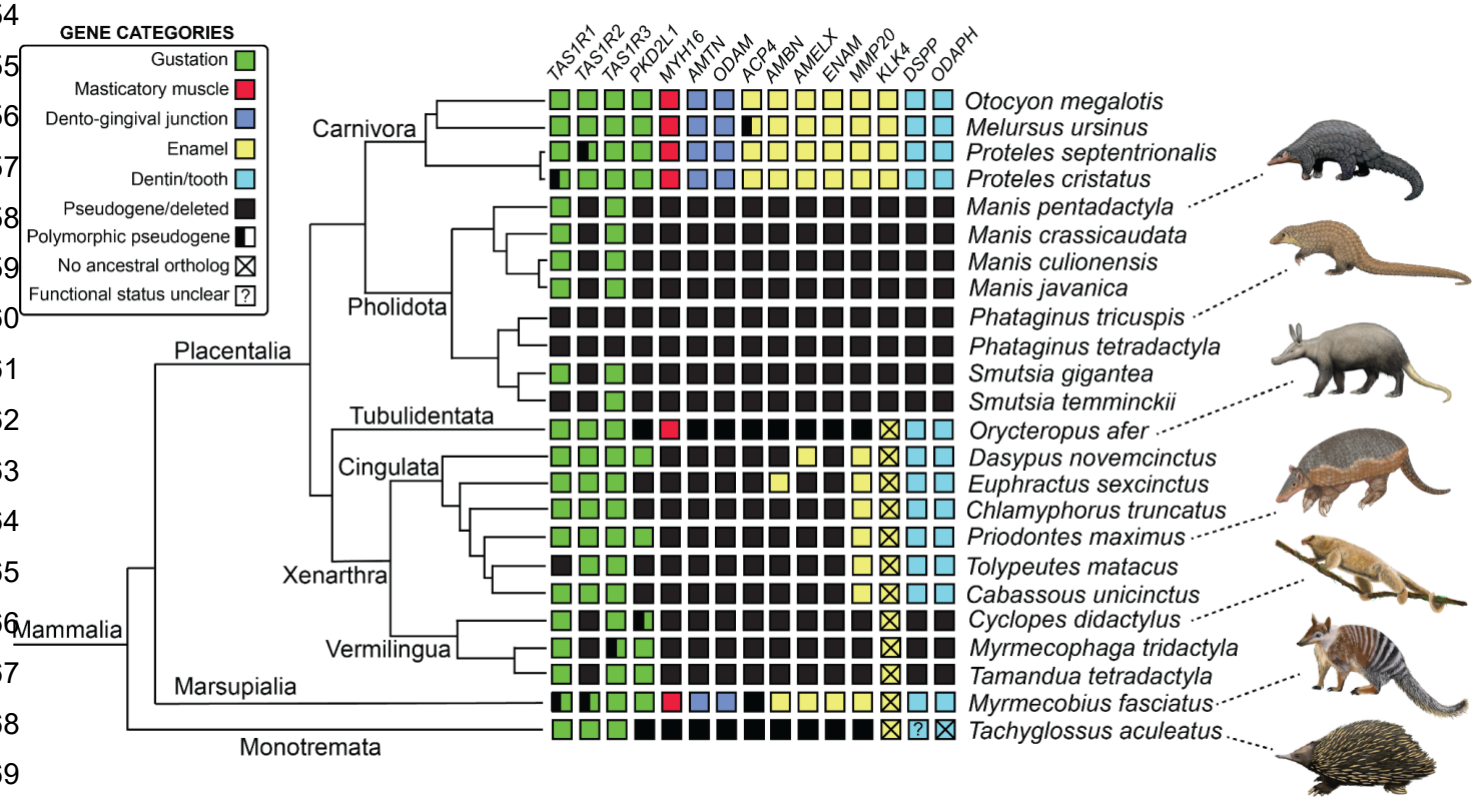
68 nests, enlarged salivary glands which produce a high volume of viscous saliva to capture prey,
69 and the reorganization and reduction of jaw musculature to accommodate this feeding
70 mechanism (Reiss 2001; Ferreira-Cardoso et al. 2019; Ferreira-Cardoso et al. 2020). These
71 morphological characteristics are exemplified in xenarthran anteaters ("anteaters" throughout
72 text) and pangolins, wherein extreme specialization for myrmecophagy and accompanying
73 anatomical modifications had led systematists to formerly conclude that they constitute a
74 monophyletic group (Edentata) (Glass 1985; Reiss 2001). It was not until molecular
75 phylogenetics was applied to this question towards the turn of the 21st century that it was firmly
76 demonstrated that their shared features were the result of convergent evolution (Murphy et al.
77 2001; Delsuc et al. 2002; Springer et al. 2013).
78 Given that convergent evolution of myrmecophagy can lead to the complete
79 reorganization of the oral apparatus, it raises the question of whether this is detectable at the
80 genomic level. One way to test for convergent molecular evolution is by studying the
81 phenomenon of regressive evolution, or vestigialization, whereby phenotypic traits are reduced
82 (e.g., in size or complexity) or completely lost over evolutionary time (Fong et al. 1995; Lahti et
83 al. 2009; Albalat and Cañestro 2016). While adaptations for myrmecophagy frequently include
84 the enlargement or extension of certain traits (e.g., hypertrophied salivary glands, lengthening of
85 tongue), several others can be classified as evolutionary regressions, including the thinning of
86 the mandible, the weakening of the jaw musculature, and reduction of the dentition. When such
87 regressed traits are linked to non-pleiotropic or marginally pleiotropic genes, these loci can
88 accumulate mutations that disrupt gene expression and/or translation (Albalat and Cañestro
89 2016). The end result is 'gene loss', either in the form of nonfunctional unitary pseudogenes,
90 characterized by frameshift indels, nonsense mutations and other inactivating mutations, or
91 genes that are deleted from the genome. Numerous examples of such gene loss linked to
92 vestigialization have been documented, ranging from the loss of claw keratins in snakes and
93 worm lizards (Emerling 2017; Holthaus et al. 2025), the loss of insect-digesting chitinase and
94 trehalase genes in mammals that have shifted from insectivory to carnivory and herbivory
95 (Emerling et al. 2018; Janiak et al. 2018; Jiao et al. 2019), and the inactivation of melatonin
96 synthesis and receptor genes in mammals that have lost their pineal glands (Huelsmann et al.
97 2019; Emerling et al. 2021; Valente et al. 2021; Yin et al. 2021).
98 Here, we test whether convergent myrmecophagous mammals show signals of
99 convergent molecular regression, by examining candidate genes related to three components of
100 oral anatomy: dentition, gustation, and masticatory muscle contraction. The development and
101 maintenance of vertebrate teeth relies on coordination between a suite of genes associated with

102 the formation of dentin and overlying enamel, and maintaining their point of contact with the
103 gingiva. Pseudogenizations have been extensively documented for a subset of these genes in
104 edentulous (toothless) and enamelless vertebrates, including turtles (Meredith et al. 2013), birds
105 (Meredith et al. 2014), toads (Shaheen et al. 2021), and baleen whales (Randall et al. 2022).
106 Dental simplification and tooth loss in myrmecophagous mammals is expected to lead to the
107 accumulation of inactivating mutations in these genes, which should vary depending on the
108 degree of dental reduction. Analyses of gustatory cells demonstrate that certain genes are
109 linked to specific taste modalities (Chandrashekar et al. 2006; Yarmolinsky et al. 2009;
110 Chaudhari and Roper 2010; Roper and Chaudhari 2017). Reductions in gustation from the
111 hypothesized ancestral vertebrate complement have been observed in certain species
112 exhibiting dietary specializations (Policarpo et al. 2024), particularly hypercarnivorous mammals
113 and the bamboo-specialist giant and red pandas (Zhao et al. 2010; Jiang et al. 2012), as well as
114 in taxa that have reduced or completely lost taste buds, such as cetaceans (Feng et al. 2014;
115 Kishida et al. 2015) and snakes (Emerling 2017). Given the extreme degree of dietary
116 specialization and modifications to the tongue in myrmecophagous mammals, losses of
117 gustatory genes are expected. Finally, a particular myosin protein (myosin heavy chain 16;
118 MYH16) is expressed exclusively in the jaw closing muscles of multiple vertebrate lineages, and
119 has been associated with species that possess a strong bite force (Hoh 2002; Toniolo et al.
120 2008; Lee et al. 2019). For myrmecophagous mammals, in which mastication has been greatly
121 reduced or even lost (Ferreira-Cardoso et al. 2020), this gene is a promising candidate for
122 pseudogenization.
123 Using these candidate genes, we addressed two main questions. First, do these genes
124 show evidence of similar patterns of gene loss in myrmecophagous mammals, linked to the
125 degree of commitment to myrmecophagy? Second, given that these genes are related to the
126 processing of food, despite being expressed in distinctly different tissues (teeth, taste buds,
127 skeletal muscles), is the timing of gene losses in these three functional categories of genes
128 similar in convergently evolved lineages? If so, this may point to a concomitant sequence of
129 regressive morphological evolution of the oral apparatus associated with the convergent
130 evolution of myrmecophagy. We addressed these questions by interrogating genomic data in a
131 diverse collection of myrmecophagous mammals, including placentals (three anteaters, six
132 armadillos, eight pangolins, the aardvark, two aardwolves, the bat-eared fox, the sloth bear), a
133 marsupial (the numbat), and a monotreme (the short-beaked echidna), representing at least
134 nine convergent adaptations to myrmecophagy.
135

RESULTS

Overall patterns

Each of the examined genes was pseudogenized or deleted in at least some of our focal species (Figure 1). *TAS1R3* was inactivated the least frequently (3/24, 12.5%) with *ACP4* showing evidence of pseudogenization in the vast majority of the focal taxa (21/24, 87.5%). The most anatomically extreme myrmecophagous species, pangolins and anteaters, have the highest proportion of gene losses, with 86.7% (13/15) to 100% of the genes lost in pangolins, and 78.6% (11/14) to 85.7% (12/14) lost in anteaters. The white-bellied (*Phataginus tricuspis*) and black-bellied (*P. tetradactyla*) tree pangolins were the only species to present evidence of inactivation for every examined gene. By contrast, myrmecophagous carnivorans, which present far less anatomical regression in the oral apparatus, ranged from 0 to 6.7% (1/15) gene inactivation. In terms of the timing of gene losses, some appear very recent, including evidently incipient examples (*i.e.*, polymorphic pseudogenes), whereas the earliest pseudogenization estimates pre-date the K/Pg boundary (66 Mya). Some of the non-myrmecophagous outgroup species we examined also possessed pseudogenes for some of the genes we studied (*e.g.*, sloths [Folivora]), and these are due in part to a combination of shared history of gene loss with myrmecophagous species and in other cases may represent adaptations associated with other dietary niches.



170 **FIGURE 1.** Summary of gene functional status in the focal taxa. Evolutionary relationships
171 depicted using a composite phylogeny with branch lengths proportional to time (Dos Reis et al.
172 2012; Gibb et al. 2016; Hassanin et al. 2021; Foley et al. 2023; Heighton et al. 2023). Pholidota
173 = pangolins; Tubulidentata = aardvark; Cingulata = armadillos; Vermilingua = anteaters.
174 Paintings by Carl Buell except giant armadillo, pygmy anteater and numbat (Michelle S.
175 Fabros).
176
177 **Pangolins**
178 Nearly all of the genes examined in this study are pseudogenes or deleted in pangolins (13/15
179 to 15/15; Figure 1; Supplementary Figures S2–S6). Most of these genes show strong evidence
180 of loss prior to the last common ancestor (LCA) of pangolins, particularly due to shared
181 inactivating mutations, suggesting early losses in this clade. For example, the masticatory
182 myosin gene (*MYH16*), is pseudogenized in all five examined species (Figure 2), with 12
183 inactivating mutations shared among them (Supplementary Figure S4). The tooth genes
184 likewise share inactivating mutations (Figure 2; Supplementary Figures S2–S4), but the taste
185 receptor genes are more variable in their patterns of gene loss (Supplementary Figures S4–S6).
186 The tooth genes we examined have been described for some pangolin species,
187 particularly *Manis pentadactyla* (Meredith et al. 2009; Meredith et al. 2014; Choo et al. 2016;
188 Springer et al. 2016; Springer et al. 2019), but here we expand upon these results by
189 characterizing ten dental genes in three Asian pangolins (*Manis* spp.) and two African pangolins
190 (*Smutsia gigantea*, *Phataginus tricuspis*), representing the two major subclades within
191 Pholidota. All ten are pseudogenes or show evidence of whole gene deletion in all five pangolin
192 species investigated. In almost every case, there is evidence of inactivating mutations shared
193 between Asian and African pangolins. For the enamel genes, *ACP4* has 11 shared mutations,
194 *AMELX* has a single shared splice acceptor mutation (intron 3), *ENAM* has four shared
195 mutations, and both *AMBN* and *MMP20* have three shared mutations. *KLK4* returned no BLAST
196 results for all three *Manis* spp. *KLK5* and *KLK15* flank this gene in other boreoeutherian
197 mammals, and both were found on the RefSeq assembly for *Manis javanica* on chromosome
198 17, suggesting a whole gene deletion. For the African pangolins, we were only able to recover
199 the first two exons of *KLK4*, with exon 2 presenting two shared premature stop codons. For the
200 dento-gingival junction genes, *AMTN* exons 1–4 have been deleted in all pangolins, and *ODAM*
201 has eight shared mutations. The dentin development gene *DSPP* has been completely deleted
202 in Asian pangolins, confirmed by the retention of homologous sequence upstream and
203 downstream of the CDS, whereas African pangolins retain it but possess 10 shared mutations.

Finally, *ODAPH*, whose function is correlated with tooth retention, has six mutations shared between all five species (Figure 2).

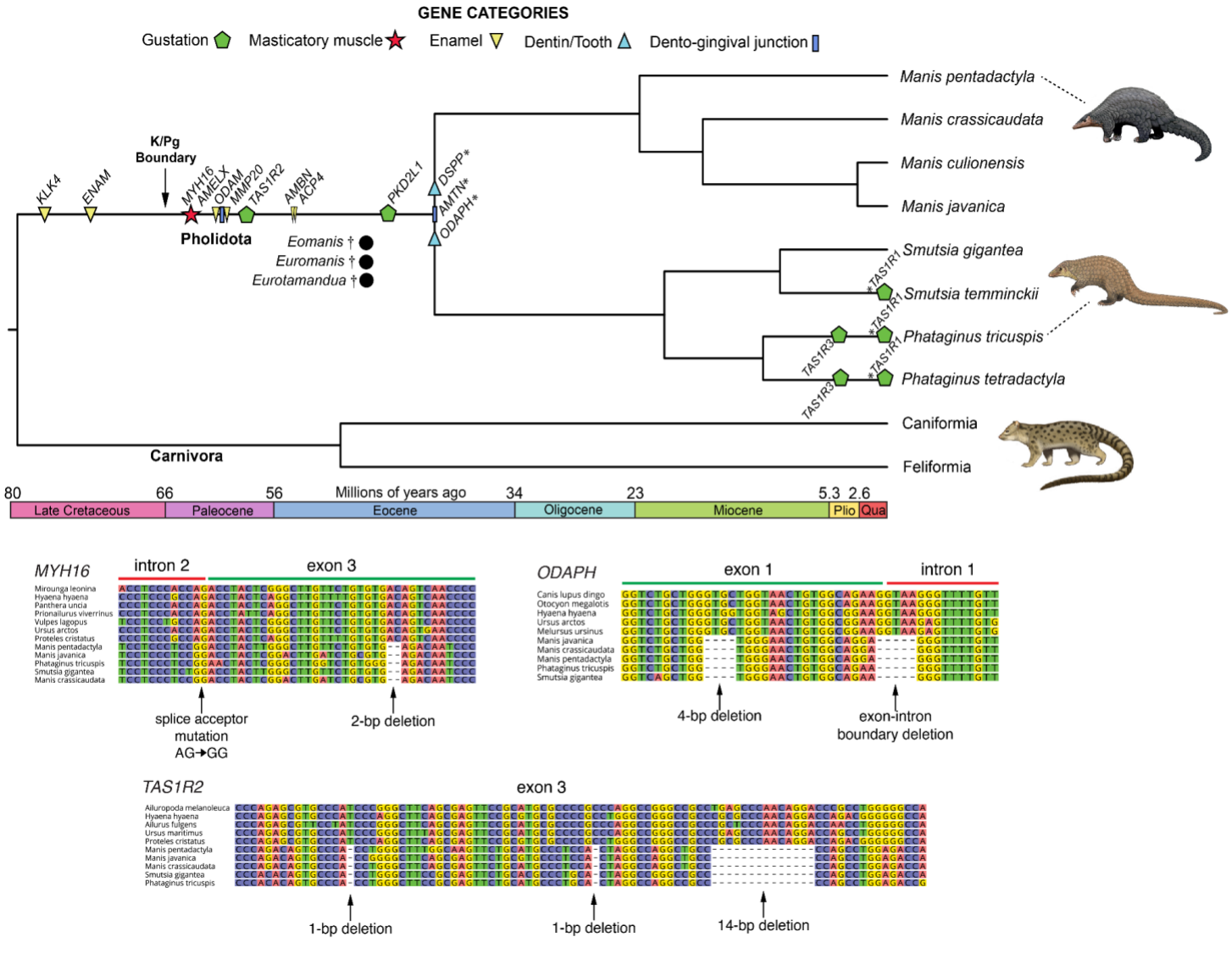


FIGURE 2. Pseudogene dating estimates for pangolins (Pholidota) and DNA sequence alignments for representative genes. Average dates are mapped on a timetree based on divergence times from Heighton et al. (2023). Also included are relevant fossil taxa, indicated by black circles and daggers (see Discussion). * indicates minimum date of inactivation based on distribution of inactivating mutations. These genes were not dated using dN/dS due to violated assumptions. Paintings by Carl Buell.

238 Among the gustatory genes, *TAS1R2* is inactivated in all five pangolins we examined,
239 with 10 shared disabling mutations (Figure 2). At the other extreme, *TAS1R3* is only inactivated
240 in the two *Phataginus* species with up to nine mutations in *P. tricuspis* and five in *P.*
241 *tetradactyla*, although none of these are shared between the two species. A premature stop
242 codon is present in the *Smutsia gigantea TAS1R3*, though this is only 14 codons upstream from
243 the ancestral stop codon, making the functional impact uncertain. Likewise, *TAS1R1* shows
244 relatively minimal evidence of inactivation, presenting as a pseudogene in both *Phataginus*
245 species and *Smutsia temminckii*. Finally, *TAS1R1* in the Sunda (*Manis javanica*), Indian (*M.*
246 *crassicaudata*) and Philippine (*M. culionensis*) pangolins may likewise be a pseudogene, based
247 on an 8-bp insertion in exon 6. However, this is near the ancestral stop codon and results in an
248 additional five residues translated prior to the next stop codon, rendering its status uncertain. All
249 pangolins have inactivating mutations in *PKD2L1* and although none of these are
250 unambiguously shared among all five species, four mutations are shared between the African
251 pangolins, and Asian pangolins possibly share a splice acceptor mutation in intron 8 relative to
252 the ancestral AG condition (GG in *M. crassicaudata* and *M. javanica*; CG in *M. pentadactyla*).
253 Given that shared inactivation mutations only provide a minimum date for
254 pseudogenization, we performed numerous pseudogene dating analyses, based on dN/dS
255 calculations (Supplementary Tables S3–S17), to provide estimates of when relaxed selection
256 effectively began on each locus. In most cases, our dN/dS estimates were consistent with
257 assumptions of transitioning from a fully-functional gene to a pseudogene. Specifically, on a
258 transitional ("mixed") branch, the ω parameter value is expected to be higher than the
259 background ("functional") branches but less than fully relaxed selection or neutral evolution
260 ("pseudogene", $\omega = 1$).
261 Of the 12 genes with mutational evidence suggesting loss on the stem Pholidota branch,
262 nine had ω estimates for the crown Pholidota branches that were not a significantly better fit
263 ($p > 0.05$) relative to models in which ω was fixed at 1 for crown branches (range ω: 0.66–2.75;
264 mean ω: 1.18; median ω: 1.04). This suggests selection was largely relaxed after inactivation.
265 However, three genes deviated from this pattern for crown Pholidota: *DSPP* ($\omega = 0.54$, $p =$
266 0.02); *MYH16* ($\omega = 0.76$, $p = 0.009$); *TAS1R2* ($\omega = 0.52$, $p < 0.00001$). One interpretation is that
267 even after becoming a pseudogene in the stem Pholidota lineage, these genes still experienced
268 some selective constraint. Alternatively, rapid evolution of CpG sites or other features of the
269 nucleotide substitution process that are not modelled by codeml might have distorted ω
270 estimates for these genes. When performing pseudogene dating analyses, we assumed the
271 "pseudogene" branch category to be $\omega = 1$ for the former nine genes, and implemented the

direct estimates of ω for the latter three. For the remaining three genes, which all lack direct evidence of inactivation on the stem Pholidota branch, two showed clear evidence of selective constraint in crown Pholidota, estimating an ω much lower than one, and statistically distinguishable from models in which the crown ω was fixed as one (*TAS1R1*, ω = 0.29; *TAS1R3*, ω = 0.24). However, despite the absence of typical inactivating mutations shared across all pangolins, *PKD2L1* had a crown ω of 1 and a stem ω that was statistically distinguishable from the "functional" category (ω = 0.28 stem branch vs 0.17 background; p = 0.037). This suggests loss of the gene, or at least extensive relaxed selection, on the apical end of the stem Pholidota branch, prior to the acquisition of clear disabling mutations in the descendant lineages.

For the 13 genes with evidence of gene loss on the stem Pholidota branch (including *PKD2L1*), nine have ω estimates that are consistent with assumptions for a "mixed" or transitional branch, being higher than the "functional" ω of outgroups and lower than the "pseudogene" ω. The three that violated this assumption had mixed branches wherein ω was higher than one: *AMTN* (mixed ω: 1.01, functional ω: 0.49, pseudogene ω: 1.14), *DSPP* (mixed ω: 1.08, functional ω: 0.75, pseudogene ω: 0.54), and *ODAPH* (mixed ω: 1.67, functional ω: 0.5, pseudogene ω: 1.47). These all are suggestive of issues with how PAML is modelling these genes, or perhaps a mixed history of positive selection along with purifying and/or relaxed selection on the stem Pholidota branch. Regardless, they were excluded from pseudogene dating estimates due to this assumption violation. For the stem Pholidota ω estimates, nine of 15 genes were statistically distinguishable from the "functional" category. The exceptions were *AMBN*, *AMTN*, *DSPP*, *ODAM*, *TAS1R1* and *TAS1R3*, of which the latter two show no evidence of gene loss on the stem Pholidota branch.

We performed pseudogene dating analyses for all genes that were consistent with our assumptions (Figure 2). The inactivation dates here and for other taxa below indicate the average and a range provided by assuming one synonymous substitution rate vs. assuming two (see Meredith et al. 2009) during the transition from functional genes to pseudogenes: *KLK4*: 77 Mya (76.6–77.4); *ENAM*: 72.8 Mya (71.8–73.8); *MYH16*: 63.7 Mya (62–65.3); *AMELX*: 61.3 Mya (59.6–63); *ODAM*: 60.8 Mya (59.6–61.3); *MMP20*: 60.4 Mya (58.7–60.4); *TAS1R2*: 58.6 Mya (56.9–60.2); *AMBN*: 54.3 Mya (52.7–55.8); *ACP4*: 54.1 Mya (52.6–55.7); *PKD2L1*: 45.8 Mya (45.1–46.5).

306 **Anteaters**

307 Like pangolins, anteaters show a large number of gene losses for our dataset (11/14 to 12/14)
308 (Figure 1; Supplementary Figures S7–S8). The dental genes for xenarthrans have already been
309 reported by Emerling et al. (2023). Among the gustatory genes (Supplementary Figures S7–S8),
310 *TAS1R1* is intact in all three anteaters, *TAS1R3* and *PKD2L1* are variably inactivated, and
311 *TAS1R2* is pseudogenic in all three. For *TAS1R2*, all anteaters share eight inactivating
312 mutations, and their sister group, sloths (Folivora), also have a pseudogenized *TAS1R2*
313 ortholog. The two clades of sloths (*Choloepus* spp., *Bradypus* spp.) share two mutations with
314 anteaters: a 19-bp deletion of the exon 4-intron 4 boundary and an 8-bp deletion in exon 5
315 (Figure 3). *Myrmecophaga tridactyla* presents a premature stop codon in exon 3 of *TAS1R3* in
316 one genome assembly, but this was not reproduced in a second. Furthermore, short-read
317 sequence data suggest that this mutation is polymorphic, as this individual is heterozygous for
318 the stop codon. Only *Cyclopes didactylus* shows evidence of *PKD2L1* pseudogenization,
319 possessing a 2-bp deletion in exon 13, which is polymorphic based on short-read sequence
320 data.

321 *MYH16* is a pseudogene in all three anteaters (Supplementary Figure S8), with three
322 shared disabling mutations in *Tamandua tetradactyla* and *Myrmecophaga tridactyla* (Figure 3).
323 While *Cyclopes didactylus* also has a nonfunctional *MYH16*, possessing 10 disabling mutations,
324 none of these are shared with the other anteaters. Among the sloth outgroups, the two-fingered
325 sloths (*Choloepus* spp.) have an intact *MYH16*, while it is a pseudogene with shared mutations
326 in both of the three-fingered species that we examined (*Bradypus* spp.).

327 Given that the lineage leading to anteaters and armadillos is broken up by cladogenic
328 events (Figure 3), we did not perform ω-based pseudogene dating analyses on xenarthrans,
329 and instead used shared inactivating mutations to provide minimum pseudogenization dates,
330 bracketed by likely maxima. First, we found no evidence of shared inactivating mutations
331 between pilosans (anteaters + sloths) and armadillos, suggesting that the entire complement of
332 genes we examined was functional in the LCA of xenarthrans. After this, assuming the dates
333 from Gibb et al. (2016), *ACP4*, *AMELX*, *ENAM*, *MMP20*, *AMTN,* and *TAS1R2* were all lost on
334 the stem Pilosa (anteaters + sloths) branch (68.1–58.9 Mya). Next, on the stem vermilinguan
335 branch (58.9–38 Mya), *ODAM*, *ODAPH* and *DSPP* were inactivated. On the stem
336 Myrmecophagidae (*Myrmecophaga* + *Tamandua*) branch (38–12.8 Mya), *MYH16* and *AMBN*
337 were disabled, and *TAS1R3* shows evidence of incipient pseudogenization in *Myrmecophaga*
338 *tridactyla*. Likewise, along the *Cyclopes didactylus* branch (38–0 Mya), *MYH16* and *AMBN* were
339 inactivated, with evidence of incipient loss of *PKD2L1*.

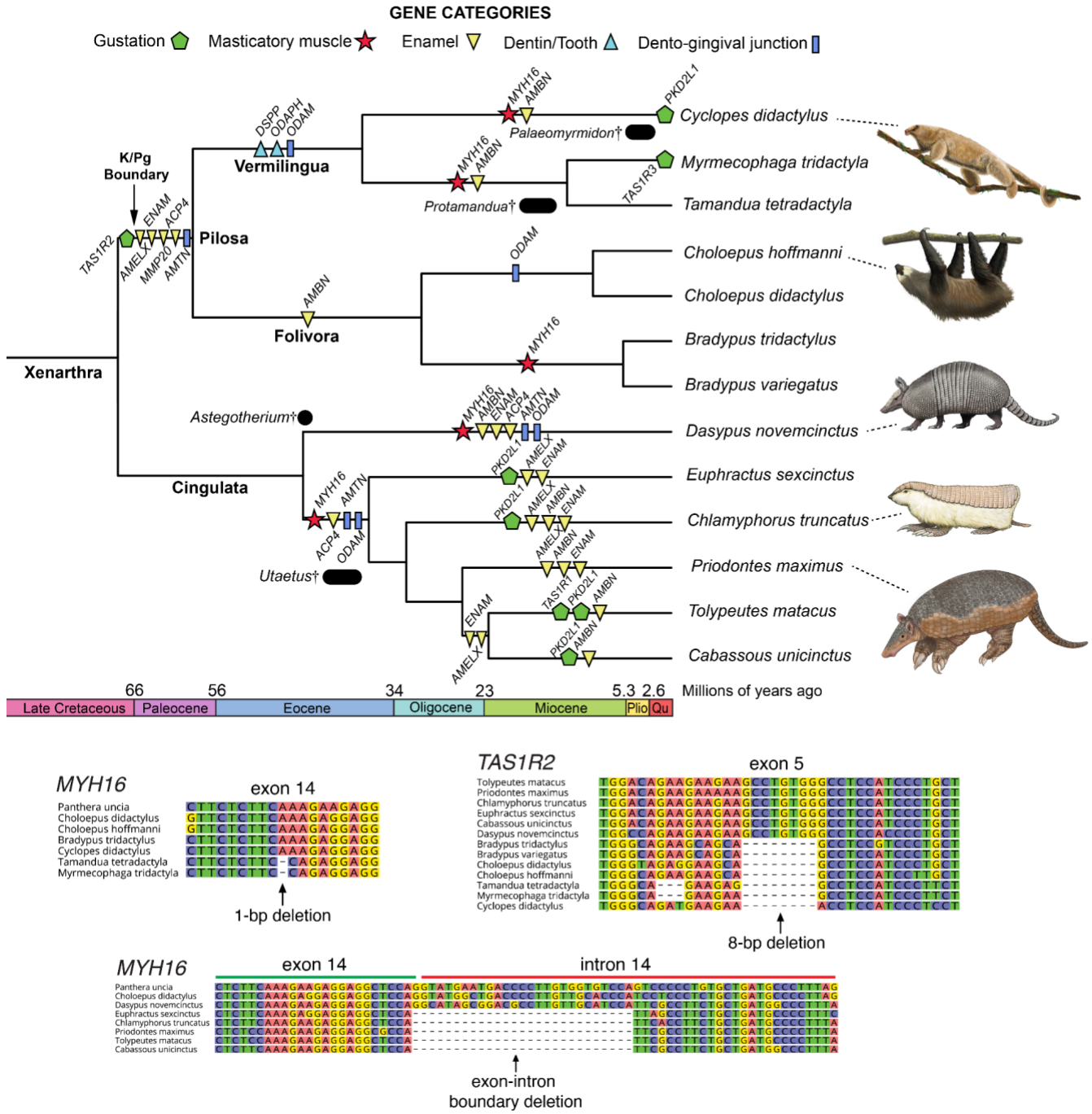


FIGURE 3. Minimum pseudogenization dates in xenarthrans based on shared inactivating mutations and DNA sequence alignments for representative genes. Gene losses were arbitrarily placed midway on pertinent branches of a timetree, with divergence times derived from Gibb et al. (2016). Exceptions are *PKD2L1* in *Cyclopes didactylus* and *TAS1R3* in *Myrmecophaga tridactyla*, which were placed at the branch tips due to being polymorphic pseudogenes. Also included are relevant fossil taxa, indicated by black circles or bars and daggers. Bars indicate the date range for these fossils (see Discussion). Vermilingua = anteaters; Folivora = sloths; Cingulata = armadillos. Paintings by Carl Buell except *Priodontes maximus* and *Cyclopes didactylus* (Michelle S. Fabros).

374 **Armadillos**

375 Armadillos represent the third clade (Cingulata) within Xenarthra, and possess fewer
376 pseudogenes (6/14–8/14) than their anteater relatives (Figure 1). Much like their dental genes
377 (Emerling et al. 2023), armadillo gustatory genes are variably pseudogenic (Supplementary
378 Figures S7–S8). *TAS1R1* has clear positive evidence of inactivation in only *Tolypeutes*
379 *matacus*, with eight inactivating mutations across the gene, validated by two different genome
380 assemblies. Furthermore, *TAS1R2* and *TAS1R3* appear functionally intact in all six armadillos
381 we examined. By contrast, *PKD2L1* is present as a pseudogene in multiple armadillos including
382 *T. matacus* (three inactivating mutations), *Cabassous unicinctus* (two), *Chlamyphorus truncatus*
383 (two) and *Euphractus sexcinctus* (two). By contrast, *MYH16* is universally inactivated in
384 armadillos (Supplementary Figure S8). We found seven mutations shared among all
385 chlamyphorid armadillos (Figure 3). *Dasypus novemcinctus* also has an *MYH16* pseudogene,
386 though none of its mutations are clearly shared with the five chlamyphorids.

387 Given the absence of shared inactivating mutations across dasypodid and chlamyphorid
388 armadillos, the available evidence suggests that the examined genes were intact in the LCA of
389 Cingulata (Figure 3). After the split between dasypodid and chlamyphorid armadillos (45.3 Mya),
390 shared mutations suggest gene losses occurred repeatedly in parallel. On the stem
391 chlamyphorid branch (45.3–37.2 Mya), *MYH16*, *ACP4, AMTN*, and *ODAM* were lost. This was
392 followed by multiple parallel losses of enamel genes (*AMELX*, *ENAM*, *AMBN*) and *PKD2L1*, and
393 a single loss of *TAS1R1*. In dasypodids, between 45.3 Mya and the present, *MYH16*, *AMBN*,
394 *ENAM*, *ACP4*, *AMTN* and *ODAM* were lost. Based on results from Emerling et al. (2023), *ACP4*,
395 *AMTN* and *ODAM* show evidence of loss prior to crown Dasypodidae (12.9 Mya), with *AMBN*
396 and *ENAM* being inactivated more recently. Based on the number of inactivating mutations in
397 *MYH16* (47), it is highly likely this gene was also lost on the stem Dasypodidae branch.

398

399 ***Aardvark***

400 The aardvark's (*Orycteropus afer*) dental pseudogene distribution has been described
401 previously (Meredith et al. 2009; Meredith et al. 2014; Springer et al. 2019), which we confirm
402 here. The aardvark has inactivated orthologs for all of the enamel and dento-gingival junction
403 genes while retaining functional *DSPP* and *ODAPH* (Figure 1). *MYH16* is intact in this species,
404 as are the three *TAS1R*s, whereas *PKD2L1* possesses a 1-bp deletion in exon 8
405 (Supplementary Figure S9).

406 We performed pseudogene dating analyses based on dN/dS modelling (Figure 4;
407 Supplementary Tables S3–S5, S8, S10, S11, S13), with the estimated dates as follows: *ODAM*

51.5 Mya (49.5–53.5), *MMP20* 28.6 Mya (25.7–31.5), *ACP4* 26.4 Mya (23.6–29.3), *AMBN* 23.3 Mya (20.6–25.9), *AMELX* 23 Mya (20.3–25.7), *ENAM* 17.7 Mya (15.4–20), and *PKD2L1* 11.6 Mya (9.9–13.4). We did not date *AMTN* because it violated our assumptions ("mixed branch" $\omega > 1$). dN/dS estimates showed statistically elevated values for five pseudogenes (*ACP4*, *AMTN*, *MMP20*, *ODAM*, *PKD2L1*), whereas *TAS1R1* was statistically reduced compared to the background (Supplementary Tables S3–S8, S10–17).

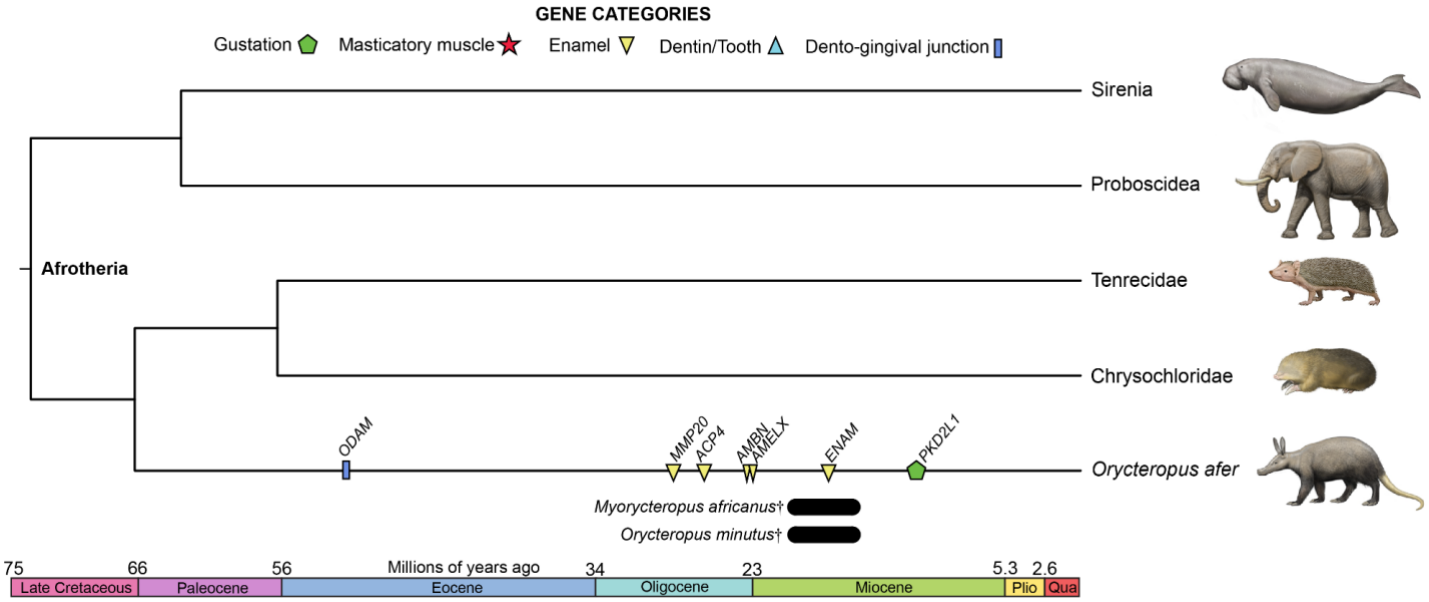


FIGURE 4. Pseudogene dating estimates for aardvark (*Orycteropus afer*). Average dates are mapped on a timetree based on dates from Foley et al. (2023). Also included are relevant fossil taxa, indicated by black bars and daggers. Bars indicate the date range for these fossils (see Discussion). Paintings by Carl Buell.

Carnivorans

Among the carnivorans we examined, disabling mutations were very rare, and when present they always presented as polymorphic (functional/nonfunctional alleles; Supplementary Figures S2, S5 and S6), demonstrating that these mutations are not fixed in their respective species. First, we found no evidence of pseudogenes in the bat-eared fox. In the sloth bear (*Melursus ursinus*), we found a single inactivating mutation in *ACP4*, being heterozygous for a splice donor mutation (GT → GG) in intron 3 (Supplementary Figure S2). The two aardwolves only show evidence of incipient pseudogenization in the *TAS1R*s (Supplementary Figures S5–S6). The

441 Southern aardwolf (*Proteles cristatus*) genome assembly has an 8-bp deletion in exon 1 of
442 *TAS1R1*, which is supported by short read data. However, one individual (NMB12641) is
443 homozygous for this mutation, whereas another individual (NMB12667) is heterozygous. By
444 contrast, its sister-species, the Eastern aardwolf (*P. septentrionalis*), does not possess this
445 mutation, but instead is heterozygous for a 1-bp insertion in exon 3 of *TAS1R2*, which is absent
446 in *P. cristatus*. We did not perform pseudogene dating estimates, given that disabling mutations
447 were not fixed in these species.
448
449 ***Numbat***
450 The numbat (*Myrmecobius fasciatus*) similarly had sparse evidence of gene dysfunction,
451 although more than one gene was affected in this species (Supplementary Figures S10–S11).
452 First, a single enamel gene (*ACP4*) presents evidence of pseudogenization (Figure 5),
453 possessing three 1-bp deletions spread across exons 10 and 11. Much of the gene was
454 irretrievable from the genome assembly due to a gap (10,000 N's), though some upstream
455 exons were assembled using short reads from our specimen. However, we were able to validate
456 these mutations using short read data from three specimens: NCBI BioProject PRJNA786364
457 (Peel et al. 2022), PRJNA512907 (https://www.dnazoo.org/), and our specimen
458 (PRJNA1268346). All three 1-bp deletions were supported by PRJNA786364 and
459 PRJNA512907, and two of the deletions were supported by PRJNA1268346 (assembly gap for
460 the third mutation), suggesting this is a fixed pseudogene in numbats. The remaining two genes
461 have polymorphic mutations. We found inactivating mutations in *TAS1R1*, in contrast to Peel et
462 al. (2022), with one premature stop codon in exon 2 and another in exon 4. Both stop codons
463 are heterozygous in the PRJNA786364 specimen, but absent from PRJNA512907 and
464 PRJNA1268346. Similarly, while *TAS1R2* of the genome assembly lacked evidence for gene
465 inactivation, PRJNA1268346 possesses a single heterozygous 1-bp deletion in exon 6, absent
466 in PRJNA786364 and PRJNA512907. A study of the tongue transcriptome in numbat noted the
467 absence of *TAS1R2* transcripts (Peel et al. 2022), in contrast to *TAS1R1* and *TAS1R3*, pointing
468 to the possibility that regulatory elements may be under relaxed selection, with some numbats
469 beginning to present inactivation mutations.
470 Given evidence of inactivation of *ACP4* in the numbat, we estimated its timing of
471 pseudogenization (Figure 5), and likewise performed dN/dS ratio analyses on the remaining
472 genes (Supplementary Tables S3–S8, S10–S17). For *ACP4*, we used the sequence
473 constructed from our specimen rather than the genome assembly given that the latter was much
474 more fragmented, estimating inactivation at 13 Mya (11.5–14.4). Among all of the genes, eight

had ω values lower on the numbat branch than on the background branches, and of the six with higher ω, only *ACP4* (ω: 0.47, functional ω: 0.12, p = 0.008) and *TAS1R3* had values statistically distinguishable from the background.

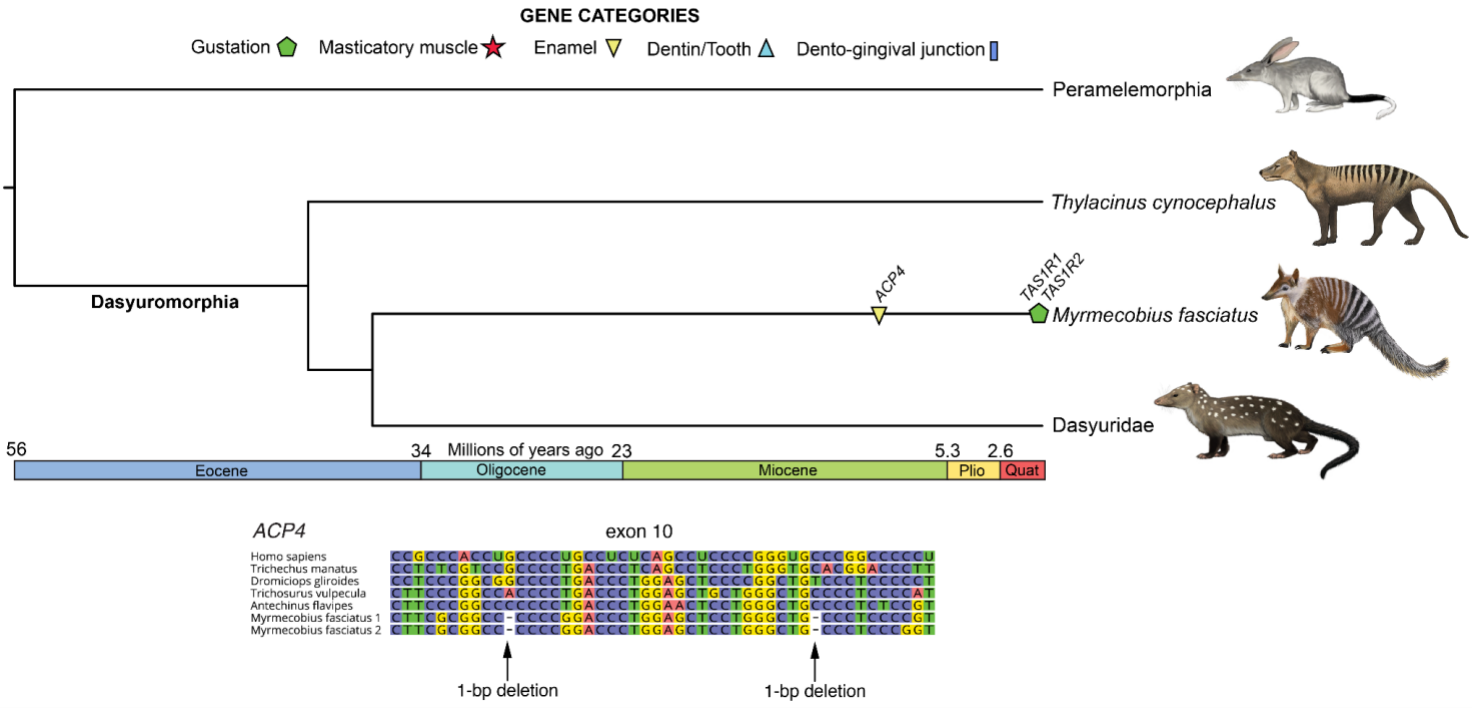


FIGURE 5. Pseudogene dating estimate and DNA sequence alignment for *ACP4* in numbat (*Myrmecobius fasciatus*). Average date is mapped on a timetree based on divergence times from Westerman et al. (2016). *TAS1R1* and *TAS1R2* were placed at the branch tip due to being polymorphic pseudogenes. Paintings by Carl Buell except numbat (Michelle S. Fabros).

Echidna

For the short-beaked echidna (*Tachyglossus aculeatus*), the patterns of gene loss are most similar to those of pangolins and anteaters (Figure 1; Supplementary Figure S12). All of the enamel and dento-gingival genes are either pseudogenes or deleted, and some of these gene losses are shared with the platypus (*Ornithorhynchus anatinus*). Zhou et al. (2021) reported the absence of *AMTN*, *ODAM*, and *MMP20* in both monotremes, *AMBN* and *ENAM* were reported lost uniquely in echidnas, and *DSPP* and *AMELX* were recorded as present and, implied, functional in both species. Here we elaborate on these results, first by confirming that *AMTN* and *ODAM* were completely deleted in both monotremes, and thereby pointing to loss in their LCA. *AMTN*, along with *AMBN* and *ENAM*, is located between *CSN3*, *MUC7* and *JCHAIN* in other mammals. While the latter three were found in synteny in both monotremes, *AMTN* was

absent. *AMBN* and *ENAM* likewise were deleted in echidna, while preserved in the platypus. *ODAM* is situated between *CSN2* and *CSN3* in other mammals, both of which were found in tandem in monotremes with *ODAM* missing. Furthermore, *MMP20* was completely deleted in the platypus, whereas exons 9 and 10 were retained in the echidna, with a 13-bp deletion in exon 9. Like other mammals, *MMP20* is positioned between *MMP7* and *MMP8*, both of which can still be found in synteny within the platypus. We recovered *AMELX* in both monotremes, but it is a pseudogene in echidna, with no recoverable exons 1–3 and two inactivating mutations in exon 4 (Figure 6). The monotreme *DSPP* is difficult to characterize, and indeed appears to have a quite different exon structure from that of placental mammals, as evidenced by BLAST and alignment approaches, as well as the gene structure predicted by NCBI genome annotations. While potential inactivation mutations were observed, RNA sequencing may provide more clarity as to whether this gene is functional in these species. We do note that the echidna predicted mRNA (XM_038769188) possesses a 2-bp deletion that was 'corrected' during model construction, further pointing to potential pseudogenization. However, based on reports of dentin in the egg tooth of echidnas, this may be a false positive (see Discussion).

To these results we add those for *ACP4* and *ODAPH*. The former is a pseudogene in both monotremes, with major deletions (exons 2–9) in the echidna, and a stop codon in exon 10 of the platypus. We found no evidence of *ODAPH* in either monotreme, and both BLAST searches and synteny analyses across vertebrates suggest this may be a therian (Marsupialia + Placentalia) innovation. Among the non-dental genes, while the *TAS1R*s all appear intact in the echidna, *MYH16* and *PKD2L1* are both pseudogenes. The former has at least 22 inactivating mutations (Figure 6), and the latter likewise has numerous inactivating mutations. All five non-dental genes appear functional in the platypus.

We performed gene inactivation estimates for five pseudogenes in echidna (Figure 6; Supplementary Tables S3, S5, S10–S11, S14): *PKD2L1* 49.9 Mya (48.9–50.9); *MYH16* 20.3 Mya (18–22.6); *AMELX* 13.4 Mya (11.6–15.2); *ACP4* 9.9 Mya (8.4–11.3); *MMP20* 1.6 Mya (1.3–1.9). dN/dS estimates were statistically elevated compared to the background for two pseudogenes (*MYH16*, *PKD2L1*), and for all three *TAS1R*s, potentially suggestive of relaxed selection on these gustatory genes (Supplementary Tables S11, S14–S17).

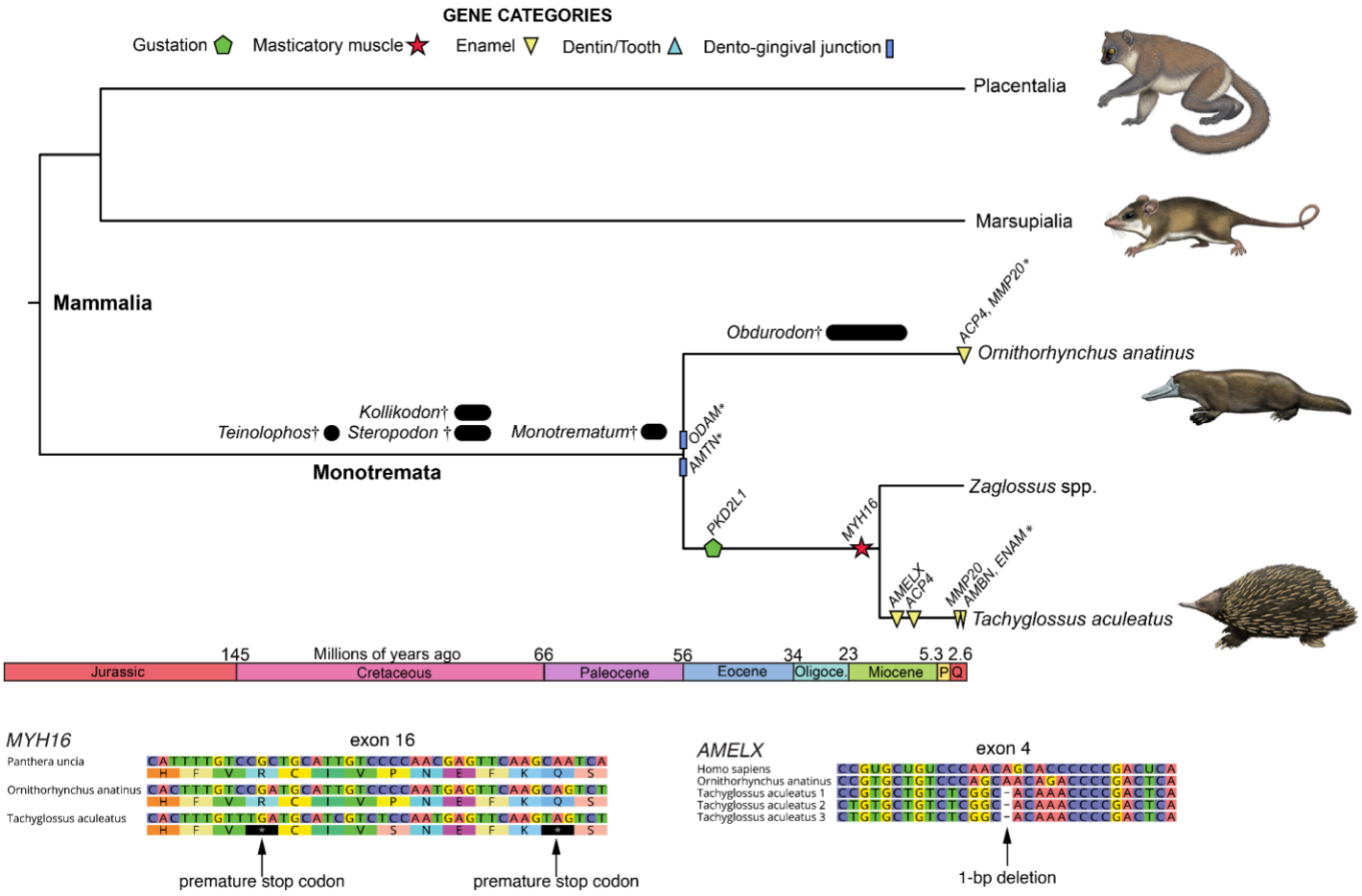


FIGURE 6. Pseudogene dating estimates for short-beaked echidna (*Tachyglossus aculeatus*) and platypus (*Ornithorhynchus anatinus*). Long-beaked echidnas (*Zaglossus* spp.) included for comparison, but not examined in this study. Average dates are mapped on a timetree derived from Dos Reis et al. (2012). * indicates minimum date of inactivation based on whole gene deletions. Also included are relevant fossil taxa, indicated by black circles or bars and daggers. Bars indicate the date range for the earliest fossils (see Discussion). Paintings by Carl Buell.

DISCUSSION

Here we showed that convergent myrmecophagous mammals have lost numerous genes underlying gustation, mastication and dentition in parallel, but their pseudogenization occured at varying degrees and at different evolutionary timescales. Of the genes examined, there was a spectrum of gene loss, ranging from 100% inactivation (African tree pangolins) to complete functional retention (bat-eared fox) or only incipient pseudogenizations (sloth bear, aardwolves). When gene inactivations did occur, the timing of the first estimate was variable, with some lineages showing pseudogenizations near the K/Pg boundary (66 Mya; anteaters), or perhaps even earlier (>70 Mya; pangolins), and others having their first gene loss estimated at 13 Mya

577 (numbat). However, while clusters of pseudogenization events did seem to occur, large variation
578 in timing occurred within lineages. For example, we estimated *KLK4* was inactivated around 77
579 Mya in a stem pangolin, but *TAS1R3* inactivations occurred near 4.5 Mya in the African tree
580 pangolins. Similarly, we estimated six pseudogene events between 68.1 and 58.9 Mya in the
581 lineage leading to anteaters, while also reporting incipient pseudogenizations in the giant and
582 pygmy anteaters, respectively.
583 It is tempting to hypothesize that there may be common patterns in terms of the timing of
584 gene loss across different lineages. Indeed, among dentition genes, three appear to be
585 commonly inactivated earlier in evolutionary history: *ACP4* (anteaters, aardvark, numbat, sloth
586 bear), *AMTN* (anteaters, chlamyphorid armadillos, echidna) and *ODAM* (anteaters, pangolins,
587 aardvark). The gustatory gene *TAS1R1* was rarely pseudogenized, but when it was, it uniformly
588 occurred late (three-banded armadillo, pangolins). However, there were likewise major contrasts
589 for certain genes. For example, while the taste gene *TAS1R2* appears to have been inactivated
590 very early in the history of pangolins and anteaters, and also being among the first to be
591 pseudogenized in an aardwolf and the numbat, it remains intact in all other species. Likewise,
592 the enamel development gene *MMP20* was inactivated early in several convergent lineages
593 (pangolins, anteaters, aardvark), but it perplexingly appears functional in all of the enamelless
594 chlamyphorid armadillos. *PKD2L1* was the most frequently inactivated gustatory gene, but was
595 lost relatively early (echidna), relatively late (xenarthrans, aardvark), or intermediate in timing
596 (pangolins) during the evolution of myrmecophagous mammals. Similarly, we estimated that the
597 masticatory myosin gene *MYH16* has been lost early in pangolins, chlamyphorid armadillos, and
598 echidna, later in anteaters, and was maintained as functional in aardvark. The explanations for
599 this variation in patterns of gene loss may be manifold, with some possibilities discussed below.
600
601 ***Myrmecophagy is associated with some degree of taste loss***
602 One of the more remarkable convergent adaptations of myrmecophagous mammals involves
603 the evolutionary remodelling of the tongue. A typically lengthened, often vermiform tongue,
604 paired with rapid protrusive movements and the production of sticky saliva, allows for the
605 consumption of large quantities of ants and termites during feeding bouts. Modifications are also
606 apparent in the papillae of the dorsum of the tongue. For example, while pangolins are reported
607 to have filiform, fungiform and circumvallate papillae (Prapong et al. 2009), the former two are
608 reported to be poorly-developed (Kubota, Kubota, Nakamura, et al. 1962), with taste buds only
609 being situated in two to three circumvallate papillae (Doran and Allbrook 1973; Prapong et al.
610 2009) or completely absent altogether (Kubota, Kubota, Nakamura, et al. 1962; Abayomi et al.

611 2009). Similarly, in anteaters, fungiform papillae appear to be absent, with taste buds restricted
612 to a pair of weakly innervated circumvallate papillae (Kubota, Kubota, Fukuda, et al. 1962;
613 Casali et al. 2017). Given the reshaping of the tongue into a prey-capturing structure, paired
614 with the tendency towards an extremely specialized diet, we hypothesized that a reduction of
615 taste perception might result from such extreme anatomical and behavioral evolution.
616 We found that there are divergent evolutionary patterns for gustation in these species. At
617 one extreme, some myrmecophagous species retained functional orthologs for the four genes
618 we investigated, suggesting the preservation of the sweet, umami, and sour taste modalities.
619 These included species with a presumably more recent and/or weaker commitment to this
620 dietary habit (sloth bear, bat-eared fox, nine-banded armadillo) but also a species with a
621 relatively long history of myrmecophagy (giant armadillo). At the opposite extreme are the
622 African tree pangolins (*Phataginus* spp.), both of which appear to have lost all three gustatory
623 pathways (Figure 1), suggesting a reduction in taste only matched so far by cetaceans (whales,
624 dolphins) among mammals (Feng et al. 2014). The remainder of the myrmecophagous taxa are
625 along a spectrum, losing one to three gustatory genes, with some being polymorphic
626 pseudogenes. Beyond the genes we examined, the limited studies that have looked at bitter
627 taste receptor genes (TAS2Rs) suggest that this gene family similarly has contracted in
628 myrmecophagous species (Liu et al. 2016; Seymen et al. 2016; Zhou et al. 2021; Li et al. 2024).
629 Loss of gustatory genes has been linked with major shifts in dietary habits, occurring in
630 mammalian hypercarnivores, bamboo specialists and sanguivores (Zhao et al. 2010; Jiang et al.
631 2012; Zhao et al. 2012; Hu et al. 2017). A narrowing of diet may largely explain the reduction of
632 taste observed in most myrmecophages, but there are other possible explanations. First, gene
633 loss may be related to the modification of the tongue to a prey-capturing organ rather than being
634 used to manipulate food and sample tastants. An analogous situation has been described in a
635 few squamate lineages, including snakes, varanids and teiids, which have modified (forked)
636 tongues specialized for vomeronasal sensation, evidently at the expense of most or all taste
637 buds (Schwenk 1985; Young 1997). Indeed, most snakes have lost their sweet and umami
638 receptor genes, with their bitter taste receptor genes likewise reduced (Emerling 2017). Second,
639 perhaps food consumption in myrmecophages is so rapid that it minimizes the need for
640 gustation. This hypothesis has also been put forth to explain taste reduction in cetaceans and
641 penguins (Feng et al. 2014; Cole et al. 2022). All three hypotheses – dietary specialization,
642 tongue modification for prey capture, and rapid consumption of prey – are certainly not mutually
643 exclusive as explanations for the pattern seen in myrmecophagous mammals, and indeed these
644 diverse factors may have reinforced each other. Yet, the retention of some taste modalities in

645 even some of the most specialized myrmecophagous species (*e.g.*, anteaters, most pangolins)
646 point to the possibility that complete taste gene loss is rarely adaptive. The fact that some of
647 these genes are expressed in extragustatory tissues (Roper and Chaudhari 2017) and have
648 been associated with processes as varied as carbohydrate metabolism (Kochem et al. 2024),
649 cerebrospinal fluid monitoring (Liu et al. 2023), skeletal muscle fitness (Serrano et al. 2024),
650 milk production (Kobayashi et al. 2023), testosterone synthesis (Gong et al. 2024), and
651 intestinal inflammation (Shon et al. 2023) may signify that they are unlikely to be lost with
652 regularity due to pleiotropy. Simultaneously, these patterns may indicate how significant it is that
653 gustation loss occurs at all, and has done so repeatedly, in convergent myrmecophagous
654 mammals.
655
656 ***Myrmecophagy is associated with loss of masticatory myosin***
657 Myosin proteins, along with actin, tropomyosin and troponin, drive muscle contraction in
658 mammals. Myosins are heterohexamer proteins that are expressed in muscle cells, with the
659 underlying genes belonging to two families: MYHs (myosin heavy chains) and MYLs (myosin
660 light chains). Most myosin genes are expressed in more than one muscle type (Hoh 2002;
661 Toniolo et al. 2008), whereas the masticatory myosin, encoded by *MYH16*, appears to be
662 exclusively expressed in jaw-closing muscles, including the temporalis, masseters and
663 pterygoideus (Hoh 2002; Hoh et al. 2006; Toniolo et al. 2008; Lee et al. 2019). It has been
664 characterized in mammals such as primates, bats, carnivorans, opossums, dasyurids, and
665 bandicoots, as well as in non-mammalian gnathostomes as distantly-related as crocodylians,
666 turtles and sharks (Hoh 2002; Hoh et al. 2006; Toniolo et al. 2008). This myosin is distinctive in
667 that it seems to allow for a powerful bite for species that possess it, which makes it all the more
668 notable that it is not found in all jawed vertebrates. Specifically, it is absent in distinct lineages of
669 herbivorous mammals, such as kangaroos, ruminants and rabbits. While these species have
670 jaw-closing muscles, their sarcomeres are composed of other myosin classes (Hoh 2002).
671 Given that masticatory myosin provides a powerful bite, shifting from faunivory to dietary habits
672 that rely more on lateral, grinding movements may have rendered *MYH16* superfluous.
673 We expand on these observations, demonstrating that myrmecophagous mammals
674 likewise frequently lose *MYH16*. While faunivorous, species that consume social insects tend to
675 have reduced masticatory muscles and minimal to no chewing capacity (Murray 1981; Redford
676 1987; Smith and Redford 1990; Naples 1999; Ferreira-Cardoso et al. 2020; Thomas et al. 2024;
677 Shoshani et al. 1988; Endo et al. 2017; Endo et al. 2007; Endo et al. 1998). Furthermore, much
678 of the jaw musculature and associated skeletal elements have been rearranged in anteaters,

679 pangolins and echidnas such that they are able to facilitate rapid protrusion of the tongue
680 (Murray 1981; Endo et al. 1998; Endo et al. 2007; Endo et al. 2017; Ferreira-Cardoso et al.
681 2020). Indeed, *MYH16* has been pseudogenized in all three of these clades, in addition to
682 armadillos.
683 Still, some myrmecophagous taxa have retained a functional *MYH16*. The carnivoran
684 myrmecophages retain a functional *MYH16* ortholog, which may reflect more recent adaptations
685 to eating social insects or perhaps phylogenetic constraints associated with an ancestral
686 carnivorous diet. Alternatively, a strong bite force may be retained for behaviors unrelated to
687 diet. The strictly termitivorous aardwolves are territorial and protect their breeding dens from
688 jackals, both of which benefit from jaw muscles that provide a powerful bite (Koehler and
689 Richardson 1990).
690
691 ***Contrasting patterns of dental pseudogenization between myrmecophagous lineages***
692 One of the best characterized examples of convergent evolution in the context of mammalian
693 myrmecophagy concerns what is apparently a ubiquitous simplification of the dentition. Different
694 taxa display various combinations of the following: a reduction in tooth number, supernumerary
695 teeth, enamel thinning and loss, simplification of crowns and roots, weak to no occlusion,
696 monophyodonty, intermolar diastemata, and variation in tooth number on the left versus right
697 sides of the jaws (Redford 1987; Koyasu 1993; Stefen and Rensberger 1999; Davit-Béal et al.
698 2009; Cooper 2011; Kupczik and Stynder 2012; Charles et al. 2013; Hayssen et al. 2013; Carter
699 et al. 2016; Pérez-Ramos et al. 2019). Seemingly paradoxically, myrmecophagous mammals
700 range from being completely edentulous (anteaters, pangolins, echidnas) to having the most
701 teeth of any terrestrial mammals (up to 100 in the giant armadillo), suggesting that the genomic
702 elements controlling the formation of teeth are generally under relaxed selection in these
703 species.
704 The association of tooth and enamel loss with dental pseudogenes is well-documented
705 (Deméré et al. 2008; Meredith et al. 2013; Meredith et al. 2014; Choo et al. 2016; Springer et al.
706 2019; Shaheen et al. 2021; Zhou et al. 2021; Randall et al. 2022; Emerling et al. 2023). Indeed,
707 the evidence is so extensive that these species have been considered genomic models that
708 may aid in the discovery of loci implicated in congenital dental diseases in humans (Emerling et
709 al. 2017). In this study, we have examined species at opposite extremes of dental gene loss,
710 particularly the edentulous pangolins, anteaters, and the short-beaked echidna at one end of the
711 spectrum and the myrmecophagous carnivorans and the numbat at the other.

712 Consistent with patterns seen in other edentulous vertebrates, pangolins, anteaters and
713 the short-beaked echidna present evidence of gene loss in all dental genes we examined
714 (Figure 1), with the possible exception of the dentin matrix gene *DSPP* in the echidna (see
715 below). Species with intermediate levels of dental regression with teeth constituted solely of
716 dentin, including the enamelless aardvark and armadillos, all retain genes necessary for dentin
717 and tooth formation (*DSPP*, *ODAPH*), but lack the dento-gingival genes and most to all of the
718 enamel genes. Myrmecophagous species that retain teeth with enamel crowns, namely
719 myrmecophagous carnivorans and the numbat, largely retain intact dental gene complements,
720 with two exceptions. First, the sloth bear has a pseudogenic allele of *ACP4*, a gene that
721 participates in enamel development and for which inactivation leads to hypoplastic (thin) enamel
722 in humans and mice (Kim et al. 2022; Liang et al. 2022). Likewise, the sole dental pseudogene
723 in the numbat is *ACP4*. Notably, when the pseudogenization pattern of this gene has been
724 compared with the other genes in detailed taxonomic datasets, it appears to be one of the first
725 to become inactivated (Randall et al. 2022; Emerling et al. 2023; Hautier et al. 2023; Randall et
726 al. 2024), which may suggest that dental regression in the sloth bear and numbat is indeed
727 underway.
728
729 ***Pseudogenes of the oral apparatus and their implications for the convergent evolution of***
730 ***myrmecophagy***
731 Organisms can, and often do, evolve analogous solutions to the same adaptive problem (Blount
732 et al. 2018; Cerca 2023), sometimes to the point of exhibiting highly similar traits. Despite
733 sharing their most recent common ancestor over 100 Mya (Foley et al. 2023), anteaters and
734 pangolins evolved remarkably similar anatomy and behavior as a result of their independent
735 adaptations for myrmecophagy. Indeed their striking similarities led anatomical systematic
736 studies to almost invariably group them together in a taxon known as Edentata (Glass 1985;
737 Reiss 2001; Springer et al. 2013). Indeed, we found that genes linked to their distinctive oral
738 anatomy were lost in high numbers in both anteaters and pangolins, providing a genomic record
739 for some of these commonalities (Figure 1). It was only with the advent of molecular
740 phylogenetics that it became abundantly clear that these shared traits were actually the result of
741 convergent evolution, with anteaters being the sister group to sloths, and pangolins recognized
742 as the closest living relatives of carnivorans (Murphy et al. 2001; Delsuc et al. 2002).
743 However, while other convergent myrmecophagous mammals display similar anatomical
744 modifications associated with this specific dietary habit, they are rarely as extreme in their
745 morphology as these two clades. One hypothesis is that this is the result of more recent

746 adaptive forays into eating ants and termites, in which not enough time has elapsed for natural
747 selection to fine-tune oral anatomy to the extremes found in pangolins and anteaters. For
748 example, the termitivorous aardwolves (*Proteles* spp.) are hyaenids, belonging to a radiation of
749 predominantly carnivorous mammals (Carnivora). The three other hyaenid species are
750 characterized by powerful jaws and teeth, capable of crushing bones using specialized enamel,
751 whereas aardwolves have small, peg-like teeth, which can vary in size and number, and
752 possess thin enamel, with cheek teeth separated by intermolar diastemata (Koehler and
753 Richardson 1990; Stefen and Rensberger 1999; Charles et al. 2013). Aardwolves consume
754 termites with the aid of sticky saliva produced by large submaxillary glands, and possess broad,
755 spatulate tongues that collect termites from soil surfaces. While clearly distinct from other
756 hyaenid species, the retention of enamel-capped teeth, a relatively typical mammalian tongue,
757 and powerful jaw muscles paired with robust zygomatic arches and jaws (Koehler and
758 Richardson 1990), likewise contrasts aardwolves strongly with anteaters and pangolins.
759 Aardwolves diverged from other hyaenids about 10 Mya (Koepfli et al. 2006; Hassanin et al.
760 2021), while the fossil aardwolf *Proteles transvaalensis*, which had larger teeth than *P. cristatus*
761 and even retained functional carnassials (Hendey 1974), dates to about 2.6–0.8 Mya. Together,
762 these suggest a relatively small temporal window for aardwolves to adapt to a myrmecophagous
763 niche. In support of this hypothesis, we found evidence of only incipient pseudogenization of
764 taste receptor genes in each aardwolf species (Figure 1).
765 In contrast with aardwolves, the highly specialized pangolins and anteaters show strong
766 evidence of far more ancient forays into myrmecophagy. First, both crown clades originated
767 around 40 million years ago (Gibb et al. 2016; Heighton et al. 2023), and their ancestors likely
768 strongly resembled modern species, being completely edentulous, possessing reorganized jaw
769 muscles and skull elements to facilitate rapid protrusion of a vermiform tongue covered in sticky
770 saliva. Their respective stem lineages mirror each other in the parallel loss of eight dental genes
771 and the sweet gustation gene (Figures 2 & 3). Furthermore, a high concentration of gene losses
772 appears to have taken place soon after the K/Pg boundary (six in stem Pilosa, five in stem
773 Pholidota), suggesting that at least some of these adaptations predated the crown clades by
774 nearly 20 million years, occurring during the radiation of placental mammals in the wake of the
775 end-Cretaceous mass extinction. However, analyzing the timing of gene losses reveals a
776 pattern in which these ancient adaptations may not have been derived all at once, but instead
777 occurred gradually over extensive periods of time.
778 For anteaters (Vermilingua), we have the benefit of a sister-group (sloths, Folivora),
779 which possesses genomic signatures for some of the regressive evolutionary events in their

780 shared history (stem Pilosa). Furthermore, this ancestral branch length is relatively short (9.2
781 Myr; Gibb et al. 2016), allowing for some temporal precision. Five genes underlying enamel
782 development possess inactivating mutations shared by anteaters and sloths, strongly
783 suggesting that enamel was lost on the stem pilosan branch during the Paleocene (Emerling et
784 al. 2023). Furthermore, sweet taste gustation likely was also lost on this ancestral branch
785 (Figure 3). While the fossil record is silent on xenarthran cranial anatomy during this epoch,
786 comparative analyses of chitinase genes (*CHIA*s) suggest that the earliest pilosans were highly
787 insectivorous (Emerling et al. 2018). This, paired with enamelless teeth and reduced gustation,
788 implies that myrmecophagy was the dietary habit of these ancestors. Following the split of
789 sloths (Folivora) and anteaters (Vermilingua), 21 Myr elapsed until the origin of crown
790 vermilinguans (Gibb et al. 2016). Stem vermilinguans likely completely lost their teeth, as
791 evidenced by shared inactivating mutations in *DSPP*, which encodes a dentin matrix protein,
792 and *ODAPH*, a gene correlated with the retention of teeth (Springer et al. 2016; Emerling et al.
793 2023). After the split of vermilinguans into pygmy (Cyclopedidae) and other anteaters
794 (Myrmecophagidae), the last remaining dental gene as well as the masticatory myosin gene
795 were inactivated independently in each lineage. Consistent with these genomic data, the earliest
796 fossil myrmecophagid and cyclopedid anteaters, *Protamandua* (18–15.2 Mya) and
797 *Palaeomyrmidon* (5–3 Mya), respectively, were both edentulous and possessed incomplete
798 zygomatic arches (Gaudin and Branham 1998). Finally, in the terminal branches, evidence for
799 additional incipient taste losses are suggested by polymorphic pseudogenes. Our genomic
800 results suggest that adaptations for myrmecophagy in the lineages leading to modern anteaters
801 have been ongoing for more than 60 million years, even continuing to the present day.
802 Our data on pangolins indicate that the pattern of protracted adaptation in anteaters may
803 not be anomalous. Crown Pholidota dates to 41.3 Mya, diverging from their closest living
804 relatives (Carnivora) 79.5 Mya (Heighton et al. 2023). While we can provide minimum dates for
805 gene losses based on shared mutations, with at least 11 genes lost by 41.3 Mya, pseudogene
806 dating estimates give an additional sense of how long these genes may have evolved neutrally
807 (Figure 2). Our results suggest that two enamel genes were soon after their divergence with
808 carnivorans (77–72.8 Mya), well before the K/Pg boundary. Next, we estimated that five gene
809 losses occurred over a 5.1 Myr window (63.7–58.6 Mya), including dental genes, the
810 masticatory myosin gene, and the sweet taste receptor gene. We dated the inactivation of two
811 additional tooth genes between 54.7–54.1 Mya, with the gustatory *PKD2L1* being inactivated
812 45.8 Mya. The dentin/tooth genes *DSPP* and *ODAPH* could not be dated beyond the LCA of
813 Pholidota, but they affirm the complete absence of teeth in this ancestor. We inferred that most

814 pseudogenization events substantially predate the earliest fossil stem pangolins, all of which
815 date to ~45 Mya (Figure 2) and were already edentulous (Storch 1978; Storch 1981; Gaudin et
816 al. 2009; Gaudin et al. 2020). Despite this long history of myrmecophagy, subsequent
817 inactivations of gustatory genes (*TAS1R1*, *TAS1R3*) occurred in parallel within Pholidota, as
818 recently as 4.8–4.3 Mya. The progressive regression and rearrangement of oral anatomy tied to
819 myrmecophagy is not uniquely suggested by DNA, but rather is further supported by analyses of
820 fossil and extant pangolin crania. From *Eurotamandua* and *Eomanis* (45 Mya), to *Patriomanis*
821 (37.7–33.9 Mya), *Necromanis* (28–14 Mya) and modern pangolins, there is a gradual sequence
822 of reductions and/or losses of masticatory anatomy, including the mandibular coronoid and
823 angular processes, the jugal bone and zygomatic arch (Gaudin et al. 2009; Gaudin et al. 2020).
824 These data raise questions about why such adaptations might occur over such extended
825 periods, and whether it can help explain why other myrmecophagous taxa have oral anatomy
826 that is not quite as extreme as those of anteaters and pangolins. One possibility is that there are
827 other selection pressures acting on the maintenance of certain genes and their emergent
828 phenotypes. Perhaps the most important of these concerns the maintenance of dietary breadth.
829 Given that a complete commitment to myrmecophagy generally involves the loss and reduction
830 of certain oral traits, corresponding with the inactivation and likely permanent loss of critical
831 underlying genes, it may be extremely difficult to switch to certain alternative dietary niches.
832 For example, the bat-eared fox consumes large quantities of ants and termites, but
833 frequently consumes other arthropods, and even vertebrate prey and plant material (Redford
834 1987; Clark Jr. 2005). The sloth bear uses mobile, protrusible lips, and lacks two upper incisors,
835 which facilitate suction of ants and termites (Laurie and Seidensticker 1977), but it also ingests
836 food as diverse as grass, fruits, potatoes, honey and other insects (Laurie and Seidensticker
837 1977). Long-nosed armadillos (*Dasypus* spp.), which commonly consume ants and termites,
838 have diets that can vary geographically and by species, with other invertebrates, plant material
839 and sometimes vertebrates being included in their diets (Redford 1986; Redford 1987; da
840 Silveira Anacleto 2007; Superina and Abba 2014). Even the aardvark, a mammal highly
841 committed to eating ants and termites, does not do so exclusively. This myrmecophage
842 possesses an elongated tongue (up to 30 cm), enlarged mandibular glands that produce
843 strongly adhesive saliva, weak jaw muscles and simplified, enamelless teeth (Shoshani et al.
844 1988; Davit-Béal et al. 2009). Fossils with enamelless teeth very similar to the extant aardvark
845 date to 20–16 Mya (Pickford 1974; Lehmann 2009), with our pseudogene dating estimates
846 pushing back enamel loss, and presumably myrmecophagy, to around 28.6–23 Mya (Figure 4).
847 However, the retention of most gustatory genes, masticatory myosin and dentition may point to

848 the need for a more generalized, versatile anatomy in this species. Indeed, the aardvark has a
849 mutualistic relationship with the aardvark cucumber (*Cucumis humifructus*). This plant fruits
850 underground and the cucumbers are dug out and consumed by aardvarks, which then spread
851 their seeds upon burying their feces (Shoshani et al. 1988). These fruits presumably provide
852 water and essential nutrients that would otherwise be inaccessible if the oral anatomy of
853 aardvarks had regressed to the condition seen in anteaters and pangolins.
854 Competing selection pressures that lead to gradual oral regression may not be solely
855 linked to diet. As a unique example, the short-beaked echidna has perhaps the morphology
856 most comparable to pangolins and anteaters, including an elongated, vermiform tongue, sticky
857 saliva, a highly reduced mandible, reorganized jaw musculature and edentulous jaws (Rismiller
858 and Grutzner 2019). Fossils indicate that edentulism in Monotremata followed a protracted
859 period of dental reduction over >100 Myr (Flannery et al. 2022), with enamel gene loss
860 estimates dating to only 13.4–1.6 Mya (Figure 6). Such a long history of dental regression,
861 culminating in much more recent losses of teeth in the short-beaked echidna may be tied to the
862 fact that monotremes are the only extant oviparous mammals. To aid in their emergence from
863 calcified eggs, monotreme hatchlings use an egg 'tooth'. A recent study characterized the egg
864 tooth of short-beaked echidnas (Fenelon et al. 2023), and found that this 'tooth' lacks enamel
865 but retains dentin. This critical function may have slowed and presumably prevented the loss of
866 certain genomic elements linked to tooth development (Davit-Béal et al. 2009), even though
867 echidnas are edentulous as adults.
868 Answers to how and why myrmecophagous mammals have evolved their specialized
869 anatomy over protracted periods of time may be as diverse as the lineages from which they
870 arose. While oral pseudogenes may reveal part of the picture, further insights will undoubtedly
871 be gained from further exploration of the genomes, transcriptomes, anatomies, and fossils of
872 these iconic examples of convergent evolution.
873
874 **MATERIALS AND METHODS**
875 ***Candidate genes***
876 The genes selected for this project were based on studies that have demonstrated a strong link
877 between gene function and phenotype. Furthermore, these genes are all known to become
878 pseudogenes in at least some instances. For example, numerous genes are known to
879 participate in tooth development, but some do not degrade into pseudogenes in toothless
880 species, likely due to pleiotropy (Deméré et al. 2008; Springer et al. 2016; Emerling et al. 2023).
881 We placed the tooth genes into one of three categories based on their expression patterns and

882 functions inferred from association with human congenital diseases, mouse knockout studies,
883 and interspecies comparisons (Meredith et al. 2013; Meredith et al. 2014; Smith et al. 2017;
884 Randall et al. 2022; Emerling et al. 2023): 1) Enamel-development: *AMELX* (amelogenin),
885 *ENAM* (enamelin), *AMBN* (ameloblastin), *MMP20* (enamelysin), *KLK4* (kallikrein-related
886 peptidase 4), and *ACP4* (acid phosphatase 4) (Lagerstrom et al. 1991; Rajpar et al. 2001; Kim
887 et al. 2005; Meredith et al. 2009; Poulter et al. 2014; Seymen et al. 2015; Seymen et al. 2016);
888 2) Expression during both enamel development and throughout adulthood at the dento-gingival
889 junction (*i.e*., junctional epithelium): *AMTN* (amelotin) and *ODAM* (odontogenic ameloblast-
890 associated) (Nishio et al. 2010; Nakayama et al. 2015; Wazen et al. 2015; Smith et al. 2016;
891 Springer et al. 2019); and 3) Dentin development and tooth retention: *DSPP* (dentin
892 sialophosphoprotein) and *ODAPH* (odontogenesis-associated phosphoprotein) (Xiao et al.
893 2001; Parry et al. 2012; Springer et al. 2016; Emerling et al. 2023). *KLK4* is specific to
894 boreoeutherian mammals (Laurasiatheria + Euarchontoglires), so we only searched for
895 orthologs in pertinent species (Kawasaki et al. 2014). Furthermore, we ignored exons 7–9 in
896 *AMBN* and exon 2 in *ODAPH* due to these exons not being consistently present across
897 mammals (Toyosawa et al. 2000; Springer et al. 2016). Finally, *DSPP* possesses a highly
898 variable repetitive region towards the 3' end of exon 4 (McKnight and Fisher 2009; Fisher 2011),
899 making this region difficult to assemble and align with confidence. As such, this region was often
900 excluded from analyses.
901 While multiple genes are associated with gustation, we examined four genes with clear
902 orthologs that are pseudogenized in at least some vertebrates (Zhao et al. 2010; Jiang et al.
903 2012; Zhao et al. 2012; Feng et al. 2014; Emerling 2017) and known to be linked to specific
904 taste modalities: *TAS1R1* (taste receptor type 1 member 1), *TAS1R2* (taste receptor type 1
905 member 2), *TAS1R3* (taste receptor type 1 member 3), and *PKD2L1* (polycystic kidney disease
906 2-like 1). The protein TAS1R1 dimerizes with TAS1R3 to make the umami taste receptor,
907 responding to certain amino acids and nucleotides and leading to the perception of savory or
908 umami flavors. TAS1R2 likewise forms a heterodimer with TAS1R3, which instead detects
909 sweet tastants, especially monosaccharides and disaccharides (Yarmolinsky et al. 2009; Roper
910 and Chaudhari 2017). Although the specific function of *PKD2L1* is unclear, this gene is
911 expressed in type III cells, which facilitate the perception of sour tastants, making it a strong
912 genetic marker for this taste modality (Tu et al. 2018; Liman and Kinnamon 2021).
913 Finally, we examined the jaw myosin gene, *MYH16* (myosin heavy chain 16). While
914 vertebrate jaw muscles can express several types of myosin proteins, MYH16, also known as
915 the masticatory myosin, appears to be uniquely expressed in jaw-closing muscles. While the

916 gene itself has not been examined in much detail, outside of being a pseudogene in humans
917 (Stedman et al. 2004), the protein has been lost in multiple mammalian species (Hoh 2002; Lee
918 et al. 2019).
919
920 ***Genomic dataset***
921 We used a mixture of publicly available genomes and newly generated data to compile gene
922 sequences from the focal myrmecophagous species, as well as from related 'outgroup' species
923 for comparison. Most sequences were taken from whole genome assemblies on NCBI's RefSeq
924 (Reference Sequence) and WGS (whole genome shotgun) databases (Supplementary Table
925 S1). Gene sequences for the following species were extracted from whole genome assemblies
926 generated by the ConvergeAnt project (Allio 2021; Bioproject ID: PRJNA1238091): the pale-
927 throated three-fingered sloth, *Bradypus tridactylus* (108x); giant anteater, *Myrmecophaga*
928 *tridactyla* (119x); southern tamandua, *Tamandua tetradactyla* (118x); pygmy anteater, *Cyclopes*
929 *didactylus* (91x); giant armadillo, *Priodontes maximus* (71x); pink fairy armadillo, *Chlamyphorus*
930 *truncatus* (92x; PRJNA1238091); six-banded armadillo, *Euphractus sexcinctus* (71x); and giant
931 pangolin, *Smutsia gigantea* (90x; now published, GCA_032199065.1 [Heighton et al. 2023]).
932 Details on the hybrid assembly procedure and genome annotation steps can be found in Allio et
933 al. (2021). The Southern tamandua (*Tamandua tetradactyla*), giant anteater (*Myrmecophaga*
934 *tridactyla*), and bat-eared fox (*O. megalotis*) hybrid genome assemblies were later upgraded to
935 chromosome-length by the DNA Zoo using complementary Hi-C data (Dudchenko et al. 2017;
936 Dudchenko et al. 2018) and re-annotated using the same strategy. Similarly, we used the DNA
937 Zoo annotations of the Hi-C genome assembly for the white-bellied tree pangolin (*Phataginus*
938 *tricuspis*). We also constructed a genomic library for numbat (*Myrmecobius fasciatus*;
939 PRJNA1268346) using a kit from SeqOnce Biosciences. The average insert size for the numbat
940 library was ~410 bp. Paired-end sequencing (150 bp) was performed on a NovaSeq platform at
941 the University of California San Francisco and generated 300 Gb of sequence, which
942 corresponds to ~88X coverage assuming a genome size of 3.42 Gb (Peel et al. 2022, doi:
943 10.46471/gigabyte.47). Finally, we used short read sequences downloaded from NCBI's SRA
944 (Sequence Read Archive) database for the Eastern aardwolf (*Proteles septentrionalis*;
945 SRX9615643; Allio et al. 2021) and the sloth bear (*Melursus ursinus*; ERX1025771; Kumar et
946 al. 2017).
947
948
949

950 ***Gene assembly***

951 A critical goal of this study is to find inactivated genes that may have been evolving neutrally as
952 pseudogenes for perhaps tens of millions of years. This is a challenging task that requires a
953 diversity of approaches to extract ancient pseudogenes from both assembled genomes and, in
954 some cases, short read data. Conventional bioinformatic pipelines typically have limitations in
955 their capabilities to discern ancient, highly degraded pseudogenes, score certain inactivating
956 mutations or validate whole gene deletions. By using a diversity of approaches, including
957 BLAST searches at various thresholds and reference sequences, synteny, and read mapping to
958 target loci sequences, we were able to mine highly divergent inactivated gene loci.

959 To assemble these genes, we first obtained reference mRNA sequences from NCBI
960 (GenBank) for the targeted genes in our study, typically for *Homo sapiens*. When this was not
961 possible, we used alternative curated mammal mRNA sequences or gene models derived from
962 the NCBI eukaryotic genome annotation pipeline (Thibaud-Nissen et al. 2016). We imported the
963 full mRNA sequence into Geneious Prime (v2019.2.3; Kearse et al. 2012) and used the
964 annotations provided by GenBank to identify exon/intron boundaries and coding sequence
965 structure. Since we did not evaluate non-coding DNA in this study, we did not analyze the
966 structure of untranslated regions. As such, exon numbering throughout this study is based on
967 coding exons only.

968 Once we obtained the coding regions of the reference mRNA sequences, we used these
969 to obtain sequences derived from whole genome assemblies. We began by obtaining
970 sequences from NCBI's RefSeq and WGS databases by performing similarity searches with
971 BLAST (discontiguous megablast) using the mRNA reference sequence against target
972 assemblies. We obtained the hits with the highest E-values that likewise spanned the largest
973 portions of the reference mRNA sequence, especially those found towards the 5' and 3' ends.
974 We then downloaded the contig or scaffold regions encompassing the entire coding sequence
975 (CDS) and some flanking sequence (100+ bp) and imported them into Geneious. Next, we
976 performed automated sequence alignments using MUSCLE (ver 3.8.425; Edgar 2004) within
977 Geneious, aligning the reference mRNA against the contig or scaffold to identify coding exons.
978 Some genome assemblies also had exon annotations for comparison. In cases where
979 sequences had long strings of unknown bases (*e.g.*, >50 Ns), all but ~10 Ns were deleted to
980 facilitate better sequence alignments. If after aligning the sequence the orthology appeared
981 dubious, we performed reciprocal BLAST (megablast) searches against NCBI's core nucleotide
982 database to confirm gene identity. This was validated a second time using gene tree analyses

983 (see below). After we obtained a whole gene sequence closely-related to a focal taxon, we then
984 used this in subsequent BLAST searches for a subset of representative species..
985 After extracting whole gene sequences from representative taxa from NCBI, these data
986 were used to mine targeted gene sequences locally from new genome assemblies. Once these
987 assemblies were imported into Geneious, we used a combination of BLAST searches and short-
988 read mapping to pull out gene sequences. BLAST searches followed the same basic
989 parameters as above in NCBI, but we also set BLAST results to return an additional 100-bp
990 flanking on each hit. For mapping, we used Minimap2 v2.24 (Li 2018) via the long-read spliced
991 alignment mode (with default parameters) as implemented in Geneious Prime v2022.1.1. This
992 splice-aware alignment enables the mapping of CDS, cDNA, or mRNA sequences against
993 reference genomic sequences containing introns by dividing the query protein-coding sequence
994 to its different mapping locations. The genomic regions that included both exons and
995 intercalating introns were then extracted for downstream analyses.
996 We used short read data derived from whole genome sequencing to assemble genes for
997 the sloth bear (*Melursus ursinus*; NCBI's SRA), a numbat (*Myrmecobius fasciatus*) specimen,
998 and the Eastern aardwolf (*Proteles septentrionalis*). For the sloth bear and numbat, we first
999 imported whole genome sequencing reads into Geneious, and obtained reference gene
1000 sequences from closely-related species (ursids and dasyurids, respectively) using BLAST
1001 methods on NCBI (WGS, RefSeq). We mapped the paired-end short reads to their respective
1002 reference sequences using the Geneious Prime mapper, with settings at Medium-Low
1003 Sensitivity/Fast with no fine-tuning (fast / read mapping), no sequence trimming and mapping
1004 multiple best matches randomly. Each mapping result was then examined by eye, and in cases
1005 of abundant erroneous mappings (*i.e.*, nonhomologous sequences), we remapped those same
1006 reads at Low Sensitivity/Fastest up to five iterations, in which reads were mapped to the
1007 consensus from the previous iteration. Again, we examined each mapping alignment by eye and
1008 manually adjusted as needed.
1009 For the Eastern aardwolf (*Proteles septentrionalis*), paired-end short reads from SRA
1010 were mapped to orthologs from the Southern aardwolf (*P. cristatus*) using bowtie2 v2.3.4.3
1011 (Langmead and Salzberg 2012), with default parameters. Mapped reads were extracted from
1012 the bam mapping files and converted to fastq files with SAMtools v1.10 (Danecek et al. 2021).
1013 These reads were then imported into Geneious Prime and remapped against reference gene
1014 sequences for visualization.
1015
1016

1017 ***Sequence alignments and gene functionality analyses***

1018 While obtaining orthologs for a single gene, we performed successive DNA sequence
1019 alignments with MUSCLE in Geneious, starting by aligning two closely related taxa, then
1020 aligning these two to a third, those three to a fourth, and so on. By doing this, we were able to
1021 anchor the alignment using highly similar sequences and then progressively add more divergent
1022 ones, thereby minimizing alignment errors. In addition, this allowed us to examine each
1023 alignment by eye to search for errors, a common issue when aligning divergent pseudogene
1024 and intron sequences. Each alignment generally only included orthologs from species belonging
1025 to a specific clade, to maximize our ability to determine homology.

1026 Next, we examined the gene structure for every ortholog to determine gene
1027 completeness. When data appeared to be missing, based on exon annotations from the
1028 reference mRNA, we executed additional BLAST searches with more relaxed parameters
1029 and/or alternative reference sequences. In some cases, there was evidence of partial or whole
1030 gene deletion, which required us to examine additional sequence data upstream or downstream
1031 from our original BLAST searches. In some cases where genes appeared to be deleted, we did
1032 synteny analyses to validate their absence. This involved two strategies used in tandem. First,
1033 we would examine the vertebrate-wide synteny comparisons provided in Genomicus v110.01
1034 (Nguyen et al. 2022) to determine which genes should be flanking the putatively deleted gene.
1035 In most cases, we then obtained NCBI RefSeq sequences from closely-related species that
1036 contained the gene of interest along with annotated genes upstream and downstream of the
1037 focal gene, comparing them to the Genomicus synteny maps. We then searched genome
1038 assemblies (NCBI RefSeq, WGS) for these flanking genes in the species appearing to lack the
1039 gene, and then downloaded and aligned these multigenic sequences to the related comparison
1040 species. If the resulting alignment showed a major gap where the focal gene should be, not
1041 explainable by an assembly gap indicated by N's, we considered this a whole gene deletion.

1042 Once our gene alignments were complete (Supplementary Dataset S1), we searched for
1043 evidence of the following categories of inactivating mutations: start codon mutations, frameshift
1044 insertions and deletions, premature stop codons, splice site mutations, alterations to the
1045 ancestral stop codon, exon-intron boundary deletions, and whole exon deletions. To search for
1046 premature stop codons and mutations in the ancestral start and stop codons, we generated
1047 alignments encompassing the coding sequences only. To do so, we removed all introns, deleted
1048 all frameshift insertions, and inserted Ns to restore the correct reading frame due to frameshift
1049 deletions.

1050 When we found evidence of at least one disabling mutation in a species, we typically
1051 attempted to rule out sequencing or assembly errors by using data from at least one of four
1052 different sources. First, if the sister taxon shared the mutation, we considered this validation that
1053 the mutation was inherited from a common ancestor (see below). Second, in some cases,
1054 alternative genome assemblies, sometimes from separate individuals, were available either in
1055 NCBI's WGS database or in our local datasets to validate the mutation. Third, for some NCBI-
1056 derived sequences, we BLASTed (megablast) the relevant regions against short reads in
1057 available experiments in SRA, downloaded them and mapped them in Geneious. Then we
1058 examined the short reads mapped to the mutation by eye. Finally, for others (*Cyclopes*
1059 *didactylus*, *Proteles cristatus*, *Myrmecobius fasciatus*), we mapped short reads using *de novo*
1060 sequence data on the gene sequence with bowtie2 or the Geneious mapper, as described
1061 above. Two individuals were used for *P. cristatus* (NMB12641 and NMB12667), one for *C.*
1062 *didactylus* (M2300) and three for *M. fasciatus* (NCBI PRJNA512907, PRJNA786364,
1063 PRJNA1268346). Together, these methods enabled us to confirm that an inactivating mutation
1064 was not the result of a sequencing or assembly error, and in a few cases determine whether a
1065 particular mutation was heterozygous or homozygous.
1066 In addition to characterizing inactivating mutations in individual species, we checked
1067 whether any were shared among two or more closely-related species. Such shared mutations
1068 are suggestive of loss of gene function in a common ancestor, thereby providing a minimum
1069 pseudogenization date. The only shared mutations were found in xenarthrans and pangolins, so
1070 we referred to the phylogenies of Gibb et al. (2016) for the former and Heighton et al. (2023) for
1071 the latter. Most of the gene sequences we collected for pangolins were from five species only
1072 (*Smutsia gigantea*, *Phataginus tricuspis*, *Manis pentadactyla*, *M. javanica*, and *M.*
1073 *crassicaudata*), because genome assemblies for *S. temminckii*, *P. tetradactyla* and *M.*
1074 *culionensis* were not available until we were far into our analyses. These new assemblies were
1075 rendered largely redundant by our initial sample of five species due to the presence of shared
1076 inactivating mutations in all five, Asian pangolins only (*Manis* spp.) or African pangolins only
1077 (*Phataginus* spp., *Smutsia* spp.), representing the deepest splits in Pholidota. For the subset of
1078 candidate genes that did not show evidence of loss on these early branches, we obtained,
1079 aligned and examined the orthologs for the remaining three pangolin species.
1080
1081
1082
1083

1084 ***Selection analyses based on dN/dS and gene inactivation dating***

1085 Given the phylogenetic distribution of gene losses during the xenarthran radiation (see Results
1086 below), we were able to provide minimum date estimates of gene inactivation based on patterns
1087 of shared mutations in this clade. However, for other myrmecophagous taxa, there are long
1088 branches on which genes were lost as autapomorphies, which do not permit any resolution of
1089 the relative timing of inactivation events in different genes. As such, we estimated the timing of
1090 pseudogenization in the remaining species that possess fixed (*i.e.*, not polymorphic) inactivating
1091 mutations (pangolins, the aardvark, short-beaked echidna, numbat) by employing the dN/dS-
1092 based method of Meredith et al. (2009).

1093 To begin with, for each gene, we aligned functional orthologs from outgroup taxa
1094 (Supplementary Table S2) and modeled their evolution over a fixed phylogeny (Supplementary
1095 Dataset S2, Supplementary Tables S3-S17), to estimate the expected background ω value
1096 when the gene is under functional constraint ("functional" branches). These orthologs were
1097 provided from annotated mRNAs and predicted mRNA models from GenBank, the latter of
1098 which are derived from NCBI's Eukaryotic Genome Annotation Pipeline
1099 (https://www.ncbi.nlm.nih.gov/genome/annotation_euk/process/). Our taxonomic sampling for
1100 outgroup taxa was broadly representative of major lineages of mammals, with the following
1101 clades being represented, when possible, for every dN/dS analysis (Supplementary Table S2):
1102 Marsupialia: Didelphimorphia, Microbiotheria, Dasyuromorphia, Diprotodontia; Afrotheria:
1103 Afrosoricida, Macroscelidea, Sirenia, Proboscidea; Euarchontoglires: Scandentia, Dermoptera,
1104 Primates, Lagomorpha, Rodentia; Laurasiatheria: Eulipotyphla, Chiroptera, Perissodactyla,
1105 Carnivora, Cetartiodactyla. In some cases, the targeted genes were not available from some
1106 outgroup taxa, perhaps due to genome assembly gaps, being a heavily degraded pseudogene
1107 or the entire gene having been deleted. The Meredith et al. (2009) method assumes that
1108 outgroup orthologs are functional, so we also examined the annotations of predicted mRNA
1109 models for information suggestive of pseudogenization. At NCBI, a gene's product is given a
1110 "LOW QUALITY PROTEIN" designation when frameshift indels and/or premature stop codons
1111 are 'corrected' during the construction of the gene model. Any such gene models were removed
1112 from the dN/dS analysis dataset. The final datasets had as few as nine outgroup species (*KLK4*)
1113 and as many as 18 (median of 17; Supplementary Table S2).

1114 These outgroup sequences were imported into Geneious, where we removed any
1115 portions of the gene models outside of the coding sequence annotation, and then performed
1116 translation alignments (MUSCLE), which we then examined by eye. We manually removed stop
1117 codons, taxon-specific predicted exons and insertions, and any codons with questionable

1118 homology (*e.g.*, straddling a deletion with ambiguous homology). Next, we performed a
1119 maximum likelihood phylogenetic analysis using RAxML (Stamatakis 2014; v8.2.11;
1120 GTR+GAMMA model, Rapid hill-climbing) to validate orthology of the genes by comparing
1121 topology and branch length to what is predicted by published species trees (Supplementary
1122 Figure S1).

1123 We then analyzed the outgroup taxa for each candidate gene using PAML v4.8 codeml
1124 analyses (Yang 2007) to determine the best fit codon frequency model. For placental mammal
1125 relationships, we implemented the topology from Fig. 4a in Foley et al. (2023), and for marsupial
1126 relationships, we referred to Mitchell et al. (2014). We then performed one ratio model analyses
1127 using the 1/61 each, F1X4 and F3X4 codon frequency models and determined the best fit using
1128 the Akaike information criterion. The best fitting model was then implemented for subsequent
1129 taxon-specific pseudogene dating analyses (Supplementary Table S3–S17).

1130 To estimate gene inactivation dates for specific taxa, we ran separate analyses for
1131 pangolins, the aardvark, the short-beaked echidna and the numbat, respectively
1132 (Supplementary Dataset S2). In each case, coding sequences of focal taxa were aligned to the
1133 corresponding outgroup alignment (MUSCLE, Translation align). Each resulting alignment was
1134 examined by eye, and we manually removed stop codons, partial codons resulting from
1135 frameshift deletions, lineage-specific insertions and any codons with ambiguous homology. We
1136 then implemented branch models in codeml of PAML to estimate the ω values for the branches
1137 on which pseudogenization occurred ("mixed" branches) and, when relevant (*i.e.*, most pangolin
1138 analyses), branches on which selection is expected to be completely relaxed ("pseudogene"
1139 branches). We then entered the relevant ω categories into the calculations provided by
1140 Meredith et al. (2009), utilizing divergence times provided by different studies.

1141 For pangolins, we usually only included four species in the analyses: *Smutsia gigantea*,
1142 *Phataginus tricuspis*, *Manis pentadactyla* and *M. javanica* or *M. crassicaudata*. Since these taxa
1143 represent the deepest divergences within Pholidota, they are probably sufficient for generating
1144 dN/dS estimates for genes lost on the stem Pholidota branch. Given the lack of shared
1145 inactivating mutations in *TAS1R1* and *TAS1R3*, we included all eight pangolin species in those
1146 analyses. We implemented the topology from Fig. 2 of Heighton et al. (2023), and set the
1147 following branch categories: the "functional" background branches received one ω, the "mixed"
1148 branch(es), on which gene inactivation likely occurred, received its own ω, and then the
1149 branches that post-dated the pseudogenization event ("pseudogene") received their own ω.
1150 Though the "pseudogene" branches are expected to have a ω of 1 (relaxed selection), we ran a
1151 second model in which the "pseudogene" branches were fixed at 1. We compared the two

models with a likelihood ratio test, and if there was no statistically significant difference between the models, we assumed the "pseudogene" **ω** was 1 when performing pseudogene dating calculations. By contrast, if there was a significant difference in models, we used the estimated **ω** for the calculations. Divergence times used for these calculations were likewise derived from Fig. 2 of Heighton et al. (2023).

For the aardvark, short-beaked echidna, and numbat dN/dS models, the branch leading to the focal taxon always represented the "mixed" branch, and all other branches were "functional" branches. No branches were designated as "pseudogene" branches, so the pseudogene **ω** was assumed to be 1. For the aardvark, we assumed that its lineage split from other afroinsectiphilian mammals 66.3 Myr ago (Foley et al. 2023, neutral 241 species average). For the short-beaked echidna, we assumed it diverged from the platypus (*Ornithorhynchus anatinus*) 56.4 Mya and monotremes from therian mammals at 184.7 Mya (Dos Reis et al. 2012). Since the platypus shows evidence for relevant gene losses (Zhou et al. 2021), we obtained orthologs by interrogating genome assemblies for this species using the methods described above and included these sequences in dN/dS analyses. Finally, for the numbat we assumed a divergence from Dasyuridae 36.5 Mya (Westerman et al. 2016).

For each taxon set (pangolins, aardvark, echidna, numbat), we performed dN/dS estimates for cases in which there was no evidence of gene loss on the branch of interest, to compare patterns of dN/dS estimates in pseudogenes versus putatively functional orthologs as controls. We also compared each of these to models in which the focal branch **ω** was fixed as part of the background. This allowed us to test whether the dN/dS was statistically distinguishable from the background, as determined by a likelihood ratio test.

ACKNOWLEDGEMENTS

This is dedicated to Averie Jeanne-Lorraine 玉蘭. We thank Lionel Hautier for insightful discussions and comments on the manuscript, and two anonymous reviewers for their constructive feedback. This work has been supported by grants from the European Research Council (ConvergeAnt project: ERC-2015-CoG-683257; FD) and Investissements d'Avenir of the Agence Nationale de la Recherche (CEBA: ANR-10-LABX-25-01; CEMEB: ANR-10-LABX-0004; FD); a National Science Foundation Postdoctoral Research Fellowship in Biology (award no. 1523943; CAE); a National Science Foundation Postdoctoral Fellow Research Opportunities in Europe award (CAE); the People Programme (Marie Curie Actions) of the European Union's Seventh Framework Programme (FP7/2007-2013) under REA grant agreement no. PCOFUND-

GA-2013-609102, through the PRESTIGE programme coordinated by Campus France (CAE).
This is contribution ISEM 2025-XXX of the Institut des Sciences de l'Evolution de Montpellier.

Data availability statement

Supporting data are available from zenodo.org (https://doi.org/10.5281/zenodo.14895250).

REFERENCES

Abayomi TA, Ofusori DA, Ayoka OA, Odukoya SA, Omotoso EO, Amegor FO, Ajayi SA, Ojo GB, Oluwayinka OP. 2009. A Comparative Histological Study of the Tongue of Rat (*Rattus norvegicus*), Bat (*Eidolon helvum*) and Pangolin (*Manis tricuspis*). *Int. J. Morphol.* 27.

Albalat R, Cañestro C. 2016. Evolution by gene loss. *Nat. Rev. Genet.* 17:379–391.

Allio R. 2021. Phylogenomics and comparative genomics in ant-eating mammals.

Allio R, Tilak M-K, Scornavacca C, Avenant NL, Kitchener AC, Corre E, Nabholz B, Delsuc F. 2021. High-quality carnivoran genomes from roadkill samples enable comparative species delineation in aardwolf and bat-eared fox. *eLife* 10:e63167.

Blount ZD, Lenski RE, Losos JB. 2018. Contingency and determinism in evolution: Replaying life's tape. *Science* 362:eaam5979.

Carter TS, Superina M, Leslie DM. 2016. *Priodontes maximus* (Cingulata: Chlamyphoridae). *Mamm. Species* 48:21–34.

Casali DM, Martins-Santos E, Santos ALQ, Miranda FR, Mahecha GAB, Perini FA. 2017. Morphology of the tongue of Vermilingua (Xenarthra: Pilosa) and evolutionary considerations. *J. Morphol.* 278:1380–1399.

Cerca J. 2023. Understanding natural selection and similarity: Convergent, parallel and repeated evolution. *Mol. Ecol.* 32:5451–5462.

Chandrashekar J, Hoon MA, Ryba NJP, Zuker CS. 2006. The receptors and cells for mammalian taste. *Nature* 444:288–294.

Charles C, Solé F, Rodrigues HG, Viriot L. 2013. Under pressure? Dental adaptations to termitophagy and vermivory among mammals. *Evolution* 67:1792–1804.

Chaudhari N, Roper SD. 2010. The cell biology of taste. *J. Cell Biol.* 190:285–296.

Choo SW, Rayko M, Tan TK, Hari R, Komissarov A, Wee WY, Yurchenko AA, Kliver S, Tamazian G, Antunes A, et al. 2016. Pangolin genomes and the evolution of mammalian scales and immunity. *Genome Res.* 26:1312–1322.

Clark Jr. HO. 2005. *Otocyon megalotis*. *Mamm. Species*:1–5.

Cole TL, Zhou C, Fang M, Pan H, Ksepka DT, Fiddaman SR, Emerling CA, Thomas DB, Bi X, Fang Q, et al. 2022. Genomic insights into the secondary aquatic transition of penguins. *Nat. Commun.* 13:3912.

Cooper CE. 2011. Myrmecobius fasciatus (Dasyuromorphia: Myrmecobiidae). *Mamm. Species* 43:129–140.

Danecek P, Bonfield JK, Liddle J, Marshall J, Ohan V, Pollard MO, Whitwham A, Keane T, McCarthy SA, Davies RM, et al. 2021. Twelve years of SAMtools and BCFtools. *GigaScience* 10:giab008.

Davit-Béal T, Tucker AS, Sire J. 2009. Loss of teeth and enamel in tetrapods: fossil record, genetic data and morphological adaptations. *J. Anat.* 214:477–501.

Delsuc F, Scally M, Madsen O, Stanhope MJ, De Jong WW, Catzeflis FM, Springer MS,

Douzery EJP. 2002. Molecular Phylogeny of Living Xenarthrans and the Impact of Character and Taxon Sampling on the Placental Tree Rooting. *Mol. Biol. Evol.* 19:1656–1671.

Deméré TA, McGowen MR, Berta A, Gatesy J. 2008. Morphological and Molecular Evidence for a Stepwise Evolutionary Transition from Teeth to Baleen in Mysticete Whales. *Syst. Biol.* 57:15–37.

Doran GA, Allbrook DB. 1973. The Tongue and Associated Structures in Two Species of African Pangolins, *Manis gigantea* and *Manis tricuspis*. *J. Mammal.* 54:887–899.

Dos Reis M, Inoue J, Hasegawa M, Asher RJ, Donoghue PCJ, Yang Z. 2012. Phylogenomic datasets provide both precision and accuracy in estimating the timescale of placental mammal phylogeny. *Proc. R. Soc. B Biol. Sci.* 279:3491–3500.

Edgar RC. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* 32:1792–1797.

Emerling CA. 2017. Genomic regression of claw keratin, taste receptor and light-associated genes provides insights into biology and evolutionary origins of snakes. *Mol. Phylogenet. Evol.* 115:40–49.

Emerling CA, Delsuc F, Nachman MW. 2018. Chitinase genes (*CHIA*s) provide genomic footprints of a post-Cretaceous dietary radiation in placental mammals. *Sci. Adv.* 4:eaar6478.

Emerling CA, Gibb GC, Tilak M-K, Hughes JJ, Kuch M, Duggan AT, Poinar HN, Nachman MW, Delsuc F. 2023. Genomic data suggest parallel dental vestigialization within the xenarthran radiation. *Peer Community J.* 3:e75.

Emerling CA, Springer MS, Gatesy J, Jones Z, Hamilton D, Xia-Zhu D, Collin M, Delsuc F. 2021. Genomic evidence for the parallel regression of melatonin synthesis and signaling pathways in placental mammals. *Open Res. Eur.* 1:75.

Emerling CA, Widjaja AD, Nguyen NN, Springer MS. 2017. Their loss is our gain: regressive evolution in vertebrates provides genomic models for uncovering human disease loci. *J. Med. Genet.* 54:787–794.

Endo H, Ito K, Watabe H, Nguyen ST, Koyabu D. 2017. Macroscopic and CT Examinations of the Mastication Mechanism in the Southern Tamandua. *Mammal Study* 42:89–96.

Endo H, Niizawa N, Komiya T, Kawada S, Kimura J, Itou T, Koie H, Sakai T. 2007. Three-Dimensional CT Examination of the Mastication System in the Giant Anteater. *Zoolog. Sci.* 24:1005–1011.

Endo H, Nishiumi I, Kurohmaru M, Nabhitabhata J, Chan-Ard T, Nadee N, Agungpriyono S, Yamada J. 1998. The functional anatomy of the masticatory muscles of the Malayan pangolin, *Manis javanica*. *Mammal Study* 23:1–8.

Fenelon JC, Bennetts A, Anthwal N, Pyne M, Johnston SD, Evans AR, Tucker AS, Renfree MB. 2023. Getting out of a mammalian egg: the egg tooth and caruncle of the echidna. *Dev. Biol.* 495:8–18.

Feng P, Zheng J, Rossiter SJ, Wang D, Zhao H. 2014. Massive Losses of Taste Receptor Genes in Toothed and Baleen Whales. *Genome Biol. Evol.* 6:1254–1265.

Ferreira-Cardoso S, Delsuc F, Hautier L. 2019. Evolutionary Tinkering of the Mandibular Canal Linked to Convergent Regression of Teeth in Placental Mammals. *Curr. Biol.* 29:468-475.e3.

Ferreira-Cardoso S, Fabre P-H, De Thoisy B, Delsuc F, Hautier L. 2020. Comparative masticatory myology in anteaters and its implications for interpreting morphological convergence in myrmecophagous placentals. *PeerJ* 8:e9690.

Fisher LW. 2011. DMP1 and DSPP: Evidence for Duplication and Convergent Evolution of Two SIBLING Proteins. *Cells Tissues Organs* 194:113–118.

Flannery TF, Rich TH, Vickers-Rich P, Ziegler T, Veatch EG, Helgen KM. 2022. A review of monotreme (Monotremata) evolution. *Alcheringa Australas. J. Palaeontol.* 46:3–20.

1280 Foley NM, Mason VC, Harris AJ, Bredemeyer KR, Damas J, Lewin HA, Eizirik E, Gatesy J,
1281 Karlsson EK, Lindblad-Toh K, et al. 2023. A genomic timescale for placental mammal
1282 evolution. *Science* 380:eabl8189.
1283 Fong DW, Kane TC, Culver DC. 1995. Vestigialization and Loss of Nonfunctional Characters.
1284 *Annu. Rev. Ecol. Syst.* 2:249–268.
1285 Gaudin TJ, Branham DG. 1998. The Phylogeny of the Myrmecophagidae (Mammalia,
1286 Xenarthra, Vermilingua) and the Relationship of *Eurotamandua* to the Vermilingua. *J.*
1287 *Mamm. Evol.* 5:237–265.
1288 Gaudin TJ, Emry RJ, Wible JR. 2009. The Phylogeny of Living and Extinct Pangolins
1289 (Mammalia, Pholidota) and Associated Taxa: A Morphology Based Analysis. *J. Mamm.*
1290 *Evol.* 16:235–305.
1291 Gaudin TJ, Gaubert P, Billet G, Hautier L, Ferreira-Cardoso S, Wible JR. 2020. Evolution and
1292 morphology. In: Pangolins. Elsevier. p. 5–23. Available from:
1293 https://linkinghub.elsevier.com/retrieve/pii/B9780128155073000010
1294 Gibb GC, Condamine FL, Kuch M, Enk J, Moraes-Barros N, Superina M, Poinar HN, Delsuc F.
1295 2016. Shotgun Mitogenomics Provides a Reference Phylogenetic Framework and
1296 Timescale for Living Xenarthrans. *Mol. Biol. Evol.* 33:621–642.
1297 Glass BP. 1985. History of classification and nomenclature in Xenarthra (Edentata). In: The
1298 evolution and ecology of armadillos, sloths, and vermilinguas. Smithsonian Institution
1299 Press. p. 51–64.
1300 Gong T, Mu Q, Xu Y, Wang W, Meng L, Feng X, Liu W, Ao Z, Zhang Y, Chen X, et al. 2024.
1301 Expression of the umami taste receptor T1R1/T1R3 in porcine testis of: Function in
1302 regulating testosterone synthesis and autophagy in Leydig cells. *J. Steroid Biochem.*
1303 *Mol. Biol.* 236:106429.
1304 Hassanin A, Veron G, Ropiquet A, Jansen Van Vuuren B, Lécu A, Goodman SM, Haider J,
1305 Nguyen TT. 2021. Evolutionary history of Carnivora (Mammalia, Laurasiatheria) inferred
1306 from mitochondrial genomes.Mans BJ, editor. *PLOS ONE* 16:e0240770.
1307 Hautier L, Gomes Rodrigues H, Ferreira-Cardoso S, Emerling CA, Porcher M-L, Asher RJ,
1308 Portela Miguez R, Delsuc F. 2023. From teeth to pad: tooth loss and development of
1309 keratinous structures in sirenians. *Proc. R. Soc. B Biol. Sci.* 290:20231932.
1310 Hayssen V, Ortega J, Morales-Leyva A, Martínez-Mendez N. 2013. *Cabassous centralis*
1311 (Cingulata: Dasypodidae). *Mamm. Species* 898:12–17.
1312 Heighton SP, Allio R, Murienne J, Salmona J, Meng H, Scornacavva C, Bastos ADS, Nijiokou F,
1313 Pietersen DW, Tilak M-K, et al. 2023. Pangolin Genomes Offer Key Insights and
1314 Resources for the World's Most Trafficked Wild Mammals. *Mol. Biol. Evol.* 40:msad190.
1315 Hendey QB. 1974. New fossil carnivores from the Swartkrans australopithecine site (Mammalia:
1316 Carnivora). *Ann. Transvaal Mus.* 29:27–47.
1317 Hoh JFY. 2002. 'Superfast' or masticatory myosin and the evolution of jaw-closing muscles of
1318 vertebrates. *J. Exp. Biol.* 205:2203–2210.
1319 Hoh JFY, Kang LHD, Sieber LG, Lim JHY, Zhong WWH. 2006. Myosin isoforms and fibre types
1320 in jaw-closing muscles of australian marsupials. *J. Comp. Physiol. B* 176:685–695.
1321 Holthaus KB, Steinbinder J, Sachslehner AP, Eckhart L. 2025. Convergent Evolution Has Led to
1322 the Loss of Claw Proteins in Snakes and Worm Lizards.Fraser B, editor. *Genome Biol.*
1323 *Evol.* 17:evae274.
1324 Hu Y, Wu Q, Ma S, Ma T, Shan L, Wang X, Nie Y, Ning Z, Yan L, Xiu Y, et al. 2017.
1325 Comparative genomics reveals convergent evolution between the bamboo-eating giant
1326 and red pandas. *Proc. Natl. Acad. Sci.* 114:1081–1086.
1327 Huelsmann M, Hecker N, Springer MS, Gatesy J, Sharma V, Hiller M. 2019. Genes lost during
1328 the transition from land to water in cetaceans highlight genomic changes associated with
1329 aquatic adaptations. *Sci. Adv.* 5:eaaw6671.
1330 Janiak MC, Chaney ME, Tosi AJ. 2018. Evolution of Acidic Mammalian Chitinase Genes (CHIA)

1331 Is Related to Body Mass and Insectivory in Primates. *Mol. Biol. Evol.* 35:607–622.
1332 Jiang P, Josue J, Li X, Glaser D, Li W, Brand JG, Margolskee RF, Reed DR, Beauchamp GK.
1333 2012. Major taste loss in carnivorous mammals. *Proc. Natl. Acad. Sci.* 109:4956–4961.
1334 Jiao H, Zhang L, Xie H-W, Simmons NB, Liu H, Zhao H. 2019. Trehalase Gene as a Molecular
1335 Signature of Dietary Diversification in Mammals.Satta Y, editor. *Mol. Biol. Evol.* 36:2171–
1336 2183.
1337 Kawasaki K, Hu JC-C, Simmer JP. 2014. Evolution of *Klk4* and enamel maturation in
1338 eutherians. *Biol. Chem.* 395:1003–1013.
1339 Kearse M, Moir R, Wilson A, Stones-Havas S, Cheung M, Sturrock S, Buxton S, Cooper A,
1340 Markowitz S, Duran C, et al. 2012. Geneious Basic: An integrated and extendable
1341 desktop software platform for the organization and analysis of sequence data.
1342 *Bioinformatics* 28:1647–1649.
1343 Kim J-W, Simmer JP, Hart TC, Hart PS, Ramaswami MD, Bartlett JD, Hu JC-C. 2005. MMP-20
1344 mutation in autosomal recessive pigmented hypomaturation amelogenesis imperfecta. *J.*
1345 *Med. Genet.* 42:271–275.
1346 Kim YJ, Lee Y, Kasimoglu Y, Seymen F, Simmer JP, Hu JC-C, Cho E-S, Kim J-W. 2022.
1347 Recessive Mutations in *ACP4* Cause Amelogenesis Imperfecta. *J. Dent. Res.* 101:37–
1348 45.
1349 Kishida T, Thewissen J, Hayakawa T, Imai H, Agata K. 2015. Aquatic adaptation and the
1350 evolution of smell and taste in whales. *Zool. Lett.* 1:9.
1351 Kobayashi K, Han L, Koyama T, Lu S-N, Nishimura T. 2023. Sweet taste receptor subunit T1R3
1352 regulates casein secretion and phosphorylation of STAT5 in mammary epithelial cells.
1353 *Biochim. Biophys. Acta BBA - Mol. Cell Res.* 1870:119448.
1354 Kochem MC, Hanselman EC, Breslin PAS. 2024. Activation and inhibition of the sweet taste
1355 receptor TAS1R2-TAS1R3 differentially affect glucose tolerance in humans.Chisini LA,
1356 editor. *PLOS ONE* 19:e0298239.
1357 Koehler CE, Richardson PRK. 1990. *Proteles cristatus*. *Mamm. Species*:1–6.
1358 Koepfli K-P, Jenks SM, Eizirik E, Zahirpour T, Valkenburgh BV, Wayne RK. 2006. Molecular
1359 systematics of the Hyaenidae: Relationships of a relictual lineage resolved by a
1360 molecular supermatrix. *Mol. Phylogenet. Evol.* 38:603–620.
1361 Koyasu K. 1993. Variability of dentition in the bat-eared fox, *Otocyon megalotis*. *Aichi-Gakuin J.*
1362 *Dent. Sci.* 31:63–104.
1363 Kubota K, Kubota J, Fukuda N, Asakura S. 1962. Comparative anatomical and neurohistological
1364 observations on the tongue of the great anteater (Myrmecophaga Jubata Linne). *Anat.*
1365 *Rec.* 143:15–25.
1366 Kubota K, Kubota J, Nakamura T, Fukuda N, Asakura S, Nakagawa S, Masui M. 1962.
1367 Comparative anatomical and neurohistological observations on the tongue of the
1368 pangolin (*Manis pentadactyla* , Linneus). *Anat. Rec.* 144:43–55.
1369 Kumar V, Lammers F, Bidon T, Pfenninger M, Kolter L, Nilsson MA, Janke A. 2017. The
1370 evolutionary history of bears is characterized by gene flow across species. *Sci. Rep.*
1371 7:46487.
1372 Kupczik K, Stynder DD. 2012. Tooth root morphology as an indicator for dietary specialization in
1373 carnivores (Mammalia: Carnivora): TOOTH ROOT MORPHOLOGY AND DIET. *Biol. J.*
1374 *Linn. Soc.* 105:456–471.
1375 Lagerstrom M, Dahl N, Nakahori Y, Nakagome Y, Backman B, Landegren U, Pettersson U.
1376 1991. A Deletion in the Amelogenin Gene (AMG) Causes X-Linked Amelogenesis
1377 Imperfecta (AIHI). *Genomics* 10:971–975.
1378 Lahti DC, Johnson NA, Ajie BC, Otto SP, Hendry AP, Blumstein DT, Coss RG, Donohue K,
1379 Foster SA. 2009. Relaxed selection in the wild. *Trends Ecol. Evol.* 24:487–496.
1380 Langmead B, Salzberg SL. 2012. Fast gapped-read alignment with Bowtie 2. *Nat. Methods*
1381 9:357–359.

1382 Laurie A, Seidensticker J. 1977. Behavioural ecology of the Sloth bear (*Melursus ursinus*). *J.*
1383 *Zool.* 182:187–204.
1384 Lee LA, Karabina A, Broadwell LJ, Leinwand LA. 2019. The ancient sarcomeric myosins found
1385 in specialized muscles. *Skelet. Muscle* 9:7.
1386 Lehmann T. 2009. Phylogeny and systematics of the Orycteropodidae (Mammalia,
1387 Tubulidentata). *Zool. J. Linn. Soc.* 155:649–702.
1388 Li H. 2018. Minimap2: pairwise alignment for nucleotide sequences.Birol I, editor. *Bioinformatics*
1389 34:3094–3100.
1390 Li J, Guo C, Xie M, Wang K, Wang X, Zou B, Hou F, Ran C, Bi S, Xu Y, et al. 2024. Genomic
1391 signatures of sensory adaptation and evolution in pangolins. *BMC Genomics* 25:1176.
1392 Liang T, Wang S-K, Smith C, Zhang H, Hu Y, Seymen F, Koruyucu M, Kasimoglu Y, Kim J-W,
1393 Zhang C, et al. 2022. Enamel defects in Acp4R110C/R110C mice and human ACP4
1394 mutations. *Sci. Rep.* 12:16477.
1395 Liman ER, Kinnamon SC. 2021. Sour taste: receptors, cells and circuits. *Curr. Opin. Physiol.*
1396 20:8–15.
1397 Liu X, Rich K, Nasseri SM, Li G, Hjæresen S, Finsen B, Scherberger H, Svenningsen Å, Zhang
1398 M. 2023. A Comparison of PKD2L1-Expressing Cerebrospinal Fluid Contacting Neurons
1399 in Spinal Cords of Rodents, Carnivores, and Primates. *Int. J. Mol. Sci.* 24:13582.
1400 Liu Z, Liu G, Hailer F, Orozco-terWengel P, Tan X, Tian J, Yan Z, Zhang B, Li M. 2016. Dietary
1401 specialization drives multiple independent losses and gains in the bitter taste gene
1402 repertoire of Laurasiatherian Mammals. *Front. Zool.* 13:28.
1403 McKnight DA, Fisher LW. 2009. Molecular evolution of dentin phosphoprotein among toothed
1404 and toothless animals. *BMC Evol. Biol.* 9:299.
1405 Meredith RW, Gatesy J, Murphy WJ, Ryder OA, Springer MS. 2009. Molecular Decay of the
1406 Tooth Gene Enamelin (ENAM) Mirrors the Loss of Enamel in the Fossil Record of
1407 Placental Mammals.Malik HS, editor. *PLoS Genet.* 5:e1000634.
1408 Meredith RW, Gatesy J, Springer MS. 2013. Molecular decay of enamel matrix protein genes in
1409 turtles and other edentulous amniotes. *BMC Evol. Biol.* 13:20.
1410 Meredith RW, Zhang G, Gilbert MTP, Jarvis ED, Springer MS. 2014. Evidence for a single loss
1411 of mineralized teeth in the common avian ancestor. *Science* 346:1254390.
1412 Mitchell KJ, Pratt RC, Watson LN, Gibb GC, Llamas B, Kasper M, Edson J, Hopwood B, Male
1413 D, Armstrong KN, et al. 2014. Molecular Phylogeny, Biogeography, and Habitat
1414 Preference Evolution of Marsupials. *Mol. Biol. Evol.* 31:2322–2330.
1415 Murphy WJ, Eizirik E, O'Brien SJ, Madsen O, Scally M, Douady CJ, Teeling E, Ryder OA,
1416 Stanhope MJ, De Jong WW, et al. 2001. Resolution of the Early Placental Mammal
1417 Radiation Using Bayesian Phylogenetics. *Science* 294:2348–2351.
1418 Murray PF. 1981. A Unique Jaw Mechanism in the Echidna, *Tachyglossus aculeatus*
1419 (Monotremata). *Aust. J. Zool.* 29:1–5.
1420 Nakayama Y, Holcroft J, Ganss B. 2015. Enamel Hypomineralization and Structural Defects in
1421 Amelotin-deficient Mice. *J. Dent. Res.* 94:697–705.
1422 Naples VL. 1999. Morphology, evolution and function of feeding in the giant anteater
1423 (*Myrmecophaga tridactyla*). *J. Zool.* 249:19–41.
1424 Nguyen NTT, Vincens P, Dufayard JF, Roest Crollius H, Louis A. 2022. Genomicus in 2022:
1425 comparative tools for thousands of genomes and reconstructed ancestors. *Nucleic Acids*
1426 *Res.* 50:D1025–D1031.
1427 Nishio C, Kuroda S, Moffatt P, Nanci A. 2010. Expression pattern of odontogenic ameloblast-
1428 associated and amelotin during formation and regeneration of the junctional epithelium.
1429 *Eur. Cell. Mater.* 20:393–402.
1430 Parry DA, Brookes SJ, Logan CV, Poulter JA, El-Sayed W, Al-Bahlani S, Al Harasi S, Sayed J,
1431 Raïf EM, Shore RC, et al. 2012. Mutations in C4orf26, Encoding a Peptide with In Vitro
1432 Hydroxyapatite Crystal Nucleation and Growth Activity, Cause Amelogenesis Imperfecta.

Am. J. Hum. Genet. 91:565–571.
Peel E, Silver L, Brandies P, Hayakawa T, Belov K, Hogg CJ. 2022. Genome assembly of the numbat (Myrmecobius fasciatus), the only termitivorous marsupial. *Gigabyte* 2022:1–17.
Pérez-Ramos A, Kupczik K, Van Heteren AH, Rabeder G, Grandal-D'Anglade A, Pastor FJ, Serrano FJ, Figueirido B. 2019. A three-dimensional analysis of tooth-root morphology in living bears and implications for feeding behaviour in the extinct cave bear. *Hist. Biol.* 31:461–473.
Pickford M. 1974. New Fossil Orycteropodidae (Mammalia, Tubulidentata) From East Africa. *Orycteropus Minutus* Sp. Nov. and *Orycteropus Chemeldoi* Sp. Nov. *Neth. J. Zool.* 25:57–88.
Policarpo M, Baldwin MW, Casane D, Salzburger W. 2024. Diversity and evolution of the vertebrate chemoreceptor gene repertoire. *Nat. Commun.* 15:1421.
Poulter JA, Murillo G, Brookes SJ, Smith CEL, Parry DA, Silva S, Kirkham J, Inglehearn CF, Mighell AJ. 2014. Deletion of ameloblastin exon 6 is associated with amelogenesis imperfecta. *Hum. Mol. Genet.* 23:5317–5324.
Prapong T, Liumsiricharoen M, Chungsamarnyart N, Chantakru S, Yatbantoong N, Sujit K, Patumrattanathan P, Pongket P, Duang-ngen A, Suprasert A. 2009. Macroscopic and Microscopic Anatomy of Pangolin's Tongue (*Manis javanica*). *Kasetsart Vet.* 19:9–19.
Rajpar MH, Harley K, Laing C, Davies RM, Dixon MJ. 2001. Mutation of the gene encoding the enamel-specific protein, enamelin, causes autosomal-dominant amelogenesis imperfecta. *Hum. Mol. Genet.* 10:1673–1677.
Randall JG, Gatesy J, McGowen MR, Springer MS. 2024. Molecular Evidence for Relaxed Selection on the Enamel Genes of Toothed Whales (Odontoceti) with Degenerative Enamel Phenotypes. *Genes* 15:228.
Randall JG, Gatesy J, Springer MS. 2022. Molecular evolutionary analyses of tooth genes support sequential loss of enamel and teeth in baleen whales (Mysticeti). *Mol. Phylogenet. Evol.* 171:107463.
Redford KH. 1986. Dietary specialization and variation in two mammalian myrmecophages (variation in mammalian myrmecophagy). *Rev. Chil. Hist. Nat.* 59:201–208.
Redford KH. 1987. Ants and Termites As Food. In: Genoways HH, editor. Current Mammalogy. Boston, MA: Springer US. p. 349–399. Available from: http://link.springer.com/10.1007/978-1-4757-9909-5_9
Reiss KZ. 2001. Using Phylogenies to Study Convergence: The Case of the Ant-Eating Mammals. *Am. Zool.* 41:507–525.
Rismiller PD, Grutzner F. 2019. Tachyglossus aculeatus (Monotremata: Tachyglossidae).Rose RK, Hamilton MJ, editors. *Mamm. Species* 51:75–91.
Roper SD, Chaudhari N. 2017. Taste buds: cells, signals and synapses. *Nat. Rev. Neurosci.* 18:485–497.
Schwenk K. 1985. Occurrence, Distribution and Functional Significance of Taste Buds in Lizards. *Copeia* 1985:91.
Serrano J, Boyd J, Brown IS, Mason C, Smith KR, Karolyi K, Maurya SK, Meshram NN, Serna V, Link GM, et al. 2024. The TAS1R2 G-protein-coupled receptor is an ambient glucose sensor in skeletal muscle that regulates NAD homeostasis and mitochondrial capacity. *Nat. Commun.* 15:4915.
Seymen F, Kim YJ, Lee YJ, Kang J, Kim T-H, Choi H, Koruyucu M, Kasimoglu Y, Tuna EB, Gencay K, et al. 2016. Recessive Mutations in <I>ACPT</I>, Encoding Testicular Acid Phosphatase, Cause Hypoplastic Amelogenesis Imperfecta. *Am. J. Hum. Genet.* 99:1199–1205.
Seymen F, Park J-C, Lee K-E, Lee H-K, Lee D-S, Koruyucu M, Gencay K, Bayram M, Tuna EB, Lee ZH, et al. 2015. Novel *MMP20* and *KLK4* Mutations in Amelogenesis Imperfecta. *J. Dent. Res.* 94:1063–1069.

1484 Shaheen J, Mudd AB, Diekwisch TGH, Abramyan J. 2021. Pseudogenized Amelogenin Reveals
1485 Early Tooth Loss in True Toads (Anura: Bufonidae). *Integr. Comp. Biol.* 61:1933–1945.
1486 Shon W-J, Song JW, Oh SH, Lee K-H, Seong H, You HJ, Seong JK, Shin D-M. 2023. Gut taste
1487 receptor type 1 member 3 is an intrinsic regulator of Western diet-induced intestinal
1488 inflammation. *BMC Med.* 21:165.
1489 Shoshani J, Goldman CA, Thewissen JGM. 1988. Orycteropus afer. *Mamm. Species*:1–8.
1490 da Silveira Anacleto TC. 2007. Food Habits of Four Armadillo Species in the Cerrado Area,
1491 Mato Grosso, Brazil. *Zool. Stud.* 46:529–537.
1492 Smith CEL, Murillo G, Brookes SJ, Poulter JA, Silva S, Kirkham J, Inglehearn CF, Mighell AJ.
1493 2016. Deletion of amelotin exons 3–6 is associated with amelogenesis imperfecta. *Hum.*
1494 *Mol. Genet.* 25:3578–3587.
1495 Smith CEL, Poulter JA, Antanaviciute A, Kirkham J, Brookes SJ, Inglehearn CF, Mighell AJ.
1496 2017. Amelogenesis Imperfecta; Genes, Proteins, and Pathways. *Front. Physiol.* 8:435.
1497 Smith KK, Redford KH. 1990. The anatomy and function of the feeding apparatus in two
1498 armadillos (Dasypoda): anatomy is not destiny. *J. Zool.* 222:27–47.
1499 Springer MS, Emerling CA, Gatesy J, Randall J, Collin MA, Hecker N, Hiller M, Delsuc F. 2019.
1500 Odontogenic ameloblast-associated (ODAM) is inactivated in toothless/enamelless
1501 placental mammals and toothed whales. *BMC Evol. Biol.* 19:31.
1502 Springer MS, Meredith RW, Teeling EC, Murphy WJ. 2013. Technical Comment on “The
1503 Placental Mammal Ancestor and the Post–K-Pg Radiation of Placentals.” *Science*
1504 341:613–613.
1505 Springer MS, Starrett J, Morin PA, Lanzetti A, Hayashi C, Gatesy J. 2016. Inactivation of
1506 C4orf26 in toothless placental mammals. *Mol. Phylogenet. Evol.* 95:34–45.
1507 Stamatakis A. 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of
1508 large phylogenies. *Bioinformatics* 30:1312–1313.
1509 Stayton CT. 2015. What does convergent evolution mean? The interpretation of convergence
1510 and its implications in the search for limits to evolution. *Interface Focus* 5:20150039.
1511 Stedman HH, Kozyak BW, Nelson A, Thesier DM, Su LT, Low DW, Bridges CR, Shrager JB,
1512 Minugh-Purvis N, Mitchell MA. 2004. Myosin gene mutation correlates with anatomical
1513 changes in the human lineage. *Nature* 428:415–418.
1514 Stefen C, Rensberger JM. 1999. The specialized structure of hyaenid enamel: Description and
1515 develpoment within the lineage - including percrocutids. *Scanning Microsc.* 13:363–380.
1516 Storch G. 1978. *Eomanis waldi*, ein Schuppentier aus dem Mittel-Eoza¨n der “Grube Messel”
1517 bei Darmstadt (Mammalia: Pholidota). *Senckenberg. Lethaea* 59:503–529.
1518 Storch G. 1981. *Eurotamandua joresi*, ein Myrmecophagidae aus dem Eoza¨n der “Grube
1519 Messel” bei Darmstadt (Mammalia, Xenarthra). *Senckenberg. Lethaea* 61:247–289.
1520 Superina M, Abba AM. 2014. *Zaedyus pichiy* (Cingulata: Dasypodidae). *Mamm. Species* 905:1–
1521 10.
1522 Thibaud-Nissen F, DiCuccio M, Hlavina W, Kimchi A, Kitts PA, Murphy TD, Pruitt KD, Souvorov
1523 A. 2016. P8008 the NCBI eukaryotic genome annotation pipeline. J Anim Sci 94:184–
1524 184. *J. Anim. Sci.* 94:184.
1525 Thomas VJ, Shaw J, Tay N, Warburton NM. 2024. Comparative three-dimensional jaw muscle
1526 anatomy of marsupial carnivores (*Dasyurus* spp.) and the termite-eating numbat
1527 (*Myrmecobius fasciatus*). *J. Morphol.* 285:e21684.
1528 Toniolo L, Cancellara P, Maccatrozzo L, Patruno M, Mascarello F, Reggiani C. 2008.
1529 Masticatory myosin unveiled: first determination of contractile parameters of muscle
1530 fibers from carnivore jaw muscles. *Am. J. Physiol.-Cell Physiol.* 295:C1535–C1542.
1531 Toyosawa S, Fujiwara T, Ooshima T, Shintani S, Sato A, Ogawa Y, Sobue S, Ijuhin N. 2000.
1532 Cloning and characterization of the human ameloblastin gene. *Gene* 256:1–11.
1533 Tu Y-H, Cooper AJ, Teng B, Chang RB, Artiga DJ, Turner HN, Mulhall EM, Ye W, Smith AD,
1534 Liman ER. 2018. An evolutionarily conserved gene family encodes proton-selective ion

1535 channels. *Science* 359:1047–1050.
1536 Valente R, Alves F, Sousa-Pinto I, Ruivo R, Castro LFC. 2021. Functional or Vestigial? The
1537 Genomics of the Pineal Gland in Xenarthra. *J. Mol. Evol.* 89:565–575.
1538 Wazen R, Ponce K, Kuroda S, Nishio C, Nanci A. 2015. Inactivation of the Odontogenic
1539 ameloblast-associated gene affects the integrity of the junctional epithelium and gingival
1540 healing. *Eur. Cell. Mater.* 30:187–199.
1541 Westerman M, Krajewski C, Kear BP, Meehan L, Meredith RW, Emerling CA, Springer MS.
1542 2016. Phylogenetic relationships of dasyuromorphian marsupials revisited. *Zool. J. Linn.*
1543 *Soc.* 176:686–701.
1544 Xiao S, Yu C, Chou X, Yuan W, Wang Y, Bu L, Fu G, Qian M, Yang J, Shi Y, et al. 2001.
1545 Dentinogenesis imperfecta 1 with or without progressive hearing loss is associated with
1546 distinct mutations in DSPP. *Nat. Genet.* 27:201–204.
1547 Yang Z. 2007. PAML 4: Phylogenetic Analysis by Maximum Likelihood. *Mol. Biol. Evol.*
1548 24:1586–1591.
1549 Yarmolinsky DA, Zuker CS, Ryba NJP. 2009. Common Sense about Taste: From Mammals to
1550 Insects. *Cell* 139:234–244.
1551 Yin D, Zhou R, Yin M, Chen Y, Xu S, Yang G. 2021. Gene Duplication and Loss of *AANAT* in
1552 Mammals Driven by Rhythmic Adaptations.Su B, editor. *Mol. Biol. Evol.* 38:3925–3937.
1553 Young BA. 1997. On the Absence of Taste Buds in Monitor Lizards (*Varanus*) and Snakes. *J.*
1554 *Herpetol.* 31:130.
1555 Zhao H, Xu D, Zhang S, Zhang J. 2012. Genomic and Genetic Evidence for the Loss of Umami
1556 Taste in Bats. *Genome Biol. Evol.* 4:73–79.
1557 Zhao H, Yang J-R, Xu H, Zhang J. 2010. Pseudogenization of the Umami Taste Receptor Gene
1558 Tas1r1 in the Giant Panda Coincided with its Dietary Switch to Bamboo. *Mol. Biol. Evol.*
1559 27:2669–2673.
1560 Zhou Y, Shearwin-Whyatt L, Li J, Song Z, Hayakawa T, Stevens D, Fenelon JC, Peel E, Cheng
1561 Y, Pajpach F, et al. 2021. Platypus and echidna genomes reveal mammalian biology
1562 and evolution. *Nature* 592:756–762.
1563