



HAL
open science

Transcriptomic data reveal divergent paths of chitinase evolution underlying dietary convergence in anteaters and pangolins

Rémi Allio, Sophie Teullet, Dave Lutgen, Amandine Magdeleine, Rachid Koual, Marie-Ka Tilak, Benoît de Thoisy, Christopher A Emerling, Tristan Lefébure, Frédéric Delsuc

► To cite this version:

Rémi Allio, Sophie Teullet, Dave Lutgen, Amandine Magdeleine, Rachid Koual, et al.. Transcriptomic data reveal divergent paths of chitinase evolution underlying dietary convergence in anteaters and pangolins. *Genome Biology and Evolution*, In press, pp.evaf002. 10.1093/gbe/evaf002 . hal-04890530

HAL Id: hal-04890530

<https://hal.science/hal-04890530v1>

Submitted on 16 Jan 2025

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial - NoDerivatives 4.0 International License

1 Transcriptomic data reveal divergent paths of chitinase evolution underlying 2 dietary convergence in anteaters and pangolins

3
4 Rémi Allio^{1,2,§,*}, Sophie Teullet^{1,§}, Dave Lutgen^{1,3,4,§}, Amandine Magdeleine¹, Rachid Koual¹,
5 Marie-Ka Tilak¹, Benoit de Thoisy^{5,6}, Christopher A. Emerling^{1,7}, Tristan Lefébure⁸, and
6 Frédéric Delsuc^{1,*}

7
8 ¹*ISEM, Univ. Montpellier, CNRS, IRD, Montpellier, France*

9 ²*CBGP, INRAE, CIRAD, IRD, Montpellier SupAgro, Univ. Montpellier, Montpellier, France*

10 ³*Institute of Ecology and Evolution, University of Bern, Bern, Switzerland*

11 ⁴*Swiss ornithological Institute, Sempach, Switzerland*

12 ⁵*Institut Pasteur de la Guyane, Cayenne, French Guiana, France*

13 ⁶*Kwata NGO, Cayenne, French Guiana, France*

14 ⁷*Biology Department, Reedley College, Reedley, CA, USA*

15 ⁸*Université Claude Bernard Lyon 1, LEHNA UMR 5023, CNRS, ENTPE, F-69622, Villeurbanne, France*

16 [§]*Equal contribution*

17 18 *Correspondence

19 Rémi Allio: remi.allio@inrae.fr

20 Frédéric Delsuc: frederic.delsuc@umontpellier.fr

21 22 Key words

23 Chitinases, Convergent evolution, Myrmecophagy, Mammals, Salivary glands, Transcriptomics

24 25 Significance Statement

26 This study shows that, despite numerous convergent morphological ant-eating adaptations, pangolins and
27 anteaters have distinct molecular mechanisms for digesting chitin, a key component of their highly
28 specialized diets. This divergence is reflected in their distinctive chitinase gene repertoires and expression

© The Author(s) 2025. Published by Oxford University Press on behalf of Society for Molecular Biology and Evolution. This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial License (<https://creativecommons.org/licenses/by-nc/4.0/>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited. For commercial re-use, please contact reprints@oup.com for reprints and translation rights for reprints. All other permissions can be obtained through our RightsLink service via the Permissions link on the article page on our site—for further information please contact journals.permissions@oup.com.

1 patterns across digestive organs. These findings illustrate how historical contingency has shaped gene
2 family evolution through molecular tinkering in these two convergent ant-eating lineages.

5 **ORCID**

6 Rémi Allio : 0000-0003-3885-5410
7 Sophie Teullet : 0000-0003-2693-1797
8 Dave Lutgen : 0000-0003-0793-3930
9 Amandine Magdeleine : NA
10 Rachid Koual : NA
11 Marie-Ka Tilak : 0000-0001-8995-3462
12 Benoit de Thoisy : 0000-0002-8420-5112
13 Christopher A. Emerling : 0000-0002-7722-7305
14 Tristan Lefebure : 0000-0003-3923-8166
15 Frédéric Delsuc : 0000-0002-6501-6287

17 **Abstract**

18 Ant-eating mammals represent a textbook example of convergent evolution. Among them,
19 anteaters and pangolins exhibit the most extreme convergent phenotypes with complete tooth
20 loss, elongated skulls, protruding tongues, and hypertrophied salivary glands producing large
21 amounts of saliva. However, comparative genomic analyses have shown that anteaters and
22 pangolins differ in their chitinase acidic gene (*CHIA*) repertoires, which potentially degrade the
23 chitinous exoskeletons of ingested ants and termites. While the southern tamandua (*Tamandua*
24 *tetradactyla*) harbors four functional *CHIA* paralogs (*CHIA1-4*), Asian pangolins (*Manis* spp.)
25 have only one functional paralog (*CHIA5*). Here, we performed a comparative transcriptomic
26 analysis of salivary glands in 33 placental species, including 16 novel transcriptomes from ant-
27 eating species and close relatives. Our results suggest that salivary glands play an important role
28 in adaptation to an insect-based diet, as expression of different *CHIA* paralogs is observed in
29 insectivorous species. Furthermore, convergently-evolved pangolins and anteaters express
30 different chitinases in their digestive tracts. In the Malayan pangolin, *CHIA5* is overexpressed in
31 all major digestive organs, whereas in the southern tamandua, all four functional paralogs are
32 expressed, at very high levels for *CHIA1* and *CHIA2* in the pancreas, and for *CHIA3* and *CHIA4*
33 in the salivary glands, stomach, liver, and pancreas. Overall, our results demonstrate that
34 divergent molecular mechanisms within the chitinase acidic gene family underlie convergent

1 adaptation to the ant-eating diet in pangolins and anteaters. This study highlights the role of
2 historical contingency and molecular tinkering of the chitin-digestive enzyme toolkit in this
3 classic example of convergent evolution.
4

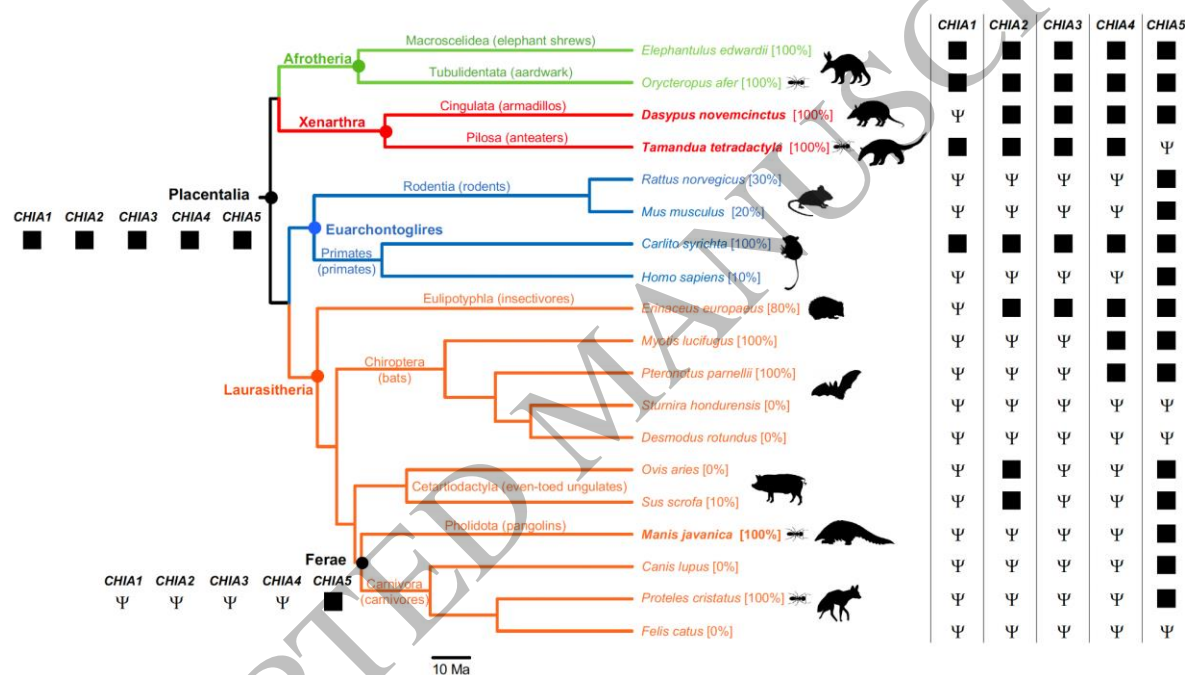
5 **Introduction**

6 Convergent evolution provides a fascinating window into the mechanisms by which similar
7 environmental pressures shape the phenotypes of phylogenetically distant taxa. Indeed, despite
8 the enormous diversity of organisms on Earth and the many potential ways to adapt to similar
9 conditions, the strong deterministic force of natural selection has led to numerous instances of
10 recurrent phenotypic adaptation (Losos 2011; McGhee 2011; Losos 2018). Although classical
11 models of convergence at the molecular level often assume identical mutations in the same genes
12 across species (Arendt and Reznick 2008), emerging evidence from comparative genomics and
13 transcriptomics suggests that the recruitment of the same or similar genes and pathways may also
14 lead to similar phenotypes across divergent lineages. For instance, convergent electric fish,
15 which have evolved independently at least six times, provide a good illustration of the
16 complexity of the selective process that follows from the interaction of contingency, constraints,
17 and convergence (Zakon et al. 2006). In this case, the same genes have been independently
18 recruited and differentially expressed in novel electric organs due to developmental constraints,
19 and their function subsequently adjusted by natural selection involving convergent amino acid
20 substitutions in functionally important domains (Galant et al. 2014; Liu et al. 2019; Wang &
21 Yang 2021). This suggests an important role for evolutionary constraints imposed by existing
22 genomic architectures and developmental pathways, leading to the repeated use of similar
23 genetic material in the origin of evolutionary novelties (Shubin et al. 2009). In this context,
24 historical contingency often leads to evolutionary tinkering as natural selection works from
25 available material (Jacob 1977). Thus, both historical contingency and deterministic evolution
26 appear to have influenced the evolution of current biodiversity, and one of the key questions is to
27 assess the relative influence of these two evolutionary processes (Blount et al. 2018).
28 As intuited by Jacob (1977), molecular tinkering appears to be particularly common and has
29 indeed shaped the evolutionary history of a number of gene families (McGlothlin et al. 2016;
30 Pillai et al. 2020; Xie et al. 2021). The particular evolutionary dynamics observed in gene
31 families can lead to both evolutionary opportunities due to gene duplications paired with the

1 acquisition of a new function but also evolutionary constraints due to ancestral loss of function.
2 A good example resides in the evolution of chitinase genes in placental mammals, which belong
3 to the large Glycosyl Hydrolase 18 (GH18) gene family (Bussink et al. 2007; Funkhouser and
4 Aronson 2007). Recent studies have shown that chitinase genes may play an important digestive
5 function in insectivorous species (Emerling et al. 2018; Janiak et al. 2018; Wang et al. 2020;
6 Cheng et al. 2022). Indeed, while the placental ancestor possessed five functional chitinase acidic
7 (*CHIA*) paralogs, the evolution of this gene family was subsequently shaped through multiple
8 pseudogenization events associated with dietary adaptation during the placental radiation
9 (Emerling et al. 2018). Interestingly, the widespread gene loss observed in carnivorous and
10 herbivorous lineages in particular, resulted in a global positive correlation between the number of
11 functional *CHIA* paralogs and the percentage of invertebrates in the diet across placentals.
12 Indeed, mammals with a low proportion of insects in their diet present none or only a few
13 functional *CHIA* paralogs and those with a high proportion of insects in their diet generally have
14 retained four or five functional *CHIA* paralogs (Emerling et al. 2018; Janiak et al. 2018; Wang et
15 al. 2020; Fig. 1).

16 Myrmecophagous mammals, with more than 90% of their diet consisting of social insects
17 (Redford 1987), have convergently evolved dietary adaptation such as powerful claws used to
18 dig into ant and termite nests, tooth reduction culminating in complete tooth loss in anteaters and
19 pangolins (Ferreira-Cardoso et al. 2019), an elongated muzzle with an extensible tongue
20 (Ferreira-Cardoso et al. 2020), and viscous saliva produced by hypertrophied salivary glands
21 (Reiss 2001). With regards to their chitinase gene repertoire, they are generally grouped with the
22 most insectivorous species (Fig. 1). Specifically, the southern tamandua (*Tamandua*
23 *tetradactyla*) and the armadillo (*Orycteropus afer*) indeed possess four (*CHIA1-4*) and five
24 (*CHIA1-5*) functional paralogs, respectively. However, pangolins appear as a striking exception.
25 Despite their strict myrmecophagous diet and many associated convergent features shared with
26 other myrmecophagous species (anteaters in particular), the two investigated species (*Manis*
27 *javanica* and *M. pentadactyla*) possess only one functional *CHIA* paralog (*CHIA5*). The presence
28 of the sole *CHIA5* in pangolins was hypothesized to be the consequence of historical contingency
29 on the evolution of the chitinase family with the probable loss of *CHIA1-4* functionality in the
30 almost recent common ancestor of Pholidota and Carnivora (Ferae; Emerling et al. 2018; Fig. 1).
31 It has indeed recently been confirmed that a non insect-based diet has caused structural and

1 functional changes in the *CHIA* gene repertoire resulting in multiple losses of function in
 2 Carnivora with only few species including insects in their diet retaining a fully functional *CHIA5*
 3 gene (Tabata et al. 2022). These recent results, combined with the apparent importance of
 4 chitinase paralogs in insect digestion, have prompted questions regarding how pangolins succeed
 5 in digesting chitin with only one functional paralog.



9
 10 **Figure 1:** Dated placental mammal phylogeny including representative species of the four major
 11 clades (Afrotheria, Xenarthra, Euarchontoglires, and Laurasiatheria) for which *CHIA* gene
 12 repertoires have been previously characterized. Numbers between brackets represent
 13 percentages of invertebrates included in the diet with myrmecophagous species indicated by an
 14 ant silhouette. Ψ symbols indicate *CHIA* pseudogenes as determined in previous studies
 15 (Emerling et al. 2018; Janiak et al. 2018; Wang et al. 2020). Ancestral *CHIA* gene repertoires for
 16 Placentalia and Ferae (Pholidota + Carnivora) as inferred by Emerling et al. (2018) are presented.
 17 The chronogram was extracted from www.timetree.org (Kumar et al. 2022). Silhouettes were
 18 obtained from www.phylopic.org.

19
 20 One possible evolutionary solution for inheriting a depleted gene family resides in the
 21 modification of gene expression patterns in the remaining functional paralogs. Indeed, *CHIA5*
 22 was recently found to be highly expressed in the main digestive organs of the Malayan pangolin
 23 (Ma et al. 2017; Ma et al. 2019; Cheng et al. 2023) suggesting that pangolins might compensate

1 for their reduced chitinase repertoire by an increased ubiquitous expression of their only
2 remaining functional *CHIA5* paralog in multiple organs. While this result is very encouraging, it
3 lacks a general comparison with *CHIA* paralogs expression in other mammals and more
4 specifically with other myrmecophagous mammals that present more functional *CHIA* paralogs.
5 If gene expression indeed plays a compensatory role, one can expect that *CHIA5* expression in
6 pangolins would be comparatively higher and more ubiquitous among digestive organs than the
7 expression of the other *CHIA* paralogs in convergent myrmecophagous species.
8 To further explore *CHIA* paralog expression in mammals and more particularly in convergent
9 myrmecophagous species, we adopted a threefold approach. First, with the aim of identifying all
10 functional paralogs and better understanding their function in chitin digestion, we reconstructed
11 the first detailed evolutionary history of the chitinase-like gene family in mammals based on
12 phylogenetic analysis of publicly available genomic and transcriptomic data. In a second step,
13 we generated a large comparative dataset of salivary gland transcriptomes encompassing 33
14 mammalian species from various lineages with diverse diets (herbivores, carnivores, frugivores,
15 insectivores, omnivores), enabling for the first time the comparison of *CHIA* expression across
16 mammalian species. The objective here was to determine whether insectivores and
17 myrmecophagous species indeed exhibit differential chitinase paralog expression in their salivary
18 glands compared to mammals with other diets. In a third step, we focused on two convergent
19 myrmecophagous species (the southern tamandua and the Malayan pangolin) and an
20 insectivorous species (the nine-banded armadillo) for which we were able to assemble and
21 generate transcriptomes of several digestive and non-digestive tissues, to compare the use of
22 their chitinase gene repertoire expression across different organs. The objective of this final step
23 was to determine whether variations in genomic chitinase repertoires were associated with
24 distinct expression patterns in digestive tissues or whether these patterns were independent of the
25 functional gene repertoire. Overall, by leveraging species diversity on the one hand and organ
26 diversity on the other, our results shed light on the molecular underpinnings of convergent
27 evolution in ant-eating mammals by revealing that divergent paths of chitinase gene family
28 evolution underlie dietary convergence between anteaters and pangolins.

29
30
31

1 **Results**

2 **Mammalian chitinase gene family evolution**

3 In order to gain further insights into the evolution and potential function of chitinase-related
4 genes in mammalian genomes, we performed the first detailed phylogenetic reconstruction of the
5 chitinase-like gene family based on functional paralogs using a gene tree/species tree
6 reconciliation approach. The reconciled maximum likelihood tree of mammalian chitinase genes
7 is presented in Fig. 2A. Our analyses showed that this gene family is constituted by nine paralogs
8 whose evolution is notably characterized by gene loss with 384 speciation events followed by
9 gene loss and 48 gene duplications as estimated by the gene tree/species tree reconciliation
10 algorithm of GeneRax. At the base of the reconciled gene tree, we found the clade *CHIA1-*
11 *2/OVGPI* (optimal root inferred by the reconciliation performed with TreeRecs) followed by a
12 duplication separating the *CHIT1/CHI3L1-2* and *CHIA3-5* groups of paralogs. Within the
13 *CHIT1/CHI3L* clade, two consecutive duplications gave rise to *CHIT1*, then *CHI3L1* and
14 *CHI3L2*. In the *CHIA3-5* clade, a first duplication separated *CHIA3* from *CHIA4* and *CHIA5*,
15 which were duplicated subsequently. Marsupial *CHIA4* sequences were located at the base of the
16 *CHIA4-5* clade suggesting that this duplication might be recent and specific to placentals. This
17 scenario of chitinase gene evolution is consistent with our new synteny analysis showing
18 physical proximity of *CHIA1-2* and *OVGPI* on one hand, and *CHIA3-5* on the other hand (Fig.
19 2B), which implies that chitinase genes evolved by successive tandem duplications.

20

1 **Figure 2: A. Mammalian chitinase-like gene family tree reconstructed using a maximum**
 2 **likelihood gene-tree/species-tree reconciliation approach on protein sequences. The nine**
 3 **chitinase paralogs are indicated on the outer circle. Scale bar represents the mean number**
 4 **of amino acid substitutions per site. B. Synteny analysis of the nine chitinase paralogs in**
 5 **humans (*Homo sapiens*), tarsier (*Carlito syrichta*), nine-banded armadillo (*Dasyopus***
 6 **novemcinctus) and the two main focal convergent ant-eating species: the southern**
 7 **tamandua (*Tamandua tetradactyla*) and the Malayan pangolin (*Manis javanica*). Assembly**
 8 **names and accession numbers are indicated below species names. Boxes represent different**
 9 **contigs with their most upstream and downstream BLAST hit positions to chitinase genes**
 10 **(colored arrows). Genes PIFO and DENND2D (grey arrows) are not chitinase paralogs but**
 11 **were used in the synteny analysis. Arrow direction indicates gene transcription direction as**
 12 **inferred in Genomicus v100.01 (Nguyen et al. 2022) for genes located on short contigs. Ψ**
 13 **symbols indicate pseudogenes as determined in Emerling et al. (2018). Genes with non**
 14 **significant BLAST hits were not represented and are probably not functional or absent.**
 15 **Silhouettes were obtained from www.phylopic.org.**
 16

17 Comparison of ancestral sequences

18 The ancestral amino acid sequences of the nine chitinase paralogs were reconstructed from the
 19 reconciled mammalian gene tree and compared to gain further insight into the potential function
 20 of the enzymes they encode (Fig. 3; Complete ancestral sequences and associated probabilities
 21 available from Zenodo). The alignment of predicted amino acid sequences locates the
 22 chitinolytic domain between positions 133 and 140 with the preserved pattern DXXDXDXE.
 23 The ancestral sequences of CHI3L1 and CHI3L2, as all contemporary protein sequences of the
 24 corresponding genes, have a mutated chitinolytic domain with absence of a glutamic acid at
 25 position 140 (Fig. 3A), which is the active proton-donor site necessary for chitin hydrolysis
 26 (Olland et al. 2009; Hamid et al. 2013). This indicates that the ability to degrade chitin has likely
 27 been lost before the duplication leading to CHI3L1 and CHI3L2 (Fig. 3B). The ancestral
 28 sequence of OVGP1 also presents a mutated chitinolytic site although the glutamic acid in
 29 position 140 is present (Fig. 3A). The evolution of the different chitinases therefore seems to be
 30 related to changes in their active site. The six cysteine residues allowing the binding to chitin are
 31 found at positions 371, 418, 445, 455, 457 and 458 (Fig. 3C). The absence of one of these
 32 cysteines prevents binding to chitin (Tjoelker et al., 2000) as this is the case in the ancestral
 33 OVGP1 protein where the last four cysteine residues are changed (Fig. 3C). The other ancestral
 34 sequences present the six conserved cysteine residues and thus can bind to chitin (Fig. 3C).

35
 36

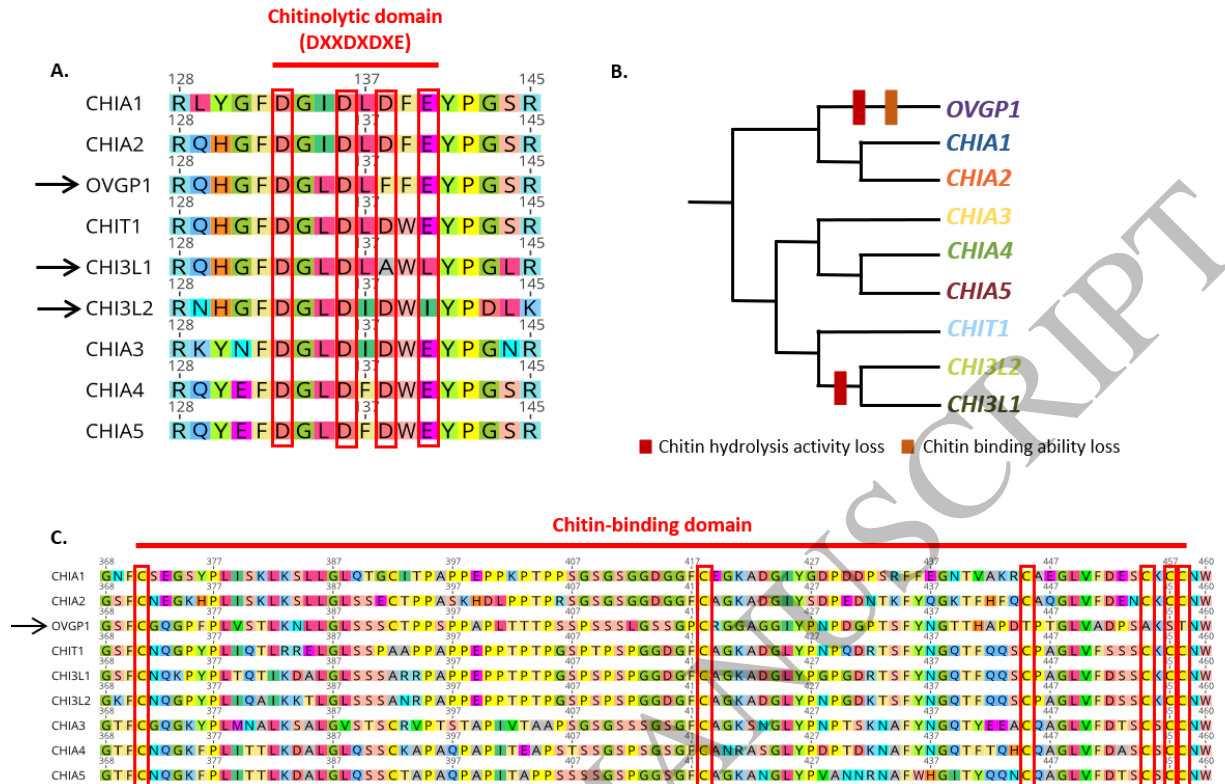
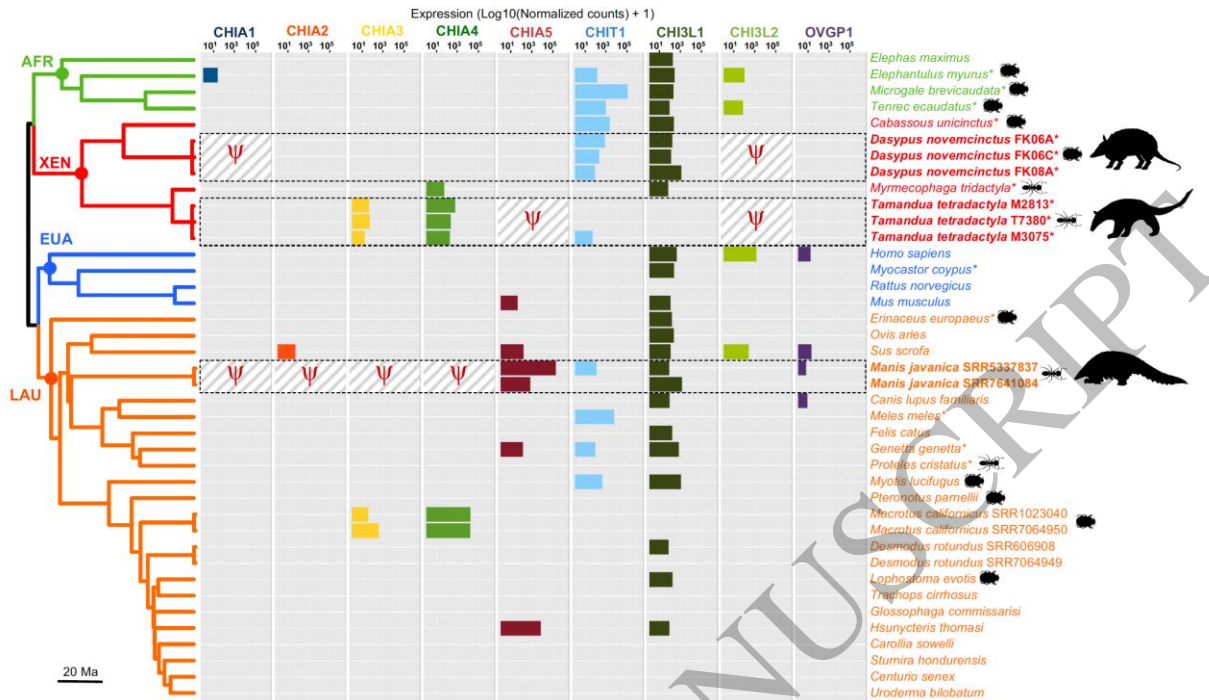


Figure 3: Comparison of predicted ancestral protein sequences of the nine mammalian chitinase paralogs. A. Conserved amino acid residues of the canonical chitinolytic domain active site (DXXDXDXE). Arrows indicate paralogs in which changes occurred in the active site. B. Summary of the evolution of chitinase paralogs functionality. C. Conserved cysteine residues of the chitin-binding domain. The arrow indicates OVGP1 in which the last four cysteines have been replaced.

Chitinase gene expression in mammalian salivary glands

To test the hypothesis that salivary glands play an important functional role in the digestion of ants and termites in ant-eating mammals, we analyzed the gene expression profiles of the nine chitinase paralogs revealed by the gene family tree reconstruction in 40 salivary gland transcriptomes representing 33 species (Fig. 4). CHIA1 was expressed only in the elephant shrew (*Elephantulus myurus*; 23.22 normalized read counts [NC]). CHIA2 was expressed only in the wild boar (*Sus scrofa*; 48.84 NC). CHIA3 was expressed in the two insectivorous California leaf-nosed bat individuals (*Macrotus californicus*; 367.70, and 35.03 NC) and in all three southern tamandua individuals (*T. tetradactyla*; 48.66, 41.52, and 15.14 NC). CHIA4 was also highly expressed in all three southern tamanduas (565.61, 214.83, and 180.26 NC) and in the two

1 California leaf-nosed bats (*M. californicus*; 17,224.06, and 16,880.24 NC), but also in the giant
2 anteater (*M. tridactyla*; 50.74 NC). Expression of CHIA5 was at least an order of magnitude
3 higher in the two Malayan pangolin individuals (*Manis javanica*; 196,778.69 and 729.18 NC)
4 and Thomas's nectar bat (*Hsunycteris thomasi*; 7,301.82 NC) than in the three other species in
5 which we detected expression of this gene: the domestic mouse (*Mus musculus*; 40.15 NC),
6 common genet (*Genetta genetta*; 132.64 NC), and wild boar (*S. scrofa*; 152.20 NC). CHIT1 was
7 expressed in many species (12 out of 40 samples) with values ranging from 46.76 NC in a single
8 southern tamandua (*T. tetradactyla*) individual to 115,739.25 NC in the short-tailed shrew tenrec
9 (*Microgale brevicaudata*). CHI3L1 was expressed in most species (24 out of 40 samples) with
10 values ranging from 61.68 NC in the giant anteater (*M. tridactyla*) to 1,297.01 NC in a Malayan
11 pangolin (*M. javanica*) individual. CHI3L2 was expressed in human (*H. sapiens*; 1334.07 NC),
12 wild boar (*S. scrofa*; 246.41 NC), elephant shrew (*E. myurus*; 94.65 NC), and common tenrec
13 (*Tenrec ecaudatus*; 68.62 NC). OVGP1 was only found expressed at very low levels in domestic
14 dog (*Canis lupus familiaris*; 6.80 NC), human (*H. sapiens*; 15.33 NC), one of the two Malayan
15 pangolins (*M. javanica*; 4.99 NC) and wild boar (*S. scrofa*; 17.84 NC). Finally, the southern
16 aardwolf (*P. cristatus*), Norway rat (*Rattus norvegicus*), Parnell's mustached bat (*Pteronotus*
17 *parnellii*) and six phyllostomid bat species (*Carollia sowelli*, *Centurio senex*, *Glossophaga*
18 *commissarisi*, *Sturnira hondurensis*, *Trachops cirrhosus*, and *Uroderma bilobatum*) did not
19 appear to express any of the nine chitinase gene paralogs in any of our salivary gland samples.
20



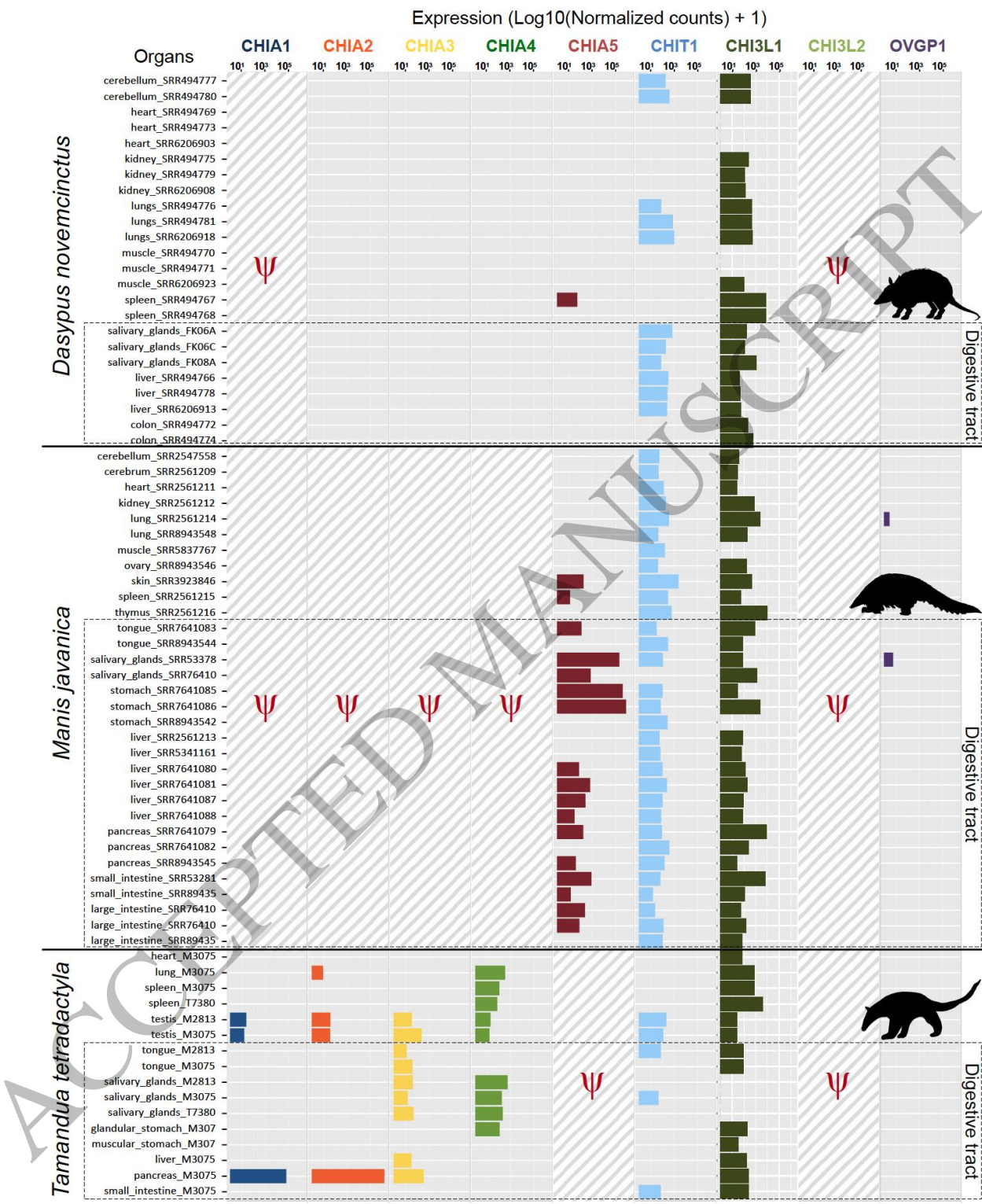
1
 2 **Figure 4: Expression of the nine chitinase paralogs in 40 mammalian salivary gland**
 3 **transcriptomes. The 33 species are presented in their phylogenetic context covering the**
 4 **four major placental clades: Afrotheria (AFR), Xenarthra (XEN), Euarchontoglires (EUA),**
 5 **and Laurasiatheria (LAU). The chronogram was extracted from www.timetree.org**
 6 **(Kumar et al. 2022). Non-functional pseudogenes are only indicated for the three focal**
 7 **species (in bold) using a Ψ symbol: nine-banded armadillo (*Dasyurus novemcinctus*),**
 8 **southern tamandua (*Tamandua tetradactyla*) and Malayan pangolin (*Manis javanica*).**
 9 **Expression level is represented as log₁₀ (Normalized Counts + 1). Asterisks indicate the 16**
 10 **new transcriptomes produced in this study. Myrmecophagous and insectivorous species are**
 11 **indicated by ant and beetle silhouettes, respectively. Silhouettes were obtained from**
 12 **www.phylopic.org.**
 13

14
15 Chitinase gene expression in digestive and non-digestive organs

16 The expression level of the nine chitinase paralogs in several organs was compared among three
 17 species including an insectivorous xenarthran (the nine-banded armadillo; *D. novemcinctus*) and
 18 two of the main convergent myrmecophagous species (the southern anteater; *T. tetradactyla*, and
 19 the Malayan pangolin; *M. javanica*). This analysis revealed marked differences in expression
 20 level of these genes among the three species and among their digestive and non-digestive organs
 21 (Fig. 5). In the nine-banded armadillo (*D. novemcinctus*), although only CHIA1 is
 22 pseudogenized and consequently not expressed, we did not detect any expression of CHIA2,
 23 CHIA3, and CHIA4 in the tissues studied here, and CHIA5 was only weakly expressed in one

1 spleen sample (51.90 NC). In the Malayan pangolin (*M. javanica*), whereas CHIA1-4 are non-
2 functional and consequently not expressed, CHIA5 was found expressed in all digestive organs
3 with particularly high levels in the stomach (377,324.73 and 735,264.20 NC) and salivary glands
4 (196,778.69 and 729.18 NC), and at milder levels in the tongue (121.24 NC), liver (254.79 NC
5 on average when expressed), pancreas (168.64 and 39.33 NC), large intestine (238.45 and 79.32
6 NC), and small intestine (847.51 and 13.72 NC), but also in skin (178.95 NC) and spleen (12.06
7 NC) samples. Conversely, in the southern tamandua (*T. tetradactyla*), only CHIA5 is
8 pseudogenized and accordingly not expressed (Fig. 5). CHIA1 was found highly expressed in the
9 pancreas (64,443.05 NC) and weakly expressed in testes (22.74 and 14.73 NC), and CHIA2 also
10 had very high expression in the pancreas (1,589,834.39 NC), and low expression in testes (36.51
11 and 34.52 NC) and lungs (8.22 NC). CHIA3 was also expressed in the pancreas (359.03 NC),
12 testes (241.79 and 35.42 NC), tongue (39.53 and 12.44 NC), salivary glands (48.66, 41.52, and
13 15.14 NC), and liver (32.40 NC). Finally, CHIA4 was expressed in the testes (19.48 and 14.59
14 NC), spleen (109.97 and 73.31 NC), lungs (340.84 NC), salivary glands (565.61, 214.83, and
15 180.26 NC), and glandular stomach (116.11 NC). More globally, CHIT1 was expressed in all
16 tissues in *M. javanica*, in the testes, tongue, salivary glands, and small intestine in *T. tetradactyla*,
17 and in the cerebellum, lungs, salivary glands, and liver in *D. novemcinctus*. CHI3L1 was found
18 to be expressed in the majority of digestive and non-digestive tissues in all three species.
19 CHI3L2 is non-functional or even absent in the genome of these three species and was
20 consequently not expressed. OVGP1 was only weakly expressed in the lungs and salivary glands
21 of *M. javanica* (2.22 and 4.99 NC, respectively).

22



1
 2 **Figure 5: Expression of the nine chitinase paralogs in 72 transcriptomes from different**
 3 **organs of the three focal species: the nine-banded armadillo (*Dasypos novemcinctus*), the**
 4 **Malayan pangolin (*Manis javanica*), and the southern tamandua (*Tamandua tetradactyla*).**
 5 **Non-functional pseudogenes are represented by a Ψ symbol and hatched background.**

1 **Boxes indicate organs of the digestive tract. Expression level is represented as log₁₀**
2 **(Normalized Counts + 1). Silhouettes were obtained from www.phylopic.org.**
3

4 5 Discussion

6 Evolution of chitinase paralogs towards different functions

7 Chitinases have long been suggested to play an important role in insect digestion within
8 mammals (Jeuniaux 1961; Jeuniaux 1966; Jeuniaux 1971; Jeuniaux and Cornelius 1997). After
9 the initial discovery of a single chitinase gene (Boot et al. 2001), comparative genomics and
10 phylogenetics have revealed a gene family (Glycosyl Hydrolase Family 18, GH18) in which
11 chitinases and chitinase-like proteins may work together to facilitate chitin digestion in the
12 digestive tracts of mammals. The first phylogenetic analyses of this gene family have revealed a
13 dynamic evolutionary history marked by gene duplication and loss following a typical birth-and-
14 death model and a high degree of synteny among mammals (Bussink et al. 2007; Funkhouser
15 and Aronson 2007; Hussain and Wilson 2013). Our new comprehensive maximum likelihood
16 phylogenetic analyses recovered nine functional paralogous chitinase gene sequences in
17 mammalian genomes (Fig. 2A). In addition to the five previously characterized CHIA paralogs
18 (Emerling et al. 2018; Janiak et al. 2018), we were able to include an additional gene (OVGP1),
19 previously identified by Hussain and Wilson (2013), which is most closely related to the
20 previously characterized CHIA1 and CHIA2 genes. In placentals, OVGP1 plays a role in
21 fertilization and embryonic development (Buhi 2002; Saint-Dizier et al. 2014; Algarra et al.
22 2016; Laheri et al. 2018). However, other aliases for OVGP1 include Mucin 9 and CHIT5
23 suggesting a possible digestive function. This result was further confirmed by synteny analyses
24 suggesting a common origin by tandem duplication for CHIA1-2 and OVGP1 within the
25 conserved chromosomal cluster that also includes CHIA3-5 and CHI3L2 (Fig. 2B). Marsupial
26 CHIA4 sequences were located at the base of the CHIA4-5 clade suggesting that this duplication
27 might be recent and specific to placentals. The physical proximity of CHIA3-5 on the same
28 chromosomal fragment implies that these three CHIA genes evolved through successive tandem
29 duplications. However, evidence of concerted evolution by gene conversion potentially
30 misleading phylogenetic reconstruction between the two more recent duplicates (CHIA4 and
31 CHIA5), at least in some taxa (Emerling et al. 2018), suggests that further analyses are needed to
32 fully unravel the origin and function of these two specific paralogs.

1 Comparison of the ancestral amino acid sequences of the nine chitinase paralogs revealed
2 differences in their ability to bind and degrade chitin (Fig. 3), suggesting that these paralogs have
3 evolved towards different functional specializations. The evolution of chitinase-like proteins was
4 accompanied by a loss of enzymatic activity for chitin hydrolysis, which occurred several times
5 independently (Bussink et al. 2007; Funkhouser and Aronson 2007; Hussain and Wilson 2013;
6 Fig. 3B). CHI3L1 and CHI3L2, which are expressed in various cell types including macrophages
7 and synovial cells, play roles in cell proliferation and immune response (Recklies et al. 2002;
8 Areshkov et al. 2011; Lee et al. 2011). In contrast to these chitinase-like proteins, CHIT1 and the
9 five CHIAs are able to degrade chitin. In humans, *CHIT1* is expressed in macrophages and
10 neutrophils and is suspected to be involved in the defense against chitin-containing pathogens
11 such as fungi (Gordon-Thomson et al. 2009; Lee et al. 2011). In addition to their role in chitin
12 digestion (Boot et al. 2001), CHIAs are also suggested to play a role in the inflammatory
13 response (Lee et al. 2011) and are expressed in non-digestive tissues, in agreement with our
14 comparative transcriptomic results. Thus, it has been proposed that the expansion of the chitinase
15 gene family is related to the emergence of the innate and adaptive immune systems in vertebrates
16 (Funkhouser and Aronson 2007).

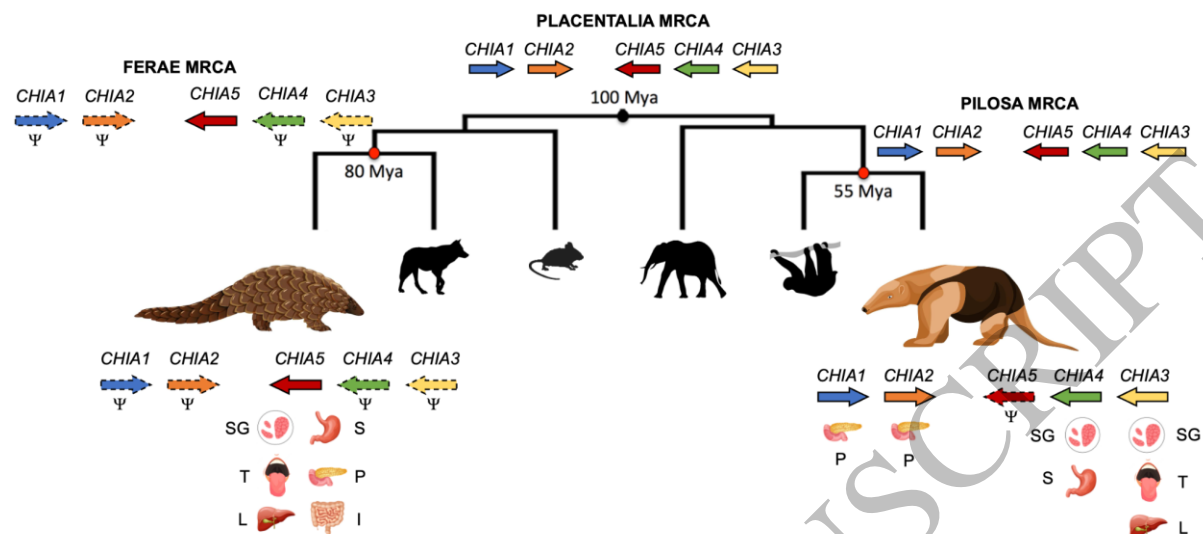
17 The evolution of the different *CHIA1-5* genes has involved changes in their catalytic sites, which
18 have consequences for the secondary structure of enzymes and potentially affect their optimal
19 pH or function, as it has recently been shown for *CHIA5* in Carnivora (Tabata et al. 2022).
20 Experimental testing of the chitin degrading activity of ancestral reconstructions of each of the
21 five CHIA enzymes, on different substrates and at different pH of enzymes, would help
22 determine if there are differences in organ specificity of each enzyme. Furthermore, studying the
23 potential molecular binding properties of these enzymes to other substrates would shed
24 additional light on their functional roles. For example, changing a cysteine in the chitin-binding
25 domain prevents binding to this substrate but not to tri-N-acetyl-chitotriose (Tjoelker et al.
26 2000), a compound derived from chitin with antioxidant properties (Chen et al. 2003; Salgaonkar
27 et al. 2015). Such functional assays, complemented by transcriptomic data to determine their
28 expression profile in different tissues and organs (as previously done in the Malayan pangolin;
29 Yusoff et al. 2016; Ma et al. 2017; Ma et al. 2019; Cheng et al. 2023), may help to decipher their
30 respective roles in mammalian digestion (see below).

31

1 **Impact of historical contingency and molecular tinkering on chitinase evolution and** 2 **expression**

3 In the specific case of adaptation to myrmecophagy, comparative genomic and transcriptomic
4 analyses of these chitinase genes, particularly those encoding chitinolytic enzymes (*CHIAs*),
5 have led to a better understanding of how convergent adaptation to myrmecophagy in placentals
6 occurs at the molecular level (Emerling et al. 2018; Cheng et al. 2022). On the one hand,
7 anteaters (*Pilosa*; *Vermilingua*) likely inherited five *CHIA* genes from an insectivorous ancestor
8 (Emerling et al. 2018), but then the *CHIA5* gene was lost at least in some of its descendants (Fig.
9 6). In the southern tamandua (*T. tetradactyla*), the inactivating mutations of *CHIA5* were
10 identified and the estimated inactivation time of this gene was 6.8 Ma, subsequent to the origin
11 of *Vermilingua* (34.2 Ma) and after the divergence with the giant anteater (*M. tridactyla*) at 11.3
12 Ma, suggesting a loss specific to lesser anteaters of the genus *Tamandua* (Emerling et al. 2018).
13 In our study, this gene was not found to be expressed in the salivary glands of the giant anteater.
14 On the other hand, *CHIA5* is functional in insectivorous carnivores (Carnivora) and pangolins
15 (Pholidota), whereas *CHIA1-4* are pseudogenized (Emerling et al. 2018; Tabata et al. 2022).
16 Similar inactivating mutations have been observed in the *CHIA1* gene in carnivores and
17 pangolins and dated to at least 67 Ma, well before the origin of carnivores (46.2 Ma) and
18 pangolins (26.5 Ma) (Emerling et al. 2018). Thus, despite relying on a fully myrmecophagous
19 diet, pangolins have only one functional *CHIA* gene (*CHIA5*), likely due to a historical
20 contingency related to their common inheritance with carnivores (Fig. 6). These analyses have
21 thus revealed contrasting pseudogenization events between convergent myrmecophagous
22 species, with lesser anteaters (genus *Tamandua*) retaining functional orthologs for four out of the
23 five chitin-degrading *CHIA* genes (*CHIA1-4*), while the Malayan pangolin (*M. javanica*)
24 inherited only the fifth one (*CHIA5*) (Emerling et al. 2018). This peculiar evolutionary history
25 raised the question whether the Malayan pangolin might compensate for the paucity of its
26 functional chitinase gene repertoire by overexpressing *CHIA5* in different digestive organs.

27
28



1
 2 **Figure 6:** Summary figure presenting the evolution and expression of chitinase acidic (*CHIA*)
 3 paralougous genes in the convergently evolved Malayan pangolin (*Manis javanica*) and southern
 4 tamandua (*Tamandua tetradactyla*) in their phylogenetic context. Reconstructed *CHIA* gene
 5 repertoires are indicated for the two myrmecophagous species and for the most recent common
 6 ancestor (MRCA) of placentals, pangolins+carnivores (Ferae) and anteaters+sloths (Pilosa).
 7 Non-functional pseudogenes are represented by the Ψ symbol and dashed line contour. Organ
 8 icons indicate expression of the corresponding gene in different digestive organs. SG: Salivary
 9 glands; S: Stomach; T: Tongue; P: Pancreas; L: Liver; I: Intestine. Silhouettes were obtained
 10 from www.phylopic.org and www.vecteezy.com.

11
 12
 13 Since the presence of enlarged salivary glands is a hallmark of ant-eating mammals, ensuring
 14 massive production of saliva to help catch and potentially digest prey, we first investigated
 15 chitinase gene expression in mammalian salivary glands. Our comparative transcriptomic study
 16 spanning a diversity of species with different diets revealed that, among ant-eating mammals, the
 17 Malayan pangolin (*M. javanica*), the southern tamandua (*T. tetradactyla*), and the giant anteater
 18 (*M. tridactyla*) all express one or more chitin-degrading genes in their salivary glands. More
 19 specifically, we found that *CHIA1* and *CHIA2* were almost never expressed in mammalian
 20 salivary glands. By contrast, *CHIA4* was found to be expressed in the giant anteater (*M.*
 21 *tridactyla*) and expression of both *CHIA3* and *CHIA4* was observed in the three southern
 22 tamandua (*T. tetradactyla*) individuals surveyed. Moreover, we were able to confirm the
 23 hypothesis implying an overexpression of the only functional *CHIA* gene possessed by the
 24 Malayan pangolin. Indeed, salivary gland expression profiles of *CHIA5* in *M. javanica* were

1 much higher than in the four other species (Thomas's nectar bat, mouse, genet, and wild boar) in
2 which we detected expression of this gene, but also substantially higher than the expression of
3 any other chitin-degrading *CHIA* in the 32 other mammalian species considered. Finally, apart
4 from anteaters, *CHIA3* and *CHIA4* were found to be highly expressed only in the two individuals
5 of the insectivorous California leaf-nosed bat (*M. californicus*), but not in any of the other 11
6 examined bat species, including insectivorous species such as *M. myotis*, *P. parnellii*, and *L.*
7 *evotis*. A possible explanation is that these genes have been pseudogenized in many of these bat
8 species, which would be concordant with the findings of comparative genomic studies reporting
9 widespread pseudogenizations of *CHIA* paralogs across multiple bat species (Emerling et al.
10 2018), with complete loss of *CHIA1-5* function in non-insectivorous old world fruit bats, most
11 frugivorous bats, and the sanguivorous common vampire bat (Wang et al. 2020). However,
12 although *CHIA4* and *CHIA5* appear to be functional in the insectivorous little brown myotis (*M.*
13 *lucifugus*; Emerling et al. 2018; Wang et al. 2020), we did not observe expression of these genes
14 in the salivary gland transcriptome we analyzed. Also, *CHIA5* was found to be highly expressed
15 in Thomas's nectar bat (*H. thomasi*). Although this bat species feeds mostly on nectar and fruits,
16 its diet also includes a substantial part of insects suggesting that *CHIA5* might play a role in
17 chitin digestion in the oral cavity, as a result of salivary gland secretion. Transcriptomic analyses
18 of additional digestive tissues besides salivary glands in bats (Vandewege et al. 2020) may
19 further clarify this pattern since chitinolytic activity has previously been reported in the stomachs
20 of seven insectivorous bat species (Strobel et al. 2013). Overall, our chitinase gene expression
21 results therefore support a primary role for salivary glands in prey digestion through the use of
22 distinct *CHIA* paralogs (*CHIA3*, *CHIA4*, and *CHIA5*) in different insect-eating placental mammal
23 species.

24 Our differential expression comparison of the distinct chitinase paralogs across different organs
25 further highlight the importance of *CHIA5* for Malayan pangolin digestive physiology by
26 confirming its ubiquitous expression in all major tissues of the digestive tract (tongue, salivary
27 glands, stomach, pancreas, liver, and large and small intestines) (Ma et al. 2017; Ma et al. 2019;
28 Cheng et al. 2023; Fig. 6). More specifically, *CHIA5* was found to be expressed at particularly
29 high levels in the stomach and salivary glands. These results are in line with previous proteomic
30 studies that have also identified *CHIA5* as a digestive enzyme (Zhang et al. 2019), which has
31 been confirmed to be highly expressed by RT-qPCR in the specialized oxyntic glands of the

1 stomach (Ma et al. 2018; Cheng et al. 2023), reflecting a key adaptation of the Malayan pangolin
2 to its strictly myrmecophagous diet. By contrast, in the southern tamandua (*T. tetradactyla*) only
3 *CHIA5* is pseudogenized (Emerling et al. 2018; Cheng et al. 2023) and all functional *CHIAs*
4 were found expressed in its digestive tract but not in the same tissues (Fig. 6). *CHIA1* and *CHIA2*
5 were particularly highly expressed in the pancreas whereas *CHIA3* and *CHIA4* were expressed
6 across several other organs of the digestive tract including tongue, salivary glands, stomach, and
7 liver. *CHIA1-4* were also expressed in other non-digestive organs (testes, lungs, and spleen), but
8 their co-expression in the salivary glands of the three southern tamandua individuals sampled
9 here strongly suggests that they play a crucial role in chitin digestion in this myrmecophagous
10 species. Conversely, in the less specialized insectivorous nine-banded armadillo (*D.*
11 *novemcinctus*), although only *CHIA1* is pseudogenized (Emerling et al. 2018) and therefore not
12 expressed, we did not detect any expression of *CHIA2*, *CHIA3*, and *CHIA4* in the diverse tissues
13 of the individuals studied here, including salivary glands, and *CHIA5* was found only weakly
14 expressed in one spleen sample. Yet, chitinases could still participate in prey digestion in the
15 nine-banded armadillo as they have been isolated from gastric tissues (Smith et al. 1998). We
16 could not confirm this result, given that the liver and colon were the only additional digestive
17 organs besides salivary glands represented in our dataset for this species. However, the
18 comparison with the two myrmecophagous species seems to fit well with its less specialized
19 insectivorous diet and actually further underscores the contrasted specific use of distinct *CHIA*
20 paralogs for chitin digestion in anteaters and pangolins.

21 Our results demonstrate that in the case of the southern tamandua (*T. tetradactyla*) and the
22 Malayan pangolin (*M. javanica*), two myrmecophagous species that diverged about 100 Ma ago
23 (Meredith et al. 2011), convergent adaptation to myrmecophagy has been achieved in part by
24 using paralogs of different chitinase genes to digest chitin (Fig. 6), probably due to phylogenetic
25 constraints leading to the loss of *CHIA1*, *CHIA2*, *CHIA3*, and *CHIA4* in the most recent common
26 ancestor of Ferae (Carnivora and Pholidota; Emerling et al. 2018). Pangolins and anteaters
27 present extreme morphological adaptations, including the complete loss of dentition, but a
28 detailed study of their feeding apparatus has shown that convergent tooth loss resulted in
29 divergent structures in the internal morphology of their mandible (Ferreira-Cardoso et al. 2019).
30 Our results combined with this observation clearly show that the evolution of convergent
31 phenotypes in myrmecophagous mammals does not necessarily imply similar underlying

1 mechanisms. Our study shows that historical contingency resulted in molecular tinkering (sensu
2 Jacob 1977) of the chitinase gene family at both the genomic and transcriptomic levels in
3 convergently evolved anteaters and pangolins. Working from different starting materials (*i.e.*
4 different *CHIA* paralogs), natural selection led pangolins and anteaters to follow different paths
5 in their convergent adaptation to the myrmecophagous diet.

6
7 Insights from paralogous gene expression in comparative transcriptomic studies
8 Conducting comparative transcriptomic studies between phylogenetically distant species is
9 challenging (Dunn et. al. 2013; Roux et. al. 2015). Confounding factors include, but are not
10 limited to, inconsistencies in tissue sampling methodology, timing of sample collection, and
11 differences in tissue preservation methods. Any of these factors could significantly affect the
12 quality of transcriptome assembly and quantification of gene expression profiles, which could
13 affect underlying expression patterns related to phylogenetic or dietary similarities between
14 species. Recognizing these challenges, we applied state-of-the-art assembly and annotation
15 methods and performed rigorous normalization of read counts to mitigate some of the
16 interspecific variability. Moreover, despite the interest in looking at overall expression patterns
17 to identify the main effect associated with gene expression variation, exploratory comparative
18 transcriptomic analyses also have some limitations. Indeed, when comparing the overall gene
19 expression pattern of different species, the first step is to identify comparable elements. These
20 comparable elements can be restricted to single-copy orthologs or extended to homologous gene
21 families containing different paralogs. However, some biases may be introduced during this step
22 (see Li et al. 2023 for a review). On the one hand, focusing only on orthologous genes
23 completely neglects the effects of paralogous gene expression. On the other hand, working with
24 at the scale of large homologous families (orthogroups) often leads to summarizing the
25 expression of multiple orthologous genes into a single expression value. In our case, for
26 example, following the orthogroup detection and summarizing the expression for each
27 orthogroup would have led to a single expression value for the entire chitinase gene family
28 (found as a single orthogroup). By contrast, thanks to our detailed investigation of the evolution
29 of this gene family, phylogenetic and expression analyses of the chitinase orthogroups revealed
30 interesting patterns that would have been missed by the global approach (*i.e.* effect of
31 contingency bypassed by the relative expression of chitinase family genes). In particular, this

1 approach highlighted differences in gene expression between closely related paralogs (i.e.
2 *CHIA*s) in the digestive organs of the southern tamandua and the Malayan pangolin, which was
3 crucial for our understanding of the molecular mechanisms involved in this case of convergent
4 dietary adaptation. This result underscores the importance of using both genome- and
5 transcriptome-wide analyses to identify novel candidate genes influencing specific traits, and
6 more targeted approaches based on existing knowledge. The latter is essential to deepen our
7 understanding of the underlying mechanisms observed in specific cases, such as those of
8 convergent evolution linked to historical contingency, as explored in this study.

11 **Material and Methods**

12 **Chitinase gene family tree reconstruction**

13 *Reconstruction of chitinase gene family evolution* - Mammalian sequences similar to the protein
14 sequence of the human *CHIA* chitinase acidic gene (NP_970615.2) were searched in the NCBI
15 non-redundant protein database using BLASTP (E-value < 10). The protein sequences identified
16 by BLASTP (n = 1,476) were then aligned using MAFFT v7.450 (Kato and Standley 2013)
17 with the following parameters (--auto --op 1.53 --ep 0.123 --aamatrix BLOSUM62). Preliminary
18 gene trees were then reconstructed with maximum likelihood using RAxML v8.2.11 (Stamatakis
19 2014) under the LG+G4 model (Le and Gascuel 2008). From the reconstructed tree, the
20 sequences were filtered according to the following criteria: (1) fast-evolving sequences with a
21 BLAST E-value greater than zero and not belonging to the chitinase family were excluded; (2) in
22 cases of multiple isoforms, only the longest was retained; (3) sequences whose length
23 represented less than at least 50% of the total alignment length were removed; (4) in case of
24 identical sequences of different lengths from the same species the longest was kept; and (5)
25 sequences labeled as "Hypothetical protein" and "Predicted: low quality protein" were discarded.
26 This procedure resulted in a dataset containing 528 mammalian sequences that were realigned
27 using MAFFT with the following parameters (--auto --op 1.53 --ep 0.123 --aamatrix
28 BLOSUM62). This alignment contained 581 amino acid positions and was then cleaned up by
29 removing sites not present in at least 50% of the sequences resulting in a total length of 460
30 amino acid sites. A maximum likelihood tree was then reconstructed with RAxML-NG v0.9.0
31 (Kozlov et al. 2019) using 10 tree searches starting from maximum parsimony trees under the

1 LG+G8+F model. The species tree of the 143 mammal species represented in our dataset was
2 reconstructed based on *COI* sequences extracted from the BOLD system database v4
3 (Ratnasingham and Hebert 2007) by searching for “Chordata” sequences in the “Taxonomy”
4 section. Sequences were aligned using MAFFT with the following parameters (--auto --op 1.53 -
5 -ep 0.123 --aamatrix BLOSUM62), the phylogeny was inferred with RAxML under the
6 GTR+G4 model and the topology was then adjusted manually based on the literature to correct
7 ancient relationships. To determine the optimal rooting scheme, a rapid reconciliation between
8 the resulting gene tree and species tree was performed using the TreeRecs reconciliation
9 algorithm based on maximum parsimony (Comte et al. 2020) as implemented in SeaView v5.0.2
10 (Gouy et al. 2010). The final chitinase gene family tree was produced using the maximum
11 likelihood gene family tree reconciliation approach implemented in GeneRax v.1.1.0 (Morel et
12 al. 2020) using the TreeRecs reconciled tree as input (source and result available from Zenodo).
13 GeneRax can reconstruct duplications, losses, and horizontal gene transfer events but since the
14 latter are negligible in mammals, only gene duplications and losses have been modeled here (--
15 rec-model UndatedDL) and the LG+G model was used.

16
17 *Ancestral sequence reconstructions* - Ancestral sequences of the different paralogs were
18 reconstructed from the reconciled tree using RAxML-NG (--ancestral function --model
19 LG+G8+F). The sequences were then aligned with MAFFT with the following parameters (--
20 auto --op 1.53 --ep 0.123 --aamatrix BLOSUM62) (source and result files available from
21 Zenodo). Given that active chitinases are characterized by a catalytic site with a conserved amino
22 acid motif (DXXDXDXE; Olland et al. 2009; Hamid et al. 2013), this motif was compared
23 among all available species. Additionally, the six conserved cysteine residues responsible for
24 chitin binding (Tjoelker et al. 2000; Olland et al. 2009) were also investigated.

25
26 *Chitinase gene synteny comparisons* - The synteny of the nine chitinase paralogs was compared
27 between the two focal ant-eating species in our global transcriptomic analysis (*T. tetradactyla*
28 and *M. javanica*), an insectivorous xenarthran species (*D. novemcinctus*), an insectivorous
29 primate species with five functional *CHIA* genes (*Carlito syrichta*), and human (*Homo sapiens*).
30 For *H. sapiens*, synteny information was added from Emerling et al. (2018) and completed by
31 using Genomicus v100.01 (Nguyen et al. 2022). For *C. syrichta* and *D. novemcinctus*, genome

1 assemblies were downloaded from the National Center for Biotechnology Information (NCBI)
2 and from the DNA Zoo (Choo et al. 2016; Dudchenko et al. 2017) for *M. javanica* and *T.*
3 *tetradactyla*. Synteny information was retrieved by blasting (*megablast*) the different CDS
4 sequences against these assemblies. Scaffold/contig names, positions and direction of BLAST
5 hits were retrieved to compare their synteny (source and result files available from Zenodo).
6 Genes with no significant BLAST hits were considered probably not functional or absent.

8 **Transcriptome assemblies**

9 *Salivary gland transcriptomes* - Biopsies of submandibular salivary glands (Gil et al. 2018)
10 preserved in RNAlater were obtained from the Mammalian Tissue Collection of the Institut des
11 Sciences de l'Evolution de Montpellier (ISEM) and the JAGUARS collection for 16 individuals
12 representing 12 placental mammal species (Table S1). Total RNA was extracted from individual
13 salivary gland tissue samples using the RNeasy extraction kit (Qiagen, Germany). Then, RNA-
14 seq library construction and Illumina sequencing on a HiSeq 2500 system using paired-end
15 2x125bp reads were conducted by the Montpellier GenomiX platform (MGX) resulting in 16
16 newly produced salivary gland transcriptomes. This sampling was completed with the 26
17 mammalian salivary gland transcriptomes available as paired-end Illumina sequencing reads in
18 the Short Read Archive (SRA) of the NCBI as of December 15th, 2022 representing an
19 additional 21 species (Table S1). This taxon sampling includes representatives from all major
20 mammal superorders Afrotheria (n = 4), Xenarthra (n = 4), Euarchontoglires (n = 4), and
21 Laurasiatheria (n = 21) and covers six different diet categories: carnivory (n = 4), frugivory and
22 herbivory (n = 8), insectivory (n = 9), myrmecophagy (n = 5), and omnivory (n = 7) (Table S1).
23 Four of the five lineages in which myrmecophagous mammals evolved are represented: southern
24 aardwolf (*P. cristatus*, Carnivora), Malayan pangolin (*M. javanica*, Pholidota), southern naked-
25 tailed armadillo (*C. unicinctus*, Cingulata), giant anteater (*M. tridactyla*, Pilosa), and southern
26 tamandua (*T. tetradactyla*, Pilosa). Species replicates in the form of different individuals were
27 included for the southern tamandua (*T. tetradactyla*; n = 3), the nine-banded armadillo (*D.*
28 *novemcinctus*; n = 3), the Malayan pangolin (*M. javanica*; n = 2), the vampire bat (*Desmodus*
29 *rotundus*; n = 2), and the California leaf-nosed bat (*Macrotus californicus*; n = 2). We
30 unfortunately were not able to obtain fresh salivary gland samples from the aardvark (*O. afer*,
31 Tubulidentata), the only missing myrmecophagous lineage in our sampling.

1
2 *Transcriptomes from additional organs* - Tissue biopsies from nine additional organs (testis,
3 lungs, heart, spleen, tongue, pancreas, stomach, liver, and small intestine) were sampled during
4 dissections of three roadkill individuals of southern tamandua (*T. tetradactyla*; Table S1). Total
5 RNA extractions from these RNAlater-preserved tissues, RNA-seq library construction, and
6 sequencing were conducted as described above resulting in 13 newly generated transcriptomes.
7 For comparative purposes, 21 additional transcriptomes of nine-banded armadillo (*D.*
8 *novemcinctus*) representing eight organs and 32 transcriptomes of Malayan pangolin (*M.*
9 *javanica*) representing 16 organs were downloaded from SRA (Table S1).

10

11 **Comparative transcriptomics**

12 *Transcriptome assemblies and quality control* - Adapters and low quality reads were removed
13 from raw sequencing data using fastp v0.19.6 (Chen et al. 2018). Reads were allowed a
14 minimum of 40% of bases with a PHRED score at least 15 (`--qualified_quality_phred ≥ 15`), as
15 suggested by (MacManes 2014). Then, *de novo* assembly was performed on each individual
16 transcriptome sample using Trinity v2.8.4 (Grabherr et al. 2011) using cleaned paired-end reads
17 (`--seqType fq --left R1.fastq --right R2.fastq`; result files available from Zenodo). For one
18 individual vampire bat (*D. rotundus*), three salivary gland transcriptomes (SRR606902,
19 SRR606908, and SRR606911) were combined to obtain a better assembly. For each of the 104
20 transcriptome assemblies, completeness was assessed by the presence of Benchmark Universal
21 Single Copy Orthologs (BUSCO v5) based on a predefined dataset (mammalia_odb10) of 9,226
22 single-copy orthologs conserved in over 90% of mammalian species (Manni et al. 2021). This
23 pipeline was run through the gVolante web server (Nishimura et al. 2017) to evaluate the
24 percentage of complete, duplicated, fragmented and missing single copy orthologs within each
25 transcriptome (Table S2).

26

27 *Transcriptome annotation and orthogroup inference* - The 104 transcriptome assemblies were
28 annotated following the pipeline implemented in assembly2ORF
29 (<https://github.com/ellefeg/assembly2orf>). This pipeline combines evidence-based and gene-
30 model-based predictions. First, potential transcripts of protein-coding genes are extracted based
31 on similarity searches (BLAST) against the peptides of Metazoa found in Ensembl (Yates et al.

1 2020). Then, using both protein similarity and exonerate functions (Slater and Birney 2005), a
2 frameshift correction is applied to candidate transcripts. Candidate open reading frames (ORFs)
3 are predicted using TransDecoder (<https://github.com/TransDecoder/TransDecoder>) and
4 annotated based on homology information inferred from both BLAST and Hmmscan searches.
5 Finally, to be able to compare the transcriptomes obtained from all species, we relied on the
6 inference of gene orthogroups. The orthogroup inference for the translated candidate ORFs was
7 performed using OrthoFinder v2 (Emms and Kelly 2019) using FastTree (Price et al. 2010) for
8 gene tree reconstructions. For expression analyses, orthogroups containing more than 20 copies
9 for at least one species were discarded, resulting in the selection of 13,392 orthogroups for
10 further analyses.

11
12 *Gene expression analyzes* - Quantification of transcript expression was performed on Trinity
13 assemblies with Kallisto v.0.46.1 (Bray et al. 2016) using the *align_and_estimate_abundance.pl*
14 script provided in the Trinity suite (Grabherr et al. 2011). Kallisto relies on pseudo-alignments of
15 the reads to search for the original transcript of a read without looking for a perfect alignment (as
16 opposed to classical quantification by counting the reads aligned on the assembled transcriptome;
17 Wolf 2013). Counts (raw number of mapped reads) and the Transcripts Per kilobase Million are
18 reported (result files available from Zenodo). Based on the previously inferred orthogroups,
19 orthogroup-level abundance estimates were imported and summarized using tximport (Soneson
20 et al. 2016). To minimize sequencing depth variation across samples and gene outlier effect (a
21 few highly and differentially expressed genes may have strong and global influence on every
22 gene read count), orthogroup-level raw reads counts were normalized using the median of the
23 ratios of observed counts using DESeq2 (Love et al. 2014).

24
25 *Chitinase expression in salivary glands* - The chitinase orthogroup was extracted from the
26 orthogroups inferred by OrthoFinder2 using BLASTX with the reference chitinase database
27 previously created. The 476 amino acid sequences composing this orthogroup were assigned to
28 the nine chitinase orthologs (*CHIA1-5*, *CHIT1*, *CHI3L1*, *CHI3L2*, *OVGP1*) using the maximum
29 likelihood Evolutionary Placement Algorithm implemented in RAxML-EPA (Berger et al. 2011)
30 with the reference chitinase sequence alignment and reconciled phylogenetic tree previously
31 inferred using GeneRax (result files available from Zenodo). This allowed excluding three

1 additional contaminant sequences and dividing the chitinase orthogroup into nine sub-
2 orthogroups corresponding to each chitinase paralog. To take advantage of the transcriptome-
3 wide expression information for the expression standardization, these new orthogroups were
4 included in the previous orthogroup-level abundance matrix estimates and the same
5 normalization approach using DESeq2 was conducted. Finally, gene-level abundance estimates
6 for all chitinase paralogs were extracted and compared on a log₁₀ scale.

9 **Data and Resource Availability**

10 Raw RNAseq Illumina reads have been submitted to the Short Read Archive (SRA) of the
11 National Center for Biotechnology Information (NCBI) and are available under BioProject
12 number PRJNA909065. Transcriptome assemblies, phylogenetic datasets, corresponding trees,
13 and other supplementary materials are available from zenodo.org
14 (<https://doi.org/10.5281/zenodo.7355329>).

17 **Acknowledgments**

18 We would like to thank Hugues Parrinello (Montpellier GenomiX platform) for advice on
19 RNAseq, Mariana Escobar Rodríguez and Gautier Debaecker for help with transcriptome
20 assembly and annotation, and Marie Sémon for providing useful advice on RNAseq statistical
21 analyses. We are also indebted to Frank Knight, Mark Scherz, Miguel Vences, Andolalao
22 Rakotoarison, Nico Avenant, Pierre-Henri Fabre, Quentin Martinez, Nathalie Delsuc, Aude
23 Caizergues, Roxanne Schaub, Lionel Hautier, Fabien Condamine, Sérgio Ferreira-Cardoso, and
24 François Catzeflis for their help with tissue sampling. We also thank the four anonymous
25 referees for their helpful comments. Computational analyses benefited from the Montpellier
26 Bioinformatics Biodiversity (MBB) platform. The JAGUARS collection is supported through a
27 grants from the Collectivité Territoriale de Guyane, from the European Union, and from
28 Direction Générale des Territoires et de la Mer / Préfet de la Région Guyane attributed to Kwata
29 NGO. This work has been supported by grants from the European Research Council
30 (ConvergeAnt project: ERC-2015-CoG-683257) and Investissements d'Avenir of the Agence

1 Nationale de la Recherche (CEBA: ANR-10-LABX-25-01; CEMEB: ANR-10-LABX-0004).
2 This is contribution ISEM 2024-313 of the Institut des Sciences de l'Evolution de Montpellier.

5 **References**

- 6 Algarra B, Han L, Soriano-Úbeda C, Avilés M, Coy P, Jovine L, Jiménez-Movilla M. 2016. The
7 C-terminal region of OVGP1 remodels the zona pellucida and modifies fertility parameters. *Sci.*
8 *Rep.* 6:32556.
- 9 Allio R, Tilak M-K, Scornavacca C, Avenant NL, Kitchener AC, Corre E, Nabholz B, Delsuc F.
10 2021. High-quality carnivoran genomes from roadkill samples enable comparative species
11 delineation in aardwolf and bat-eared fox. *eLife* 10:e63167.
- 12 Arendt J, Reznick D. 2008. Convergence and parallelism reconsidered: What have we learned
13 about the genetics of adaptation? *Trends Ecol. Evol.* 23:26–32.
- 14 Areshkov PO, Avdieiev SS, Balynska OV, LeRoith D, Kavsan VM. 2011. Two closely related
15 human members of Chitinase-like family, CHI3L1 and CHI3L2, activate ERK1/2 in 293 and
16 U373 cells but have the different influence on cell proliferation. *Int. J. Biol. Sci.* 8:39–48.
- 17 Berger SA, Krompass D, Stamatakis A. 2011. Performance, accuracy, and web server for
18 evolutionary placement of short sequence reads under maximum likelihood. *Syst. Biol.* 60:291–
19 302.
- 20 Blount ZD, Lenski RE, Losos JB. 2018. Contingency and determinism in evolution: Replaying
21 life's tape. *Science* 362:eaam5979.
- 22 Boot RG, Blommaert EF, Swart E, Ghauharali-van der Vlugt K, Bijl N, Moe C, Place A, Aerts
23 JM. 2001. Identification of a novel acidic mammalian chitinase distinct from chitotriosidase. *J.*
24 *Biol. Chem.* 276:6770–6778.
- 25 Bray NL, Pimentel H, Melsted P, Pachter L. 2016. Near-optimal probabilistic RNA-seq
26 quantification. *Nat. Biotechnol.* 34:525–527.
- 27 Buhi WC. 2002. Characterization and biological roles of oviduct-specific, oestrogen-dependent
28 glycoprotein. *Reproduction* 123:355–362.
- 29 Bussink AP, Speijer D, Aerts JMFG, Boot RG. 2007. Evolution of mammalian Chitinase(-like)
30 members of family 18 Glycosyl Hydrolases. *Genetics* 177:959–970.
- 31 Chen A-S, Taguchi T, Sakai K, Kikuchi K, Wang M-W, Miwa I. 2003. Antioxidant activities of

1 Chitobiose and Chitotriose. *Biol. Pharm. Bull.* 26:1326–1330.

2 Chen S, Zhou Y, Chen Y, Gu J. 2018. fastp: an ultra-fast all-in-one FASTQ preprocessor.

3 *Bioinformatics* 34:i884–i890.

4 Chen Y-H, Zhao H. 2019. Evolution of digestive enzymes and dietary diversification in birds.

5 *PeerJ* 7:e6840.

6 Cheng S-C, Liu C-B, Yao X-Q, Hu J-Y, Yin T-T, Lim BK, Chen W, Wang G-D, Zhang C-L,

7 Irwin DM, et al. 2023. Hologenomic insights into mammalian adaptations to myrmecophagy.

8 *Natl. Sci. Rev.* 10:nwac174.

9 Choo SW, Rayko M, Tan TK, Hari R, Komissarov A, Wee WY, Yurchenko AA, Kliver S,

10 Tamazian G, Antunes A. 2016. Pangolin genomes and the evolution of mammalian scales and

11 immunity. *Genome Res.* 26:1312–1322.

12 Christin P-A, Weinreich DM, Besnard G. 2010. Causes and evolutionary significance of genetic

13 convergence. *Trends Genet.* 26:400–405.

14 Cole TL, Zhou C, Fang M, Pan H, Ksepka DT, Fiddaman SR, Emerling CA, Thomas DB, Bi X,

15 Fang Q. 2022. Genomic insights into the secondary aquatic transition of penguins. *Nat. Commun.*

16 13:1–13.

17 Comte N, Morel B, Hasić D, Guéguen L, Boussau B, Daubin V, Penel S, Scornavacca C, Gouy

18 M, Stamatakis A, et al. 2020. Treerecs: an integrated phylogenetic tool, from sequences to

19 reconciliations. *Bioinformatics* 36:4822–4824.

20 Conway Morris S. 1999. The crucible of creation: The Burgess Shale and the rise of animals.

21 Oxford, New York: Oxford University Press

22 Delsuc F, Metcalf JL, Wegener Parfrey L, Song SJ, González A, Knight R. 2014. Convergence

23 of gut microbiomes in myrmecophagous mammals. *Mol. Ecol.* 23:1301–1317.

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

25 Douzery EJP. 2002. Molecular phylogeny of living xenarthrans and the impact of character and

26 taxon sampling on the placental tree rooting. *Mol. Biol. Evol.* 19:1656–1671.

27 Dudchenko O, Batra SS, Omer AD, Nyquist SK, Hoeger M, Durand NC, Shamim MS, Machol I,

28 Lander ES, Aiden AP. 2017. De novo assembly of the *Aedes aegypti* genome using Hi-C yields

29 chromosome-length scaffolds. *Science* 356:92–95.

30 Dunn CW, Luo X, Wu Z. 2013. Phylogenetic analysis of gene expression. *Integr. Comp. Biol.*

31 53:847–856.

1 Eizirik E, Murphy WJ, Koepfli K-P, Johnson WE, Dragoo JW, Wayne RK, O'Brien SJ. 2010.
2 Pattern and timing of diversification of the mammalian order Carnivora inferred from multiple
3 nuclear gene sequences. *Mol. Phylogenet. Evol.* 56:49–63.

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

6 Emms DM, Kelly S. 2019. OrthoFinder: Phylogenetic orthology inference for comparative
7 genomics. *Genome Biol.* 20:238.

8 Ferreira-Cardoso S, Delsuc F, Hautier L. 2019. Evolutionary tinkering of the mandibular canal
9 linked to convergent regression of teeth in placental mammals. *Curr. Biol.* 29:468–475.

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

13 Francischetti IMB, Assumpção TCF, Ma D, Li Y, Vicente EC, Uieda W, Ribeiro JMC. 2013.
14 The “Vampirome”: Transcriptome and proteome analysis of the principal and accessory
15 submaxillary glands of the vampire bat *Desmodus rotundus*, a vector of human rabies. *J.*
16 *Proteomics* 82:288–319.

17 Funkhouser JD, Aronson NN. 2007. Chitinase family GH18: Evolutionary insights from the
18 genomic history of a diverse protein family. *BMC Evol. Biol.* 7:96.

19 Galiano H, Tseng ZJ, Solounias N, Wang X, Zhan-Xiang Q, White S. 2022. A new aardwolf-line
20 fossil hyena from Middle and Late Miocene deposits of Linxia Basin, Gansu, China. *Vertebr.*
21 *Palasiat.* 60:81–116.

22 Gil F, Arencibia A, García V, Ramírez G, Vázquez JM. 2018. Anatomic and magnetic resonance
23 imaging features of the salivary glands in the dog. *Anat. Histol. Embryol.* 47:551–559.

24 Gordon-Thomson C, Kumari A, Tomkins L, Holford P, Djordjevic JT, Wright LC, Sorrell TC,
25 Moore GPM. 2009. Chitotriosidase and gene therapy for fungal infections. *Cell. Mol. Life Sci.*
26 66:1116–1125.

27 Gould SJ. 1990. Wonderful life: The Burgess Shale and the nature of history. WW Norton &
28 Company

29 Gould SJ. 2002. The Structure of Evolutionary Theory. Harvard University Press

30 Gouy M, Guindon S, Gascuel O. 2010. SeaView version 4: A multiplatform graphical user
31 interface for sequence alignment and phylogenetic tree building. *Mol. Biol. Evol.* 27:221–224.

1 Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I, Adiconis X, Fan L,
2 Raychowdhury R, Zeng Q, et al. 2011. Trinity: reconstructing a full-length transcriptome without
3 a genome from RNA-Seq data. *Nat. Biotechnol.* 29:644–652.

4 Hamid R, Khan MA, Ahmad M, Ahmad MM, Abdin MZ, Musarrat J, Javed S. 2013. Chitinases:
5 An update. *J. Pharm. Bioallied Sci.* 5:21–29.

6 Hussain M, Wilson JB. 2013. New paralogues and revised time line in the expansion of the
7 vertebrate GH18 family. *J. Mol. Evol.* 76:240–260.

8 Jacob F. 1977. Evolution and tinkering. *Science* 196:1161–1166.

9 Janiak MC, Chaney ME, Tosi AJ. 2018. Evolution of acidic mammalian chitinase genes (CHIA)
10 is related to body mass and insectivory in Primates. *Mol. Biol. Evol.* 35:607–622.

11 Jeuniaux C. 1961. Chitinase: An addition to the list of hydrolases in the digestive tract of
12 vertebrates. *Nature* 192:135–136.

13 Jeuniaux C. 1966. [111] Chitinases. In: *Methods in enzymology*. Vol. 8. Elsevier. p. 644–650.

14 Jeuniaux C. 1971. On some biochemical aspects of regressive evolution in animals. In:
15 *Biochemical evolution and the origin of life*. E. Schoffeniels. p. 304–313.

16 Jeuniaux C, Cornelius C. 1997. Distribution and activity of chitinolytic enzymes in the digestive
17 tract of birds and mammals. In: *First international conference on Chitin/Chitosan*.

18 Katoh K, Standley DM. 2013. MAFFT multiple sequence alignment software Version 7:
19 Improvements in performance and usability. *Mol. Biol. Evol.* 30:772–780.

20 Koepfli K-P, Jenks SM, Eizirik E, Zahirpour T, Van Valkenburgh B, Wayne RK. 2006.
21 Molecular systematics of the Hyaenidae: relationships of a relictual lineage resolved by a
22 molecular supermatrix. *Mol. Phylogenet. Evol.* 38:603–620.

23 Kozlov AM, Darriba D, Flouri T, Morel B, Stamatakis A. 2019. RAxML-NG: a fast, scalable
24 and user-friendly tool for maximum likelihood phylogenetic inference. *Bioinformatics* 35:4453–
25 4455.

26 Kumar S, Suleski M, Craig JM, Kasprowitz AE, Sanderford M, Li M, Stecher G, Hedges SB.
27 2022. TimeTree 5: An Expanded Resource for Species Divergence Times. *Mol. Biol. Evol.*
28 39:msac174.

29 Laheri S, Ashary N, Bhatt P, Modi D. 2018. Oviductal glycoprotein 1 (OVGP1) is expressed by
30 endometrial epithelium that regulates receptivity and trophoblast adhesion. *J. Assist. Reprod.*
31 *Genet.* 35:1419–1429.

- 1 Le SQ, Gascuel O. 2008. An improved general amino acid replacement matrix. *Mol. Biol. Evol.*
2 25:1307–1320.
- 3 Lee CG, Da Silva CA, Dela Cruz CS, Ahangari F, Ma B, Kang M-J, He C-H, Takyar S, Elias JA.
4 2011. Role of chitin and Chitinase/Chitinase-like proteins in inflammation, tissue remodeling,
5 and injury. *Annu. Rev. Physiol.* 73:479–501.
- 6 Li S, Nakayama H, Sinha NR. 2023. How to utilize comparative transcriptomics to dissect
7 morphological diversity in plants. *Curr. Opin. Plant Biol.* 76:102474.
- 8 Losos JB. 2011. Convergence, adaptation, and constraint. *Evol. Int. J. Org. Evol.* 65:1827–1840.
- 9 Losos JB. 2018. Improbable destinies: Fate, chance, and the future of evolution. Riverhead
10 Books, New York.
- 11 Love MI, Huber W, Anders S. 2014. Moderated estimation of fold change and dispersion for
12 RNA-seq data with DESeq2. *Genome Biol.* 15:550.
- 13 Ma J-E, Jiang H-Y, Li L-M, Zhang X-J, Li H-M, Li G-Y, Mo D-Y, Chen J-P. 2019. SMRT
14 sequencing of the full-length transcriptome of the Sunda pangolin (*Manis javanica*). *Gene*
15 692:208–216.
- 16 Ma J-E, Li L-M, Jiang H-Y, Zhang X-J, Li J, Li G-Y, Chen J-P. 2018. Acidic mammalian
17 chitinase gene is highly expressed in the special oxyntic glands of *Manis javanica*. *FEBS Open*
18 *Bio* 8:1247–1255.
- 19 Ma J-E, Li L-M, Jiang H-Y, Zhang X-J, Li J, Li G-Y, Yuan L-H, Wu J, Chen J-P. 2017.
20 Transcriptomic analysis identifies genes and pathways related to myrmecophagy in the Malayan
21 pangolin (*Manis javanica*). *PeerJ* 5:e4140.
- 22 MacManes M. 2014. On the optimal trimming of high-throughput mRNA sequence data. *Front.*
23 *Genet.* 5:13.
- 24 Manni M, Berkeley MR, Sepey M, Simão FA, Zdobnov EM. 2021. BUSCO update: Novel and
25 streamlined workflows along with broader and deeper phylogenetic coverage for scoring of
26 eukaryotic, prokaryotic, and viral genomes. *Mol. Biol. Evol.* 38:4647–4654.
- 27 McGhee GR. 2011. Convergent evolution: Limited forms most beautiful. MIT Press, Boston.
- 28 McGlothlin JW, Kobiela ME, Feldman CR, Castoe TA, Geffeney SL, Hanifin CT, Toledo G,
29 Vonk FJ, Richardson MK, Brodie ED, et al. 2016. Historical contingency in a multigene family
30 facilitates adaptive evolution of toxin resistance. *Curr. Biol.* 26:1616–1621.
- 31 McNab BK. 1984. Physiological convergence amongst ant-eating and termite-eating mammals.

1 *J. Zool.* 203:485–510.

2 Meredith RW, Janečka JE, Gatesy J, Ryder OA, Fisher CA, Teeling EC, Goodbla A, Eizirik E,
3 Simão TLL, Stadler T, et al. 2011. Impacts of the Cretaceous terrestrial revolution and KPg
4 extinction on mammal diversification. *Science* 334:521–524.

5 Morel B, Kozlov AM, Stamatakis A, Szöllősi GJ. 2020. GeneRax: A tool for species-tree-aware
6 maximum likelihood-based gene family tree inference under gene duplication, transfer, and loss.
7 *Mol. Biol. Evol.* 37:2763–2774.

8 Nguyen NTT, Vincens P, Dufayard JF, Roest Crollius H, Louis A. 2022. Genomicus in 2022:
9 Comparative tools for thousands of genomes and reconstructed ancestors. *Nucleic Acids Res.*
10 50:D1025–D1031.

11 Nishimura O, Hara Y, Kuraku S. 2017. gVolante for standardizing completeness assessment of
12 genome and transcriptome assemblies. *Bioinformatics* 33:3635–3637.

13 Novacek MJ. 1992. Mammalian phylogeny: Shaking the tree. *Nature* 356:121–125.

14 O’Leary MA, Bloch JI, Flynn JJ, Gaudin TJ, Giallombardo A, Giannini NP, Goldberg SL,
15 Kraatz BP, Luo Z-X, Meng J, et al. 2013. The placental mammal ancestor and the post-K-Pg
16 radiation of placentals. *Science* 339:662–667.

17 Olland AM, Strand J, Presman E, Czerwinski R, Joseph-McCarthy D, Krykbaev R, Schlingmann
18 G, Chopra R, Lin L, Fleming M, et al. 2009. Triad of polar residues implicated in pH specificity
19 of acidic mammalian chitinase. *Protein Sci.* 18:569–578.

20 Phillips CJ, Phillips CD, Goecks J, Lessa EP, Sotero-Caio CG, Tandler B, Gannon MR, Baker
21 RJ. 2014. Dietary and flight energetic adaptations in a salivary gland transcriptome of an
22 insectivorous bat. *PLoS One* 9:e83512.

23 Pillai AS, Chandler SA, Liu Y, Signore AV, Cortez-Romero CR, Benesch JLP, Laganowsky A,
24 Storz JF, Hochberg GKA, Thornton JW. 2020. Origin of complexity in haemoglobin evolution.
25 *Nature* 581:480–485.

26 Price MN, Dehal PS, Arkin AP. 2010. FastTree 2--approximately maximum-likelihood trees for
27 large alignments. *PloS One* 5:e9490.

28 Ratnasingham S, Hebert PDN. 2007. bold: The barcode of life data system
29 (<http://www.barcodinglife.org>). *Mol. Ecol. Notes* 7:355–364.

30 Recklies AD, White C, Ling H. 2002. The chitinase 3-like protein human cartilage glycoprotein
31 39 (HC-gp39) stimulates proliferation of human connective-tissue cells and activates both

1 extracellular signal-regulated kinase- and protein kinase B-mediated signalling pathways.
2 *Biochem. J.* 365:119–126.

3 Redford KH. 1987. Ants and termites as food. In: Genoways HH, editor. *Current Mammalogy*.
4 Boston, MA: Springer US. p. 349–399.

5 Reiss KZ. 2001. Using phylogenies to study convergence: the case of the ant-eating mammals.
6 *Am. Zool.* 41:507–525.

7 Roux J, Rosikiewicz M, Robinson-Rechavi M. 2015. What to compare and how: comparative
8 transcriptomics for evo-devo. *J. Exp. Zool. Part B: Mol. Dev. Evol.* 324:372–382.

9 Saint-Dizier M, Marnier C, Tahir MZ, Grimard B, Thoumire S, Chastant-Maillard S, Reynaud K.
10 2014. OVGP1 is expressed in the canine oviduct at the time and place of oocyte maturation and
11 fertilization. *Mol. Reprod. Dev.* 81:972–982.

12 Salgaonkar N, Prakash D, Nawani NN, Kapadnis BP. 2015. Comparative studies on ability of N-
13 acetylated chitooligosaccharides to scavenge reactive oxygen species and protect DNA from
14 oxidative damage. *Indian J. Biotechnol.* 14:186–192.

15 Sanders JG, Beichman AC, Roman J, Scott JJ, Emerson D, McCarthy JJ, Girguis PR. 2015.
16 Baleen whales host a unique gut microbiome with similarities to both carnivores and herbivores.
17 *Nat. Commun.* 6:8285.

18 Slater GSC, Birney E. 2005. Automated generation of heuristics for biological sequence
19 comparison. *BMC Bioinformatics* 6:31.

20 Smith SA, Robbins LW, Steiert JG. 1998. Isolation and characterization of a chitinase from the
21 nine-banded armadillo, *Dasypus novemcinctus*. *J. Mammal.* 79:486–491.

22 Sonesson C, Love MI, Robinson MD. 2016. Differential analyses for RNA-seq: Transcript-level
23 estimates improve gene-level inferences. *F1000 Res.* 4:1521.

24 Springer MS, Meredith RW, Teeling EC, Murphy WJ. 2013. Technical comment on “The
25 placental mammal ancestor and the post-K-Pg radiation of placentals.” *Science* 341:613–613.

26 Stamatakis A. 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of
27 large phylogenies. *Bioinformatics* 30:1312–1313.

28 Strobel S, Roswag A, Becker NI, Trenczek TE, Encarnação JA. 2013. Insectivorous bats digest
29 chitin in the stomach using acidic mammalian chitinase. *PloS One* 8:e72770.

30 Tabata E, Itoigawa A, Koinuma T, Tayama H, Kashimura A, Sakaguchi M, Matoska V, Bauer
31 PO, Oyama F. 2022. Noninsect-based diet leads to structural and functional changes of Acidic

1 Chitinase in Carnivora. *Mol. Biol. Evol.* 39:msab331.

2 Tjoelker LW, Gosting L, Frey S, Hunter CL, Le Trong H, Steiner B, Brammer H, Gray PW.

3 2000. Structural and functional definition of the human chitinase chitin-binding domain. *J. Biol.*

4 *Chem.* 275:514–520.

5 Tucker R. 1958. Taxonomy of the salivary glands of vertebrates. *Syst. Biol.* 7:74–83.

6 Vandewege MW, Sotero-Caio CG, Phillips CD. 2020. Positive selection and gene expression

7 analyses from salivary glands reveal discrete adaptations within the ecologically diverse bat

8 family Phyllostomidae. *Genome Biol. Evol.* 12:1419–1428.

9 Wang K, Tian S, Galindo-González J, Dávalos LM, Zhang Y, Zhao H. 2020. Molecular

10 adaptation and convergent evolution of frugivory in Old World and neotropical fruit bats. *Mol.*

11 *Ecol.* 29:4366–4381.

12 Westbury MV, Le Duc D, Duchêne DA, Krishnan A, Prost S, Rutschmann S, Grau JH, Dalén L,

13 Weyrich A, Norén K, et al. 2021. Ecological specialization and evolutionary reticulation in

14 extant Hyainidae. *Mol. Biol. Evol.* 38:3884–3897.

15 Wolf JBW. 2013. Principles of transcriptome analysis and gene expression quantification: an

16 RNA-seq tutorial. *Mol. Ecol. Resour.* 13:559–572.

17 Xie VC, Pu J, Metzger BP, Thornton JW, Dickinson BC. 2021. Contingency and chance erase

18 necessity in the experimental evolution of ancestral proteins. *eLife* 10:e67336.

19 Yates AD, Achuthan P, Akanni W, Allen James, Allen Jamie, Alvarez-Jarreta J, Amode MR,

20 Armean IM, Azov AG, Bennett R, et al. 2020. Ensembl 2020. *Nucleic Acids Res.* 48:D682–

21 D688.

22 Yusoff AM, Tan TK, Hari R, Koepfli K-P, Wee WY, Antunes A, Sitam FT, Rovie-Ryan JJ,

23 Karuppannan KV, Wong GJ. 2016. De novo sequencing, assembly and analysis of eight different

24 transcriptomes from the Malayan pangolin. *Sci. Rep.* 6:1–11.

25 Zakon HH, Lu Y, Zwickl DJ, Hillis DM. 2006. Sodium channel genes and the evolution of

26 diversity in communication signals of electric fishes: convergent molecular evolution. *Proc.*

27 *Natl. Acad. Sci. USA* 103:3675–3680.

28 Zhang F, Xu N, Yu Y, Wu S, Li S, Wang W. 2019. Expression profile of the digestive enzymes

29 of *Manis javanica* reveals its adaptation to diet specialization. *ACS Omega* 4:19925–19933.

30