Comparative transcriptomics reveals divergent paths of chitinase evolution underlying dietary convergence in ant-eating mammals

Rémi Allio[1,2,§,*], Sophie Teullet[1,§], Dave Lutgen[1,3,4,§], Amandine Magdeleine[1], Rachid Koual[1], Marie-Ka Tilak[1], Benoit de Thoisy[5,6], Christopher A. Emerling[1,7], Tristan Lefébure[8], and Frédéric Delsuc[1,*]

[1]*ISEM, Univ. Montpellier, CNRS, IRD, Montpellier, France*

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

[3]*Institute of Ecology and Evolution, University of Bern, Bern, Switzerland*

[4]*Swiss ornithological Institute, Sempach, Switzerland*

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

[6]*Kwata NGO, Cayenne, French Guiana, France*

[7]*Biology Department, Reedley College, Reedley, CA, USA*

[8]*Univ. Lyon, Université Claude Bernard Lyon 1, CNRS, ENTPE, UMR 5023 LEHNA, F-69622, Villeurbanne, France*

[§]*Equal contribution*

***Correspondence**

Rémi Allio: remi.allio@inrae.fr

Frédéric Delsuc: frederic.delsuc@umontpellier.fr

Key words

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

ORCID

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

Abstract

Ant-eating mammals represent a textbook example of convergent morphological evolution. Among them, anteaters and pangolins exhibit the most extreme convergent phenotypes with complete tooth loss, elongated skulls, protrusive tongues, and powerful claws to rip open ant and termite nests. Despite this remarkable convergence, comparative genomic analyses have shown that anteaters and pangolins differ in their chitinase gene (*CHIA*) repertoires, which potentially degrade the chitinous exoskeletons of ingested ants and termites. While the southern tamandua (*Tamandua tetradactyla*) harbors four functional CHIA paralogs (*CHIA1*, *CHIA2*, *CHIA3*, and *CHIA4*), Asian pangolins (*Manis* spp.) have only one functional paralog (*CHIA5*). These two placental mammal lineages also possess hypertrophied salivary glands producing large quantities of saliva to capture and potentially digest their social insect prey. We performed a comparative transcriptomic analysis of salivary glands in 23 representative species of placental mammals, including new ant-eating species and close relatives. Our results on chitinase gene expression suggest that salivary glands play a major role in adapting to an insect-based diet with myrmecophagous and insectivorous species highly expressing CHIA paralogs. Moreover, convergently-evolved pangolins and anteaters express different chitinases in their hypertrophied salivary glands and other additional digestive organs. *CHIA5* is overexpressed in Malayan pangolin, whereas the southern tamandua exhibits high levels of *CHIA3* and *CHIA4* expression. Overall, our results demonstrate that divergent molecular mechanisms underlie convergent adaptation to the ant-eating diet in pangolins and anteaters. This work highlights the role of historical contingency and molecular tinkering of the chitin-digestive enzyme toolkit in this classical example of convergent evolution.

Introduction

The phenomenon of evolutionary convergence is a fascinating process in which distantly related species independently acquire similar characteristics in response to the same selection pressures. A fundamental question famously illustrated by the debate between Stephen Jay Gould (Gould 2002) and Simon Conway Morris (Conway Morris 1999) resides in the relative contribution of historical contingency and evolutionary convergence in the evolution of biodiversity. While Gould (Gould 1990; 2002) argued that the evolution of species strongly depends on the characteristics inherited from their ancestors (historical contingency), Conway Morris (Conway Morris 1999) retorted that convergent evolution is one of the dominant processes leading to biodiversity evolution. Despite the huge diversity of organisms found on Earth and the numerous potential possibilities to adapt to similar conditions, the strong deterministic force of natural selection led to numerous cases of recurrent phenotypic adaptations (Losos 2011; McGhee 2011; Losos 2018). However, the role of historical contingency and evolutionary tinkering in convergent evolution has long been recognized, with evolution proceeding from available material through natural selection often leading to structural and functional imperfections (Jacob 1977). As first pointed out by François Jacob (Jacob 1977), molecular tinkering seems to be particularly frequent and has shaped the evolutionary history of a number of protein families (Pillai et al. 2020; Xie et al. 2021). Indeed, if in some cases, convergent phenotypes can be associated with similar or identical mutations in the same genes occurring in independent lineages (Arendt and Reznick 2008), in other cases, they appear to arise by diverse molecular paths (*e.g.* Christin et al. 2010). Hence, both historical contingency and evolutionary convergence seems to have impacted the evolution of the current biodiversity and the major question relies on evaluating the relative impact of these two evolutionary processes (Blount et al. 2018).

A notable example of convergent evolution is the adaptation to the specialized ant- and/or termite-eating diet (*i.e.* myrmecophagy) in placental mammals (Reiss 2001). Within placental mammals, over 200 species include ants and termites in their regime, but only 22 of them can be considered as specialized myrmecophagous mammals, eating more than 90% of social insects (Redford 1987). Historically, based on shared morphological characteristics, ant-eating mammals were considered monophyletic (*i.e.* Edentata; Novacek 1992; O'Leary et al. 2013), but molecular phylogenetic evidence now strongly supports their polyphyly (*e.g.* Delsuc et al. 2002; Meredith et al. 2011; Springer et al. 2013). This highly-specialized diet has indeed independently evolved in five placental orders: armadillos (Cingulata), anteaters

(Pilosa), aardvarks (Tubulidentata), pangolins (Pholidota), and aardwolves (Carnivora). As a consequence of foraging for small-sized prey (Redford 1987), similar morphological adaptations have evolved in these mammalian species such as powerful claws used to dig into ant and termite nests, tooth reduction culminating in complete tooth loss in anteaters and pangolins (Ferreira-Cardoso et al. 2019), an elongated muzzle with an extensible tongue (Ferreira-Cardoso et al. 2020), and viscous saliva produced by hypertrophied salivary glands (Reiss 2001). Due to strong energetic constraints imposed by a nutritionally poor diet, myrmecophagous mammals also share relatively low metabolic rates and might thus require specific adaptations to extract nutrients from the chitinous exoskeletons of their prey (McNab 1984). It has long been shown that chitinase enzymes are present in the digestive tract of mammals and vertebrates more broadly (Jeuniaux 1961; Jeuniaux 1966; Jeuniaux 1971; Jeuniaux and Cornelius 1997). More recent studies have indeed shown that chitinase genes are present in the mammalian genome and may play an important digestive function in insectivorous species (Bussink et al. 2007; Emerling et al. 2018; Janiak et al. 2018; Wang et al. 2020; Cheng et al. 2022). Elevated levels of digestive enzyme gene expression have notably been observed in placental mammal salivary glands. For instance in bat salivary glands, studies have shown that dietary adaptations can be associated with elevated expression levels in carbohydrase, lipase, and protease genes (Francischetti et al. 2013; Phillips et al. 2014; Vandewege et al. 2020).

In placental mammals, the salivary glands are composed of three major gland pairs (parotid, sublingual, and submandibular) and hundreds of minor salivary glands (Tucker 1958). In most myrmecophagous placental lineages, it has been shown that hypertrophied submandibular salivary glands are the primary source of salivary production. These enlarged horseshoe-shaped glands extend posteriorly along the side of the neck and ventrally over the chest. In the Malayan pangolin (*Manis javanica*), recent transcriptomic (Ma et al. 2017; Ma et al. 2019) and proteomic (Zhang et al. 2019) studies have shown that genes associated with digestive enzymes are highly expressed in salivary glands, which supports the hypothesis that the enlarged submandibular glands play an important functional role in social insect digestion. This result also found support in a study on the molecular evolution of the chitinase genes across 107 placental mammals that revealed the likely existence of a repertoire of five functional paralogous chitinase (CHIA, acidic mammalian chitinase) genes in the placental ancestor, which was subsequently shaped through multiple pseudogenization events associated with dietary adaptation during the placental radiation (Emerling et al. 2018). The widespread gene loss observed in carnivorous and herbivorous lineages resulted in a general

positive correlation between the number of functional *CHIA* paralogs and the percentage of invertebrates in the diet across placentals (Emerling et al. 2018). Indeed, mammals with a low proportion of insects in their diet present none or a few functional *CHIA* paralogs and those with a high proportion of insects in their diet generally have retained four or five functional *CHIA* paralogs (Emerling et al. 2018; Janiak et al. 2018; Wang et al. 2020). Among mammals, pangolins appear as an exception as the two investigated species (*M. javanica* and *Manis pentadactyla*) possess only one functional *CHIA* paralog (*CHIA5*) whereas other myrmecophagous species such as the southern tamandua (*Tamandua tetradactyla*) and the aardvark (*Orycteropus afer*) possess respectively four (*CHIA1-4*) and five (*CHIA1-5*) functional paralogs (Emerling et al. 2018). The presence of the sole *CHIA5* in pangolins was interpreted as the consequence of historical contingency with the probable loss of *CHIA1-4* functionality in the last common ancestor of Pholidota and Carnivora (Emerling et al. 2018). In Carnivora, it has recently been confirmed that a non insect-based diet has caused structural and functional changes in the *CHIA* gene repertoire resulting in multiple losses of function with only few species including insects in their diet retaining a fully functional *CHIA5* gene (Tabata et al. 2022). The fact that *CHIA5* was found to be highly expressed in the main digestive organs of the Malayan pangolin (Ma et al. 2017; Ma et al. 2019; Cheng et al. 2022) suggests that pangolins might compensate for their reduced chitinase repertoire by an increased ubiquitous expression of their only remaining functional paralog in multiple organs.

To test this hypothesis, we first reconstructed the detailed evolutionary history of the chitinase gene family in mammals. Then, we conducted a comparative transcriptomic analysis of chitinase gene expression in salivary glands of 23 mammal species including 17 newly generated transcriptomes from myrmecophagous placentals and other mammalian species. Finally, we compared the expression of chitinase paralogs in different organs between the nine-banded armadillo (*Dasypus novemcinctus*), the Malayan pangolin (*M. javanica*), and the southern tamandua (*T. tetradactyla*) for which we produced 12 new transcriptomes from eight additional organs. Our results shed light on the molecular underpinnings of convergent evolution in ant-eating mammals by revealing that divergent paths of chitinase molecular evolution underlie dietary convergence between anteaters and pangolins.

Results

Mammalian chitinase gene family evolution

The reconciled maximum likelihood tree of mammalian chitinase genes is presented in Figure 1A. The evolution of this gene family constituted by nine paralogs is characterized by the presence of numerous inferred gene losses with 384 speciation events followed by gene loss and 48 gene duplications as estimated by the gene tree/species tree reconciliation algorithm of GeneRax. At the base of the reconciled gene tree, we found the clade *CHIA1-2/OVGP1* (optimal root inferred by the reconciliation performed with TreeRecs) followed by a duplication separating the *CHIT1/CHI3L1-2* and *CHIA3-5* groups of paralogs. Within the *CHIT1/CHI3L* clade, two consecutive duplications gave rise to *CHIT1*, then *CHI3L1* and *CHI3L2*. In the *CHIA3-5* clade, a first duplication separated *CHIA3* from *CHIA4* and *CHIA5*, which were duplicated subsequently. Marsupial *CHIA4* sequences were located at the base of the *CHIA4-5* clade suggesting that this duplication might be specific to placentals. The *CHIA5* sequences of chiropterans were found at the base of the *CHIA5* clade. The duplication that gave rise to the *CHIA4* and *CHIA5* genes appears recent and specific to eutherians (marsupials and placentals) since no other taxon was found within these clades. This scenario of chitinase gene evolution is consistent with synteny analysis showing physical proximity of *CHIA1-2* and *OVGP1* on one hand, and *CHIA3-5* on the other hand (Fig. 1B), which implies that chitinase genes evolved by successive tandem duplications. However, evidence of gene conversion between the two more recent duplicates (*CHIA4* and *CHIA5*) at least in some taxa suggests that further data are necessary to fully disentangle the origins of these two paralogs (Emerling et al. 2018). Within the *CHIA5* clade of Muroidea (Spalacidae, Cricetidae and Muridae), we found four subclades (named here *CHIA5a-d*) representing potential duplications specific to the muroid rodent species represented in our dataset. From the *CHIA5a* paralog, two consecutive duplications gave rise to the three *CHIA5b-d* paralogs represented by long branches, characterizing rapidly evolving sequences. The duplication giving rise to the *CHIA5c* and *CHIA5d* paralogs concerns only the Cricetidae and Muridae, *Nannospalax galili* (Spalacidae) being present only in the clade of the *CHIA5b* paralogous gene.

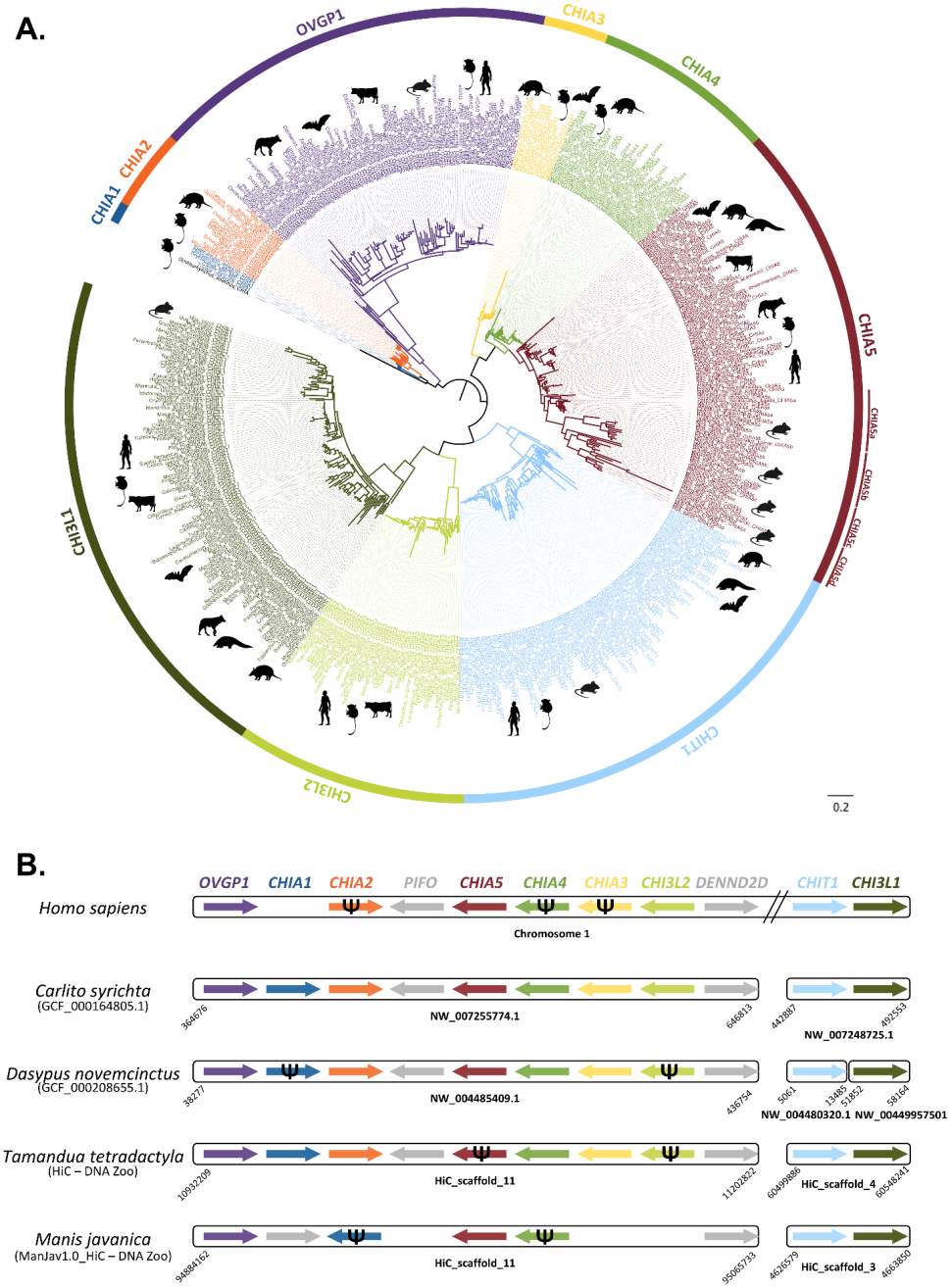


Figure 1: A. Mammalian chitinase gene family tree reconstructed using a maximum likelihood gene-tree/species-tree reconciliation approach on protein sequences. The nine chitinase paralogs are indicated on the outer circle. Scale bar represents the mean number of amino acid substitutions per site. B. Synteny of the nine chitinase paralogs in humans (*Homo sapiens*), tarsier (*Carlito syrichta*), nine-banded armadillo (*Dasypus novemcinctus*) and the two main focal convergent ant-eating species: the southern tamandua (*Tamandua tetradactyla*) and the Malayan pangolin (*Manis javanica*). Assembly names and accession numbers are indicated below species names. Arrows represent genes with scaffold/contig names and BLAST hit positions indicated below. Arrow direction indicates gene transcription direction as inferred in Genomicus v100.01 (Nguyen et al. 2022) for genes located on short contigs. Ψ symbols indicate pseudogenes as determined in Emerling et al. (2018). Genes with negative BLAST results were not represented and are probably not functional or absent.

202

203 **Ancestral sequences comparison**

204 The ancestral amino acid sequences of the nine chitinase paralogs have been reconstructed
205 from the reconciled mammalian gene tree and compared to gain further insight into the
206 potential function of the enzymes they encode (Fig. 2). The alignment of predicted amino
207 acid sequences locates the chitinolytic domain between positions 133 and 140 with the
208 preserved pattern DXXDXDXE. The ancestral sequences of CHI3L1 and CHI3L2, as all
209 contemporary protein sequences of these genes, have a mutated chitinolytic domain with
210 absence of a glutamic acid at position 140 (Fig. 2A), which is the active proton-donor site
211 necessary for chitin hydrolysis (Olland et al. 2009; Hamid et al. 2013). This indicates that the
212 ability to degrade chitin has likely been lost before the duplication leading to CHI3L1 and
213 CHI3L2 (Fig. 2B). It is also the case for the ancestral sequences of the muroid-specific
214 CHIA5b-d, which thus cannot degrade chitin (data not shown). The ancestral sequence of
215 OVGP1 also presents a mutated chitinolytic site although the glutamic acid in position 140 is
216 present (Fig. 2A). The evolution of the different chitinases therefore seems to be related to
217 changes in their active site. The six cysteine residues allowing the binding to chitin are found
218 at positions 371, 418, 445, 455, 457 and 458 (Fig. 2C). The absence of one of these cysteines
219 prevents binding to chitin (Tjoelker et al., 2000) as this is the case in the ancestral OVGP1
220 protein where the last four cysteine residues are changed (Fig. 2C). The other ancestral
221 sequences present the six conserved cysteine residues and thus can bind to chitin (Fig. 2C).

222

223

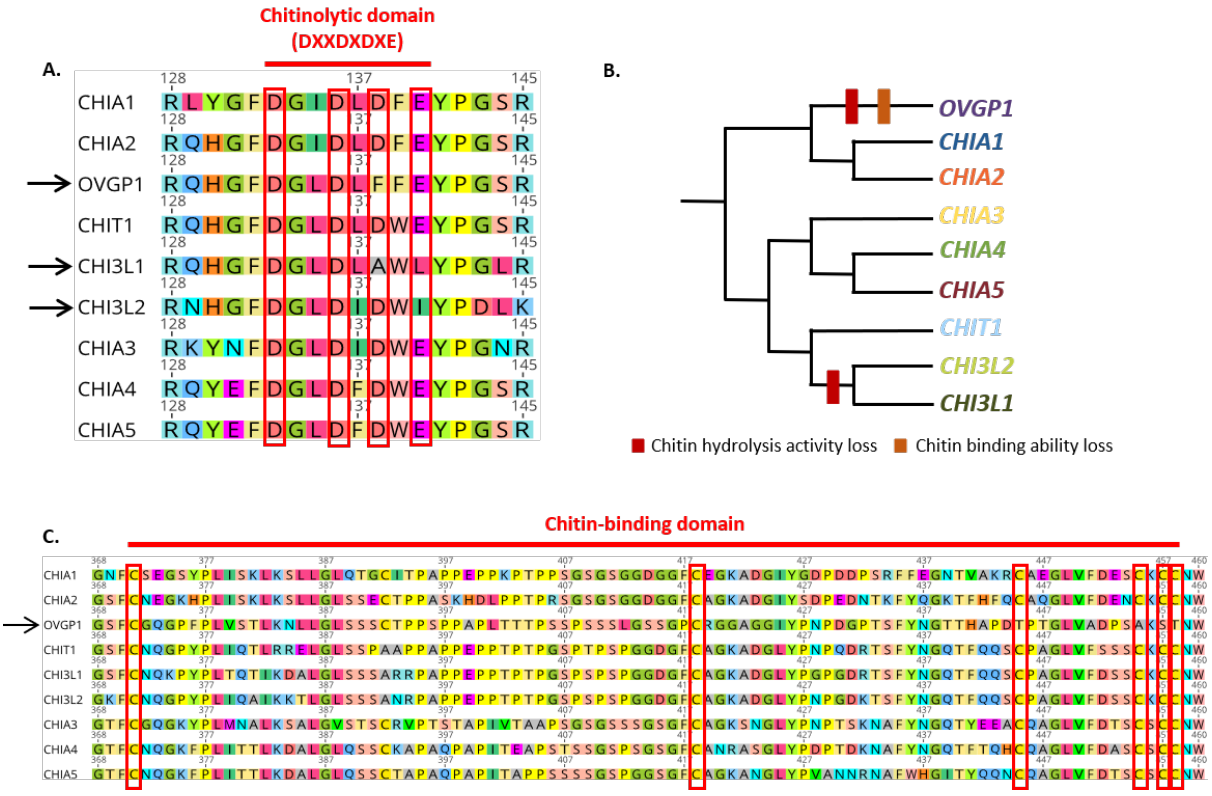


Figure 2: Comparison of predicted ancestral sequences of the nine mammalian chitinase paralogs. A. Conserved residues of the canonical chitinolytic domain active site (DXXDXDXE). Arrows indicate paralogs in which changes occurred in the active site. B. Summary of the evolution of chitinase paralogs functionality. 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 for the digestion of ants and termites in ant-eating mammals, we analyzed gene expression profiles of the nine chitinase paralogs revealed by the gene family tree reconstruction in 28 salivary gland transcriptomes (Fig. 3). *CHIA1* was only expressed in the elephant shrew (*Elephantulus myurus*; 21.71 normalized read counts [NC]). *CHIA2* was only expressed in the wild boar (*Sus scrofa*; 42.06 NC). *CHIA3* was expressed in the insectivorous California leaf-nosed bat (*Macrotus californicus*; 32.55 NC) and in all three southern tamandua individuals (*T. tetradactyla*; 45.24, 37.75, and 13.93 NC). *CHIA4* was also highly expressed in all three southern tamandua individuals (202.55, 521.48, and 168.21 NC), in the giant anteater (*M. tridactyla*; 49.85 NC), and in the California leaf-nosed bat (*M. californicus*; 16,232.31 NC). The expression of *CHIA5* was much higher in the two Malayan pangolin individuals (*Manis javanica*; 190,773.22 and 719.45 NC) than in the three other species in which we detected

246 expression of this gene: the domestic mouse (*Mus musculus*; 39.57 NC), the common genet
247 (*Genetta genetta*; 134.49 NC), and the wild boar (*Sus scrofa*; 152.78 NC). *CHIT1* was
248 expressed in many species (11 out of 28 samples) with NC values ranging from 42.08 NC in
249 a single southern tamandua (*T. tetradactyla*) individual to 105,830.61 NC in the short-tailed
250 shrew tenrec (*Microgale brevicaudata*). *CHI3L1* was expressed in most species (20 out of 28
251 samples) with NC values ranging from 67.25 for the giant anteater (*M. tridactyla*) to 1339.38
252 for a Malayan pangolin (*M. javanica*) individual. *CHI3L2* was expressed in human (*H.*
253 *sapiens*; 1357.58 NC), the wild boar (*S. scrofa*; 246.61 NC), the elephant shrew (*E. myurus*;
254 94.42 NC), and the common tenrec (*Tenrec ecaudatus*; 70.24 NC). *OVGP1* was only found
255 expressed at very low levels in the domestic dog (*Canis lupus familiaris*; 6.65 NC), human
256 (*H. sapiens*; 15.02 NC), and the wild boar (*S. scrofa*; 17.18 NC). Finally, the southern
257 aardwolf (*P. cristatus*), the Norway rat (*Rattus norvegicus*), and the tent-making bat
258 (*Uroderma bilobatum*) did not appear to express any of the nine chitinase gene paralogs in
259 any of our salivary gland samples.

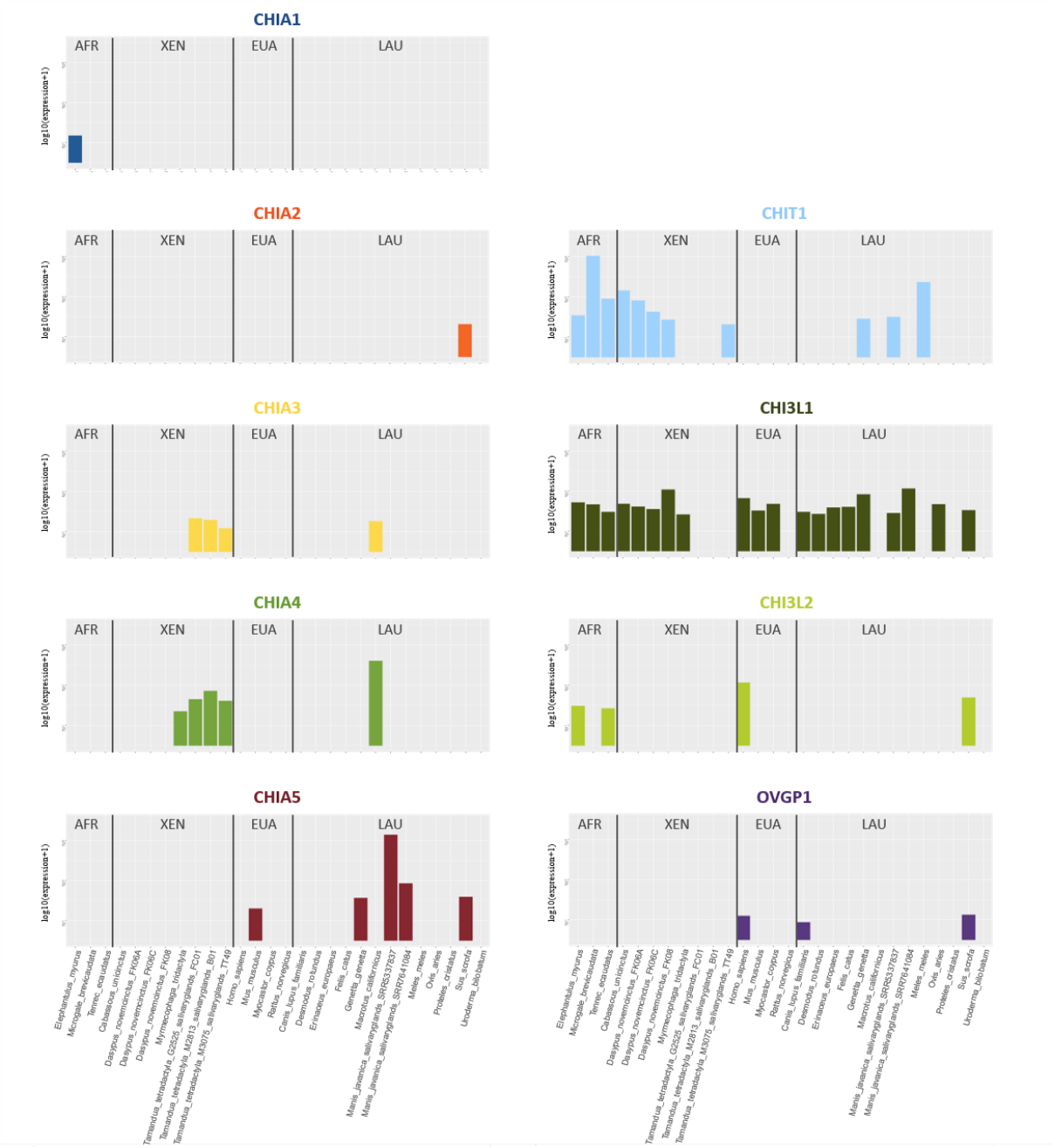


Figure 3: Comparative expression of the nine chitinase paralogs in 28 mammalian salivary gland transcriptomes. Species are ordered taxonomically into the four major placental clades: AFR: Afrotheria, XEN: Xenarthra, EUA: Euarchontoglires, and LAU: Laurasiatheria. Expression level is represented as log10 (Normalized Counts + 1).

Chitinase gene expression in other digestive and non-digestive organs

The expression level of the nine chitinase paralogs in several organs was compared among three species including an insectivorous xenarthran (the nine-banded armadillo; *D. novemcinctus*) and the two main convergent myrmecophagous species (the southern anteater; *T. tetradactyla*, and the Malayan pangolin; *M. javanica*) (Fig. 4). This analysis revealed

marked differences in expression level of these genes among the three species and among their digestive and non-digestive organs. *CHIT1* was expressed in all tissues in *M. javanica* and only in the spleen, testes, tongue and small intestine in *T. tetradactyla*, and in the cerebellum, liver, lungs and salivary glands in *D. novemcinctus*. *CHI3L1* was found to be expressed in the majority of digestive and non-digestive tissues in all three species. *CHI3L2* is non-functional or even absent in the genome of these three species and was therefore not expressed. *OVGP1* was only weakly expressed in the lungs of *M. javanica* (2.22 NC).

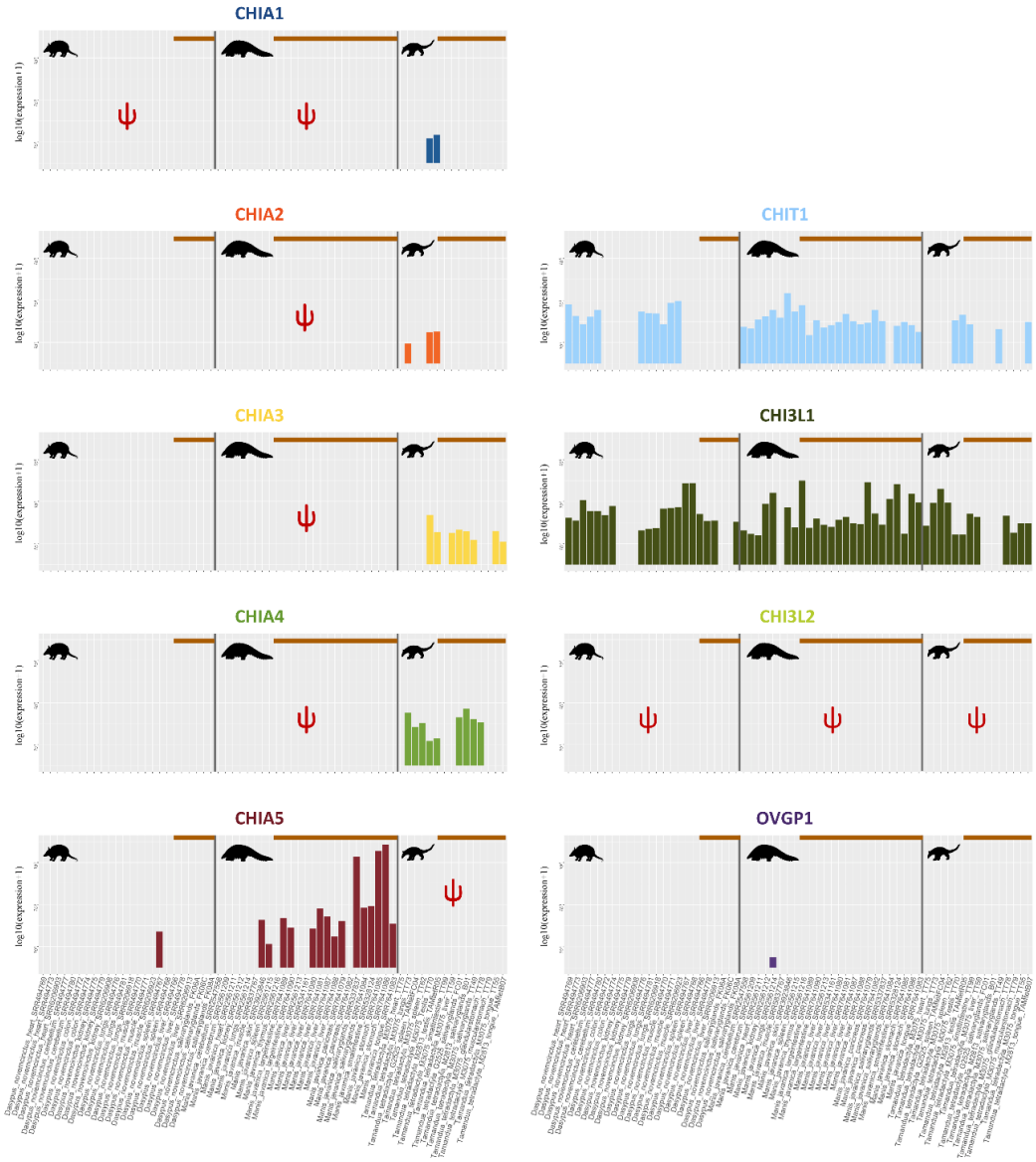


Figure 4: Comparative expression of *CHIA1-5*, *CHIT1*, *CHI3L1-2*, and *OVGP1* in 64 transcriptomes from different organs in three mammalian species: the nine-banded armadillo (*Dasypus novemcinctus*), the Malayan pangolin (*Manis javanica*), and the southern tamandua (*Tamandua tetradactyla*). Non-functional pseudogenes are represented by the Ψ symbol and horizontal bars indicate the digestive organs on the right side of the different graphs. Expression level is represented as log10 (Normalized Counts + 1).

In the nine-banded armadillo (*D. novemcinctus*), although only *CHIA1* is pseudogenized and therefore logically not expressed, we did not detect any expression of *CHIA2*, *CHIA3*, and *CHIA4* in the tissues studied here, and *CHIA5* was only weakly expressed in one spleen sample (52.40 NC) (Fig. 4). In the Malayan pangolin (*M. javanica*), whereas *CHIA1-4* are non-functional and consequently not expressed, *CHIA5* was found expressed in all digestive organs with particularly high levels in the stomach (366,485.51 and 739,987.73 NC) and salivary glands (190,773.22 and 719.45 NC), and at milder levels in the tongue (123.60 NC), liver (254.84 NC on average when expressed), pancreas (166.04 NC), large intestine (232.87 and 78.56 NC), and small intestine (837.44 NC), but also in skin (184.75 NC) and spleen (12.29 NC) samples. In the southern tamandua (*T. tetradactyla*), only *CHIA5* is pseudogenized (Fig. 4). *CHIA1* was only found weakly expressed in the testes (21.57 and 13.88 NC), and *CHIA2* also had low expression in the testes (31.78 and 29.85 NC) and lungs (7.58 NC). *CHIA3* was also expressed in testes (33.43 and 226.68 NC), tongue (11.17 and 37.81 NC), salivary glands (45.24, 37.75, and 13.93 NC), and liver (30.56 NC). Finally, *CHIA4* was expressed in the testes (18.65 and 13.88 NC), spleen (69.80 and 108.01 NC), lungs (334.64 NC), salivary glands (202.54, 521.48, and 168.21 NC), and glandular stomach (112.31 NC).

Discussion

Evolution of chitinase paralogs toward different functions

Chitinases have long been suggested to play an important role in insect digestion in mammals (Jeuniaux 1961; Jeuniaux 1966; Jeuniaux 1971; Jeuniaux and Cornelius 1997). Phylogenetic analyses of the Glycosyl Hydrolase gene family (GH18), which comprises genes encoding chitinase-like proteins, have revealed a dynamic evolutionary history despite a high degree of synteny among mammals (Bussink et al. 2007; Hussain and Wilson 2013). Our maximum likelihood phylogenetic analyses recovered nine functional paralogous chitinase gene sequences in mammalian genomes (Fig. 1A). In addition to the five *CHIA* paralogs previously characterized (Emerling et al. 2018; Janiak et al. 2018), we were able to identify an additional gene *OVGP1* that is most closely related to the previously characterized *CHIA1* and *CHIA2* genes. In mammals, OVGP1 has a role in fertilization and embryonic

development (Buhi 2002; Saint-Dizier et al. 2014; Algarra et al. 2016; Laheri et al. 2018). However, different name aliases for OVGP1 include Mucin 9 and CHIT5 (www.genecards.org), which suggests a possible digestive function. This result was further confirmed by synteny analyses suggesting a common origin by tandem duplication for *CHIA1-2* and *OVGP1* within the conserved chromosomal cluster also including *CHIA3-5* and *CHI3L2* (Fig. 1B). The comparison of ancestral amino acid sequences of the nine chitinase paralogs revealed differences in their ability to bind and degrade chitin (Fig. 2), suggesting that these paralogs have evolved towards different functional specializations. The evolution of chitinase-like proteins was accompanied by a loss of chitin hydrolysis enzymatic activity that occurred several times independently (Bussink et al. 2007; Funkhouser and Aronson 2007; Hussain and Wilson 2013; Fig. 2B). CHI3L1 and CHI3L2, which are expressed in various cell types including macrophages and synovial cells, play roles in cell proliferation and immune response (Recklies et al. 2002; Areshkov et al. 2011; Lee et al. 2011). In contrast to these chitinase-like proteins, CHIT1 and the five CHIAs are able to degrade chitin. In humans, *CHIT1* is expressed in macrophages and neutrophils and is suspected to be involved in the defense against chitin-containing pathogens such as fungi (Gordon-Thomson et al. 2009; Lee et al. 2011). In addition to their role in chitin digestion (Boot et al. 2001), CHIAs are also suspected to play a role in the inflammatory response (Lee et al. 2011) and are expressed in non-digestive tissues, in agreement with our comparative transcriptomic results. It has thus been proposed that the expansion of the chitinase gene family is linked to the emergence of the innate and adaptive immune systems in vertebrates (Funkhouser and Aronson 2007).

CHIA genes specific to muroid rodents and characterized by rapidly evolving sequences have also been described as chitinase-like rodents-specific (CHILrs) enzymes (Bussink et al. 2007; Hussain and Wilson 2013). These enzymes also appear to have evolved for functions in the immune response (Lee et al. 2011; Hussain and Wilson 2013). CHIA5b cannot bind to chitin unlike CHIA5c and CHIA5d, suggesting different roles for these three paralogous proteins. The evolution of the different *CHIA1-5* genes involved changes in their catalytic sites, which have consequences on the secondary structure of enzymes and potentially affect their optimal pH or function as it has recently been shown for *CHIA5* in Carnivora (Tabata et al. 2022). Experimentally testing the chitin degradation activity on different substrates and at different pH of enzymes produced from the ancestral sequences reconstructed for each of the five *CHIA* paralogs would allow a better understanding of their enzymatic activity. Studying the potential binding of these enzymes with other substrates

would shed more light on their functional roles. For instance, the change of a cysteine in the chitin binding domain prevents binding to this substrate but not to tri-N-acetyl-chitotriose (Tjoelker et al. 2000), a compound derived from chitin with antioxidant properties (Chen et al. 2003; Salgaonkar et al. 2015). Such functional assays, complemented by transcriptomic data to determine their expression profile across different tissues and organs (as previously done in the Malayan pangolin; Yusoff et al. 2016; Ma et al. 2017; Ma et al. 2019; Cheng et al. 2022), might help decipher their respective roles in mammalian digestion (see below).

Impact of historical contingency and molecular tinkering on chitinase gene evolution and expression

In the specific case of adaptation to myrmecophagy, comparative genomic and transcriptomic analyses of these chitinase genes, in particular chitin-degrading CHIAs, have led to a better understanding of how convergent adaptation to myrmecophagy in placentals occurs at the molecular level (Emerling et al. 2018; Cheng et al. 2022). On one hand, anteaters (Pilosa; Vermilingua) likely inherited five *CHIA* genes from an insectivorous ancestor (Emerling et al. 2018), but then the *CHIA5* gene was lost. In the southern tamandua (*T. tetradactyla*), the inactivating mutations of *CHIA5* have been identified and the estimated inactivation time of this gene was 6.8 Ma, subsequent to the origin of Vermilingua (34.2 Ma) and after the divergence with the giant anteater (*M. tridactyla*) at 11.3 Ma, suggesting a loss specific to lesser anteaters of the genus *Tamandua* (Emerling et al. 2018). Our study did not find this gene expressed in giant anteater salivary glands. On the other hand, in insectivorous carnivores (Carnivora) and pangolins (Pholidota), *CHIA5* is functional whereas *CHIA1-4* are pseudogenized (Emerling et al. 2018; Tabata et al. 2022). Similar inactivating mutations were observed in the *CHIA1* gene in carnivores and pangolins and were dated to at least 67 Mya, well before the origin of carnivores (46.2 Ma) and pangolins (26.5 Ma) (Emerling et al. 2018). Thus, despite relying on a fully myrmecophagous diet, pangolins have only one functional *CHIA* gene, probably due to historical contingency linked to their common inheritance with carnivores. These analyses have therefore highlighted opposing pseudogenization events between convergent myrmecophagous species where lesser anteaters (genus *Tamandua*) have retained four out of the five functional chitin-degrading *CHIA* genes (*CHIA1-4*), whereas the Malayan pangolin (*M. javanica*) only inherited the fifth one (*CHIA5*). This peculiar evolutionary history raised the question whether the Malayan pangolin potentially compensates for the paucity of its functional chitinase gene repertoire by overexpressing *CHIA5* in different digestive organs.

Since the presence of enlarged salivary glands is a hallmark of convergent ant-eating mammals ensuring massive production of saliva to help catch and potentially digest prey, we first investigated chitinase gene expression in mammalian salivary glands. Our comparative transcriptomic study encompassing a diversity of species with different diets revealed that in ant-eating mammals, the Malayan pangolin (*M. javanica*), the southern tamandua (*T. tetradactyla*), the giant anteater (*M. tridactyla*) and the southern naked-tailed armadillo (*C. unicinctus*) all express one or more chitinase genes in their salivary glands. More specifically, we found that *CHIA1* and *CHIA2* were almost never expressed in mammalian salivary glands. In contrast, *CHIA3* and *CHIA4* appeared to be mainly expressed in the insectivorous California leaf-nosed bat and anteaters. Our results therefore suggest that *CHIA3* and *CHIA4* are the main chitinase paralogs involved in chitin digestion in anteaters. Interestingly, *CHIA3* and *CHIA4* gene expressions were also particularly elevated in the California leaf-nosed bat (*M. californicus*), which is highly insectivorous. This result, together with previous observations made on myrmecophagous mammals, strongly supports the hypothesis that salivary glands play a preponderant adaptive role in placental mammal evolution towards insectivory. Indeed, in the sanguivorous common vampire bat (*D. rotundus*) and the frugivorous tent-making bat (*U. bilobatum*), none of the chitinase genes were expressed (Fig. 3). The most likely explanation is that these genes have been pseudogenized in both species, which would be concordant with the findings of comparative genomic studies reporting widespread pseudogenizations of *CHIA* paralogs across multiple non-insectivorous bat species (Emerling et al. 2018) with complete loss of *CHIA1-5* function in the vampire bat (Wang et al. 2020). Transcriptomic analyses of additional digestive tissues besides salivary glands in bats (Vandewege et al. 2020) may further clarify this pattern since chitinolytic activity has previously been reported in the stomachs of seven insectivorous bat species (Strobel et al. 2013). Finally, we were able to confirm the hypothesis implying an overexpression of the only chitinase gene possessed by the Malayan pangolin. Indeed, salivary gland expression profiles for the *CHIA5* gene in *M. javanica* were substantially higher than in the three other species (mouse, genet and wild boar) in which we detected expression of this gene but also substantially higher than the expression found for any other chitin-degrading *CHIA* in the mammalian species considered. Overall, our chitinase gene expression results support a primary role for salivary glands in insect-eating placental mammal prey digestion through the use of different *CHIA* paralogs in different species.

Our differential expression comparison of the distinct chitinase paralogs across different organs further highlighted the importance of *CHIA5* for Malayan pangolin digestion

physiology by confirming its ubiquitous expression in all major digestive tissue types (tongue, salivary glands, stomach, pancreas, liver, and large and small intestines) (Ma et al. 2017; Ma et al. 2019; Cheng et al. 2022; and Fig. 4). More specifically, *CHIA5* was found expressed at particularly high levels in the stomach and salivary glands. These results are in line with previous proteomic studies that have also identified *CHIA5* as a digestive enzyme (Zhang et al. 2019), which has been confirmed to be highly expressed by RT-qPCR in the specialized oxyntic glands stomach (Ma et al. 2018a; Cheng et al. 2022), reflecting a key adaptation of the Malayan pangolin to its strictly myrmecophagous diet. By contrast, in the southern tamandua (*T. tetradactyla*) only *CHIA5* is pseudogenized (Emerling et al. 2018; Cheng et al. 2022). Although *CHIA1* and *CHIA2* are functional, they did not appear to be expressed in the organs of the digestive tract of the southern tamandua individual sampled here (Fig. 4). These two chitinase genes were only found weakly expressed in the testes and lungs and may therefore not be involved in digestion in lesser anteaters. On the other hand, *CHIA3* and *CHIA4* were expressed across several digestive organs including tongue, salivary glands, stomach, and liver (Fig. 4). *CHIA3* was also expressed in testes and *CHIA4* in testes, spleen, and lungs but their co-expression in the salivary glands of the three distinct southern tamandua individuals sampled here (Figs. 3, 4) strongly suggests that they play a crucial role in chitin digestion in this myrmecophagous species. Conversely, in the insectivorous nine-banded armadillo (*D. novemcinctus*), although only *CHIA1* is pseudogenized (Emerling et al. 2018) and therefore not expressed, we did not detect any expression of *CHIA2*, *CHIA3*, and *CHIA4* in the tissues of the individuals studied here, including salivary glands (Figs. 3, 4), and *CHIA5* was only weakly expressed in one spleen sample (Fig. 4). Yet, chitinases could still participate in prey digestion in the nine-banded armadillo as they have been isolated from gastric tissues (Smith et al. 1998); results we could not confirm here, the liver being the only additional digestive organ besides salivary glands represented in our dataset for this species. However, the comparison with the two myrmecophagous species seems to fit well with its less specialized insectivorous diet and actually further underlines the contrasted specific use of distinct *CHIA* paralogs for chitin digestion in anteaters and pangolins.

Our results demonstrate that in the case of the southern tamandua (*T. tetradactyla*) and the Malayan pangolin (*M. javanica*), two myrmecophagous species that diverged about 100 Ma ago (Meredith et al. 2011), convergent adaptation to myrmecophagy has been achieved by using paralogs of different chitinase genes to digest chitin, probably due to phylogenetic constraints leading to the loss of *CHIA1*, *CHIA2*, *CHIA3*, and *CHIA4* in the ancestor of Ferae (Carnivora and Pholidota) as suggested by Emerling et al. (2018).

Pangolins and anteaters present extreme morphological adaptations including the complete loss of dentition but a detailed study of their feeding apparatus has shown that convergent tooth loss resulted in divergent structures in the internal morphology of their mandible (Ferreira-Cardoso et al. 2019). Our results combined to this observation clearly show that the evolution of convergent phenotypes in myrmecophagous mammals does not necessarily imply similar underlying mechanisms. Our study shows that historical contingency resulted in molecular tinkering (*sensu* Jacob 1977) of the chitinase gene family at both the genomic and transcriptomic levels. Working from different starting materials (*i.e.* different *CHIA* paralogs), natural selection led pangolins and anteaters to follow different paths in their adaptation to the myrmecophagous diet.

A potential complementary role of the gut microbiome?

Chitinase gene family evolution seems to have been strongly influenced by historical contingency events related to gene loss following adaptation to a specific diet (Emerling et al. 2018; Janiak et al. 2018; Chen and Zhao 2019; Tabata et al. 2022). For instance, fossil evidence showing that stem penguins primarily relied on large prey items like fish and squid has been invoked to explain the loss of all functional *CHIA* genes in all extant penguin species despite the recent specialization of some species towards a chitin-rich crustacean diet (Cole et al. 2022). One might therefore wonder why in highly specialized myrmecophagous groups which inherited a depauperate chitinase repertoire, such as pangolins and aardwolves, secondary chitinase duplications did not occur. As we demonstrated in the Malayan pangolin, one possible solution is to adjust the expression level of the remaining *CHIA5* paralog and expand its expression to multiple digestive organs. However, contrary to anteaters and pangolins, the southern aardwolf (*P. cristatus*) did not seem to express any chitinase gene in its salivary glands (Fig. 3). The presence of frameshift mutations and stop codons was inspected in all nine chitinase genes in the southern aardwolf genome (Allio et al. 2021). As expected, *CHIA1*, *CHIA2*, *CHIA3*, *CHIA4* were indeed found to be non functional, and *CHI3L2* seems to be absent from the genome of the southern aardwolf as in most members of Carnivora (Emerling et al. 2018; Tabata et al. 2022). While no inactivating mutations could be detected in the coding sequences of *CHIA5*, *CHI3L1*, *CHIT1* or *OVGP1*, we cannot rule out the possibility that some specific mutations in regulatory elements inactivating the expression of these genes could have appeared in *P. cristatus*. However, we verified that the southern aardwolf possesses the same amino acids at positions 214 and 216 of its *CHIA5*

exon 7, which control the chitinolytic activity of this chitinase, as its sister-species the striped hyena (*Hyaena hyaena*) and the other carnivore species including insects in their diet in which *CHIA5* is fully functional (Tabata et al. 2022). This intriguing result needs to be confirmed by studying the expression profiles of chitinase genes across digestive organs including the stomach in additional aardwolf specimens.

The aardwolf lineage represents the sister-group of all other hyenas (Koepfli et al. 2006; Westbury et al. 2021) from which it diverged < 10 Ma (Eizirik et al. 2010). The fossil record indicates that the adaptation of aardwolves to myrmecophagy is relatively recent (< 4 Ma; Galiano et al. 2022) and there are no clear signs of specific adaptation to an exclusive termite-based diet in the southern aardwolf genome (Westbury et al. 2021). This raises the possibility that the gut microbiome might play a key role for termite digestion in this species as suggested by results of 16S rRNA barcoding analyses of fecal samples (Delsuc et al. 2014). Aardwolves, and myrmecophagous mammals more broadly, therefore provide a model of choice for testing whether the loss of functional *CHIA* genes could be compensated by symbiotic bacteria from the gastrointestinal tract microbiota capable of degrading chitin, as previously shown in baleen whales eating krill (Sanders et al. 2015). A first metagenomic study of the fecal microbiome of the Malayan pangolin (*M. javanica*) previously identified a number of gut bacterial taxa containing chitinase genes capable of degrading chitin (Ma et al. 2018b). A more recent study has confirmed the chitin degradation potential of the Malayan pangolin gut microbiome and proposed that chitin is digested in this species by a combination of endogenous chitinolytic enzymes produced by oxyntic glands in the stomach and bacterial chitinases secreted in the colon (Cheng et al. 2022). Moreover, metagenomic data of fecal samples from captive giant anteater (*M. tridactyla*) individuals have revealed a chitin degradation potential in their gut microbiome (Cheng et al. 2022). Future genomic and metagenomic studies conducted in independent myrmecophagous mammals should allow deciphering the relative contributions of the host genome and its associated microbiome in the convergent adaptation to the myrmecophagous diet.

Material and Methods

Chitinase gene family tree reconstruction

Reconstruction of chitinase gene family evolution - The chitinase family in placental mammals appears to be composed of nine major paralogs (*CHIA1-5*, *CHIT1*, *CHI3L1*, *CHI3L2*, *OVGP1*). Mammalian sequences similar to the protein sequence of the human chitinase gene (NP_970615.2) were searched in the NCBI non-redundant protein database using BLASTP (E-value < 10). The protein sequences identified by BLASTP were then imported into Geneious Prime (Kearse et al. 2012) and aligned using MAFFT v7.450 (Katoh and Standley 2013) with the default parameters. Preliminary gene trees were then reconstructed with maximum likelihood using RAxML v8.2.11 (Stamatakis 2014) under the LG+G4 model (Le and Gascuel 2008) as implemented in Geneious Prime. From the reconstructed tree, the sequences were filtered according to the following criteria: (1) fast-evolving sequences with an E-value greater than zero and not belonging to the chitinase family were excluded; (2) in cases of multiple isoforms, only the longest was retained; (3) sequences whose length represented less than at least 50% of the total alignment length were removed; (4) in case of identical sequences from the same species the longest was kept; and (5) sequences labeled as "Hypothetical protein" and "Predicted: low quality protein" were discarded. This procedure resulted in a dataset containing 528 mammalian sequences that were realigned using MAFFT. This alignment was then cleaned up by removing sites not present in at least 50% of the sequences resulting in a total length of 460 amino acid sites. A maximum likelihood tree was then reconstructed with RAxML-NG v0.9.0 (Kozlov et al. 2019) using 10 tree searches starting from maximum parsimony trees under the LG+G8+F model. The species tree of the 143 mammal species represented in our dataset was reconstructed based on *COI* sequences extracted from the BOLD system database v4 (Ratnasingham and Hebert 2007) by searching for “Chordata” sequences in the “Taxonomy” section. Sequences were aligned using MAFFT, the phylogeny was inferred with RAxML and the topology was then adjusted manually based on the literature to correct ancient relationships. To determine the optimal rooting scheme, a rapid reconciliation between the resulting gene tree and species tree was performed using the TreeRecs reconciliation algorithm based on maximum parsimony (Comte et al. 2020) as implemented in SeaView v5.0.2 (Gouy et al. 2010). The final chitinase gene family tree was produced using the maximum likelihood gene family tree reconciliation approach implemented in GeneRax v.1.1.0 (Morel et al. 2020) using the TreeRecs reconciled tree as input (source and result

available from Zenodo). GeneRax can reconstruct duplications, losses and horizontal gene transfer events but since the latter are negligible in mammals, only gene duplications and losses have been modeled here (--rec-model UndatedDL) and the LG+G model was used.

Ancestral sequence reconstructions - Ancestral sequences of the different paralogs were reconstructed from the reconciled tree using RAxML-NG (--ancestral function, --model LG+G8+F). The sequences were then aligned in Geneious Prime with MAFFT (source and result files available from Zenodo). Given that active chitinases are characterized by a catalytic site with a conserved amino acid motif (DXXDXDXE; Olland et al. 2009; Hamid et al. 2013), this motif was compared among all available species. Additionally, the six conserved cysteine residues responsible for chitin binding (Tjoelker et al. 2000; Olland et al. 2009) were also investigated.

Chitinase gene synteny comparisons - The synteny of the nine chitinase paralogs was compared between the two focal ant-eating species in our global transcriptomic analysis (*T. tetradactyla* and *M. javanica*), an insectivorous xenarthran species (*D. novemcinctus*), an insectivorous primate species with five functional *CHIA* genes (*Carlito syrichta*) and human (*Homo sapiens*). For *H. sapiens*, synteny information was added from Emerling et al. (2018) and completed by using Genomicus v100.01 (Nguyen et al. 2022). For *C. syrichta* and *D. novemcinctus*, genome assemblies have been downloaded from the National Center for Biotechnology Information (NCBI) and from the DNA Zoo (Choo et al. 2016; Dudchenko et al. 2017) for *M. javanica* and *T. tetradactyla*. Synteny information was retrieved by blasting (*megablast*) the different CDS sequences against these assemblies. Scaffold/contig names, positions and direction of BLAST hits were retrieved to compare their synteny (source and result files available from Zenodo). Genes with negative BLAST results were considered probably not functional or absent.

Transcriptome assemblies

Salivary gland transcriptomes - Biopsies of submandibular salivary glands (Gil et al. 2018) preserved in RNAlater were obtained from the Mammalian Tissue Collection of the Institut des Sciences de l'Evolution de Montpellier (ISEM) and the JAGUARS collection for 17 individuals representing 12 placental mammal species (Table 1). Total RNA was extracted from individual salivary gland tissue samples using the RNeasy extraction kit (Qiagen,

585 Germany). Then, RNA-seq library construction and Illumina sequencing on a HiSeq 2500
586 system using paired-end 2x125bp reads were conducted by the Montpellier GenomiX
587 platform (MGX) resulting in 17 newly produced salivary gland transcriptomes. This sampling
588 was completed with the 13 mammalian salivary gland transcriptomes available as paired-end
589 Illumina sequencing reads in the Short Read Archive (SRA) of the NCBI as of April 15th,
590 2019 representing an additional 11 species (Table 1). This taxon sampling includes
591 representatives from all major mammal superorders Afrotheria (n = 3), Xenarthra (n = 4),
592 Euarchontoglires (n = 3), and Laurasiatheria (n = 13) and covers six different diet categories:
593 carnivory (n = 4), frugivory (n = 1), herbivory (n = 2), insectivory (n = 4), myrmecophagy (n
594 = 6), and omnivory (n = 6). Four of the five lineages in which myrmecophagous mammals
595 evolved are represented: southern aardwolf (*P. cristatus*, Carnivora), Malayan pangolin (*M.*
596 *javanica*, Pholidota), southern naked-tailed armadillo (*C. unicinctus*, Cingulata), giant
597 anteater (*M. tridactyla*, Pilosa), and southern tamandua (*T. tetradactyla*, Pilosa). Species
598 replicates in the form of different individuals were collected for the southern tamandua (*T.*
599 *tetradactyla*; n = 3), the nine-banded armadillo (*D. novemcinctus*; n = 3), and the Malayan
600 pangolin (*M. javanica*; n = 2). We unfortunately were not able to obtain fresh salivary gland
601 samples from the aardvark (*O. afer*, Tubulidentata), the only missing myrmecophagous
602 lineage in our sampling.
603
604

Table 1: Detailed information on the tissues sequenced or retrieved from public databases for the project.

Sample name	Species	Tissue type	Individual name	Sex	Source	Country of origin	Study
CABuniCU04	*Cabassous unicinctus*	Salivary gland (submandibular)	M2757	Male	JAGUARS	French Guiana	This study
SRR5889344	*Canis lupus familiaris*	Salivary gland	NA	NA	SRA NCBI	USA	Broad Institute, unpublished
DASnovFK06A	*Dasypus novemcinctus*	Salivary gland (submandibular)	FK06	NA	ISEM	USA	This study
DASnovFK06C	*Dasypus novemcinctus*	Salivary gland (submandibular)	FK06	NA	ISEM	USA	This study
DASnovFK08A	*Dasypus novemcinctus*	Salivary gland (submandibular)	FK08	NA	ISEM	USA	This study
SRR494766	*Dasypus novemcinctus*	Liver	0986	Male	SRA NCBI	USA	Broad Institute, unpublished
SRR494767	*Dasypus novemcinctus*	Spleen	0986	Male	SRA NCBI	USA	Broad Institute, unpublished
SRR494768	*Dasypus novemcinctus*	Spleen	0986	Male	SRA NCBI	USA	Broad Institute, unpublished
SRR494769	*Dasypus novemcinctus*	Heart	0986	Male	SRA NCBI	USA	Broad Institute, unpublished
SRR494770	*Dasypus novemcinctus*	Muscle	0986	Male	SRA NCBI	USA	Broad Institute, unpublished
SRR494771	*Dasypus novemcinctus*	Muscle	0986	Male	SRA NCBI	USA	Broad Institute, unpublished
SRR494772	*Dasypus novemcinctus*	Colon	0986	Male	SRA NCBI	USA	Broad Institute, unpublished
SRR494773	*Dasypus novemcinctus*	Heart	0986	Male	SRA NCBI	USA	Broad Institute, unpublished
SRR494774	*Dasypus novemcinctus*	Colon	0986	Male	SRA NCBI	USA	Broad Institute, unpublished
SRR494775	*Dasypus novemcinctus*	Kidney	0986	Male	SRA NCBI	USA	Broad Institute, unpublished
SRR494776	*Dasypus novemcinctus*	Lung	0986	Male	SRA NCBI	USA	Broad Institute, unpublished
SRR494777	*Dasypus novemcinctus*	Cerebellum	0986	Male	SRA NCBI	USA	Broad Institute, unpublished
SRR494778	*Dasypus novemcinctus*	Liver	0986	Male	SRA NCBI	USA	Broad Institute, unpublished
SRR494779	*Dasypus novemcinctus*	Kidney	0986	Male	SRA NCBI	USA	Broad Institute, unpublished
SRR494780	*Dasypus novemcinctus*	Cerebellum	0986	Male	SRA NCBI	USA	Broad Institute, unpublished
SRR494781	*Dasypus novemcinctus*	Lung	0986	Male	SRA NCBI	USA	Broad Institute, unpublished
SRR6206903	*Dasypus novemcinctus*	Heart	NA	NA	SRA NCBI	USA	Chen *et al*, 2019
SRR6206908	*Dasypus novemcinctus*	Kidney	NA	NA	SRA NCBI	USA	Chen *et al*, 2019
SRR6206913	*Dasypus novemcinctus*	Liver	NA	NA	SRA NCBI	USA	Chen *et al*, 2019
SRR6206918	*Dasypus novemcinctus*	Lung	NA	NA	SRA NCBI	USA	Chen *et al*, 2019
SRR6206923	*Dasypus novemcinctus*	Muscle	NA	NA	SRA NCBI	USA	Chen *et al*, 2019
SRR606902	*Desmodus rotundus*	Salivary gland	NA	NA	SRA NCBI	Brazil	National Institute of Allergy and Infectious Diseases, unpublished
SRR606908	*Desmodus rotundus*	Salivary gland	NA	NA	SRA NCBI	Brazil	National Institute of Allergy and Infectious Diseases, unpublished
SRR606911	*Desmodus rotundus*	Salivary gland	NA	NA	SRA NCBI	Brazil	National Institute of Allergy and Infectious Diseases, unpublished
ELEmyuNA02	*Elephantulus myurus*	Salivary gland (submandibular)	TDR	NA	ISEM	South Africa	This study
ERIeurRA02	*Erinaceus europaeus*	Salivary gland (submandibular)	RA03	NA	ISEM	France	This study
SRR3218717	*Felis catus*	Salivary gland	NA	Female	SRA NCBI	USA	Visser *et al* , 2019
GENgenRA01	*Genetta genetta*	Salivary gland (submandibular)	RA02	NA	ISEM	France	This study
SRR1957200	*Homo sapiens*	Salivary gland	NA	NA	SRA NCBI	USA	Duff *et al*, 2015
SRR1023040	*Macrotus californicus*	Salivary gland (submandibular)	NA	Male	SRA NCBI	USA	Texas Tech University; unpublished
SRR2547558	*Manis javanica*	Cerebellum	NA	Female	SRA NCBI	Malysia	Yusoff *et al*, 2016
SRR2561209	*Manis javanica*	Brain	NA	Female	SRA NCBI	Malysia	Yusoff *et al*, 2016
SRR2561211	*Manis javanica*	Heart	NA	Female	SRA NCBI	Malysia	Yusoff *et al*, 2016
SRR2561212	*Manis javanica*	Kidney	NA	Female	SRA NCBI	Malysia	Yusoff *et al*, 2016
SRR2561213	*Manis javanica*	Liver	NA	Female	SRA NCBI	Malysia	Yusoff *et al*, 2016
SRR2561214	*Manis javanica*	Lung	NA	Female	SRA NCBI	Malysia	Yusoff *et al*, 2016
SRR2561215	*Manis javanica*	Spleen	NA	Female	SRA NCBI	Malysia	Yusoff *et al*, 2016
SRR2561216	*Manis javanica*	Thymus	NA	Female	SRA NCBI	Malysia	Yusoff *et al*, 2016
SRR3923846	*Manis javanica*	Skin	NA	Female	SRA NCBI	Malysia	Yusoff *et al*, 2016
SRR5337837	*Manis javanica*	Salivary gland	NA	Female	SRA NCBI	China	Ma *et al*, 2017
SRR5341161	*Manis javanica*	Liver	NA	Female	SRA NCBI	China	Ma *et al*, 2017
SRR5328124	*Manis javanica*	Small intestine	NA	Female	SRA NCBI	China	Ma *et al*, 2017
SRR5837767	*Manis javanica*	Muscle	NA	NA	SRA NCBI	China	Jiangsu Normal University; unpublished
SRR7641079	*Manis javanica*	Pancreas	NA	Female	SRA NCBI	China	Ma *et al*, 2019
SRR7641080	*Manis javanica*	Liver	NA	Female	SRA NCBI	China	Ma *et al*, 2019
SRR7641081	*Manis javanica*	Liver	NA	Female	SRA NCBI	China	Ma *et al*, 2019
SRR7641082	*Manis javanica*	Pancreas	NA	Female	SRA NCBI	China	Ma *et al*, 2019
SRR7641083	*Manis javanica*	Tongue	NA	Female	SRA NCBI	China	Ma *et al*, 2019
SRR7641084	*Manis javanica*	Salivary gland	NA	Female	SRA NCBI	China	Ma *et al*, 2019
SRR7641085	*Manis javanica*	Stomach	NA	Female	SRA NCBI	China	Ma *et al*, 2019
SRR7641086	*Manis javanica*	Stomach	NA	Female	SRA NCBI	China	Ma *et al*, 2019
SRR7641087	*Manis javanica*	Liver	NA	Female	SRA NCBI	China	Ma *et al*, 2019
SRR7641088	*Manis javanica*	Liver	NA	Female	SRA NCBI	China	Ma *et al*, 2019
SRR7641089	*Manis javanica*	Large intestine	NA	Female	SRA NCBI	China	Ma *et al*, 2019
SRR7641090	*Manis javanica*	Large intestine	NA	Female	SRA NCBI	China	Ma *et al*, 2019
MELmelRA01	*Meles meles*	Salivary gland (submandibular)	RA01	Female	ISEM	France	This study
MICspMV01	*Microgale brevicaudata*	Salivary gland (submandibular)	MV03	NA	ISEM	Madagascar	This study
SRR5878900	*Mus musculus*	Salivary gland	NA	NA	SRA NCBI	USA	Metwalli *et al*, 2018
MYOcoyPH03	*Myocastor coypus*	Salivary gland (submandibular)	Myo2	NA	ISEM	France	This study
MYRtriCAY01	*Myrmecophaga tridactyla*	Salivary gland (submandibular)	M3023	Male	JAGUARS	French Guiana	This study
ERR2076303	*Ovis aries*	Salivary gland	NA	Female	SRA NCBI	USA	Clark *et al*, 2017
PROcri01_S29	*Proteles cristatus*	Salivary gland (submandibular)	TS307	NA	ISEM	South Africa	This study
PROcri01_S2	*Proteles cristatus*	Salivary gland (submandibular)	TS307	NA	ISEM	South Africa	This study
SRR3056926	*Rattus norvegicus*	Salivary gland	NA	NA	SRA NCBI	USA	Barasch *et al*, 2017
SRR5802558	*Sus scrofa*	Salivary gland	NA	Male	SRA NCBI	China	China Agricultural Univeristy ; unpublished
TAMtetFC01	*Tamandua tetradactyla*	Salivary gland (submandibular)	T7380	Female	ISEM	French Guiana	This study
TAMtetFC04	*Tamandua tetradactyla*	Spleen	T7380	Female	ISEM	French Guiana	This study
TAMtetR05	*Tamandua tetradactyla*	Testis	M2813	Male	JAGUARS	French Guiana	This study
TAMtetB01	*Tamandua tetradactyla*	Salivary gland (submandibular)	M2813	Male	JAGUARS	French Guiana	This study
TAMtetB07	*Tamandua tetradactyla*	Tongue	M2813	Male	JAGUARS	French Guiana	This study
TAMtetTT49	*Tamandua tetradactyla*	Salivary gland (submandibular)	M3075	Male	JAGUARS	French Guiana	This study
TAMtetTT55	*Tamandua tetradactyla*	Tongue	M3075	Male	JAGUARS	French Guiana	This study
TAMtetTT59	*Tamandua tetradactyla*	Liver	M3075	Male	JAGUARS	French Guiana	This study
TAMtetTT62	*Tamandua tetradactyla*	Spleen	M3075	Male	JAGUARS	French Guiana	This study
TAMtetTT70	*Tamandua tetradactyla*	Testis	M3075	Male	JAGUARS	French Guiana	This study
TAMtetTT73	*Tamandua tetradactyla*	Lung	M3075	Male	JAGUARS	French Guiana	This study
TAMtetTT75	*Tamandua tetradactyla*	Heart	M3075	Male	JAGUARS	French Guiana	This study
TAMtetTT78	*Tamandua tetradactyla*	Glandular stomach	M3075	Male	JAGUARS	French Guiana	This study
TAMtetTT79	*Tamandua tetradactyla*	Muscular stomach	M3075	Male	JAGUARS	French Guiana	This study
TAMtetTT99	*Tamandua tetradactyla*	Small intestine	M3075	Male	JAGUARS	French Guiana	This study
SETsetMV01	*Tenrec ecaudatus*	Salivary gland (submandibular)	MV01	NA	ISEM	Madagascar	This study
SRR1663490	*Uroderma bilobatum*	Salivary gland (submandibular)	NA	Male	SRA NCBI	Uruguay	Feijoo *et al*, 2017

Transcriptomes from additional organs - Tissue biopsies from eight additional organs (testis, lungs, heart, spleen, tongue, stomach, liver, and small intestine) were sampled during dissections of three individuals of southern tamandua (*T. tetradactyla*; Table 1). Total RNA extractions from these RNAlater-preserved tissues, RNA-seq library construction, and sequencing were conducted as described above resulting in 12 newly generated transcriptomes. For comparative purposes, 21 additional transcriptomes of nine-banded armadillo (*D. novemcinctus*) representing eight organs and 24 transcriptomes of Malayan pangolin (*M. javanica*) representing 16 organs were downloaded from SRA (Table 1).

Comparative transcriptomics

Transcriptome assemblies and quality control - Adapters and low quality reads were removed from raw sequencing data using fastp v0.19.6 (Chen et al. 2018) using default parameters except for the PHRED score which was defined as "--qualified_quality_phred ≥ 15", as suggested by (MacManes 2014). Then, *de novo* assembly was performed on each individual transcriptome sample using Trinity v2.8.4 (Grabherr et al. 2011) using default parameters. For each of the 28 salivary gland transcriptomes, completeness was assessed by the presence of Benchmark Universal Single Copy Orthologs (BUSCOs) based on a dataset of 4,104 single-copy orthologs conserved in over 90% of mammalian species (Waterhouse et al. 2018). This pipeline evaluates the percentage of complete, duplicated, fragmented and missing single copy orthologs within each transcriptome.

Transcriptome annotation and orthogroup inference - The transcriptome assemblies were annotated following the pipeline implemented in assembly2ORF (https://github.com/ellefeg/assembly2orf). This pipeline combines evidence-based and gene-model-based predictions. First, potential transcripts of protein-coding genes are extracted based on similarity searches (BLAST) against the peptides of Metazoa found in Ensembl (Yates et al. 2020). Then, using both protein similarity and exonerate functions (Slater and Birney 2005), a frameshift correction is applied to candidate transcripts. Candidate open reading frames (ORFs) are predicted using TransDecoder (https://github.com/TransDecoder/TransDecoder) and annotated based on homology information inferred from both BLAST and Hmmscan searches. Finally, to be able to compare the transcriptomes obtained from all species, we relied on the inference of gene orthogroups. The orthogroup inference for the translated candidate ORFs was performed using OrthoFinder v2 (Emms and Kelly 2019) using IQ-TREE (Nguyen et al. 2015) for gene tree reconstructions. For expression analyses, orthogroups containing more than 20 copies for at least one species were discarded.

Gene expression analyses - Quantification of transcript expression was performed on Trinity assemblies with Kallisto v.0.46.1 (Bray et al. 2016) using the *align_and_estimate_abundance.pl* script provided in the Trinity suite (Grabherr et al. 2011). Kallisto relies on pseudo-alignments of the reads to search for the original transcript of a read without looking for a perfect alignment (as opposed to classical quantification by counting

651 the reads aligned on the assembled transcriptome; Wolf 2013). Counts (raw number of
652 mapped reads) and the Transcripts Per kilobase Million are reported (result files available
653 from Zenodo). Based on the previously inferred orthogroups, orthogroup-level abundance
654 estimates were imported and summarized using tximport (Soneson et al. 2016). To minimize
655 sequencing depth variation across samples and gene outlier effect (a few highly and
656 differentially expressed genes may have strong and global influence on the total read count),
657 orthogroup-level raw reads counts were normalized using the median of the ratios of
658 observed counts using DESeq2 (Love et al. 2014) for orthogroups containing up to 20 gene
659 copies by species. The normalization incorporated the following conditions: diet and
660 taxonomic order.
661
662 *Chitinase expression in salivary glands* - The chitinase orthogroup inferred by OrthoFinder2
663 in previous analyses (see above) was extracted using BLASTX with the reference chitinase
664 database previously created. The amino acid sequences of this orthogroup were aligned using
665 MAFFT (--adjustdirection option) and gene tree inference was performed with IQ-TREE2
666 (LG+G4 model)(result files available from Zenodo). A visual verification of the alignments
667 and gene tree was performed to eliminate potential chimeric transcripts or erroneous
668 sequences. Then, the chitinase orthogroup was divided into sub-orthogroups for each
669 chitinase paralog (*CHIA1-5*, *CHIT1*, *CHI3L1*, *CHI3L2*, *OVGP1*). To take advantage of the
670 transcriptome-wide expression information for the expression standardization, these new
671 orthogroups were included in the previous orthogroup-level abundance matrix estimates and
672 the same normalization approach using DESeq2 was conducted. Finally, gene-level
673 abundance estimates for all chitinase paralogs were extracted and compared with a log2 scale.
674

675 **Data and Resource Availability**

676 Raw RNAseq Illumina reads have been submitted to the Short Read Archive (SRA) of the
677 National Center for Biotechnology Information (NCBI) and are available
678 under BioProject number PRJNXXXXX. Transcriptome assemblies, phylogenetic
679 datasets, corresponding trees, and other supplementary materials are available from
680 zenodo.org (DOI: 10.5281/zenodo.7355330).
681
682

Acknowledgments

We would like to thank Hugues Parrinello (Montpellier GenomiX platform) for advice on RNAseq and Marie Sémon for providing useful advice on RNAseq statistical analyses. We are also indebted to Frank Knight, Mark Scherz, Miguel Vences, Andolalao Rakotoarison, Nico Avenant, Pierre-Henri Fabre, Quentin Martinez, Nathalie Delsuc, Aude Caizergues, Roxanne Schaub, Lionel Hautier, Fabien Condamine, Sérgio Ferreira-Cardoso, and François Catzeflis for their help with tissue sampling. Computational analyses benefited from the Montpellier Bioinformatics Biodiversity (MBB) platform. The JAGUARS collection is supported through a FEDER/ERDF grant attributed to Kwata NGO, funded by the European Union, the Collectivité Territoriale de Guyane, and the DEAL Guyane. This work has been supported by grants from the European Research Council (ConvergeAnt project: ERC-2015-CoG-683257) and Investissements d'Avenir of the Agence Nationale de la Recherche (CEBA: ANR-10-LABX-25-01; CEMEB: ANR-10-LABX-0004). This is contribution ISEM 2022-XXX of the Institut des Sciences de l'Evolution de Montpellier.

References

Algarra B, Han L, Soriano-Úbeda C, Avilés M, Coy P, Jovine L, Jiménez-Movilla M. 2016. The C-terminal region of OVGP1 remodels the zona pellucida and modifies fertility parameters. *Sci. Rep.* 6:32556.

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.

Arendt J, Reznick D. 2008. Convergence and parallelism reconsidered: What have we learned about the genetics of adaptation? *Trends Ecol. Evol.* 23:26–32.

Areshkov PO, Avdieiev SS, Balynska OV, LeRoith D, Kavsan VM. 2011. Two closely related human members of Chitinase-like family, CHI3L1 and CHI3L2, activate ERK1/2 in 293 and U373 cells but have the different influence on cell proliferation. *Int. J. Biol. Sci.* 8:39–48.

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

Boot RG, Blommaart EF, Swart E, Ghauharali-van der Vlugt K, Bijl N, Moe C, Place A, Aerts JM. 2001. Identification of a novel acidic mammalian chitinase distinct from

716 chitotriosidase. *J. Biol. Chem.* 276:6770–6778.
717 Bray NL, Pimentel H, Melsted P, Pachter L. 2016. Near-optimal probabilistic RNA-seq
718 quantification. *Nat. Biotechnol.* 34:525–527.
719 Buhi WC. 2002. Characterization and biological roles of oviduct-specific, oestrogen-
720 dependent glycoprotein. *Reproduction* 123:355–362.
721 Bussink AP, Speijer D, Aerts JMFG, Boot RG. 2007. Evolution of mammalian Chitinase(-
722 like) members of family 18 Glycosyl Hydrolases. *Genetics* 177:959–970.
723 Chen A-S, Taguchi T, Sakai K, Kikuchi K, Wang M-W, Miwa I. 2003. Antioxidant activities
724 of Chitobiose and Chitotriose. *Biol. Pharm. Bull.* 26:1326–1330.
725 Chen S, Zhou Y, Chen Y, Gu J. 2018. fastp: an ultra-fast all-in-one FASTQ preprocessor.
726 *Bioinformatics* 34:i884–i890.
727 Chen Y-H, Zhao H. 2019. Evolution of digestive enzymes and dietary diversification in birds.
728 *PeerJ* 7:e6840.
729 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,
730 Irwin DM, et al. 2022. Hologenomic insights into mammalian adaptations to
731 myrmecophagy. *Natl. Sci. Rev.*:nwac174.
732 Choo SW, Rayko M, Tan TK, Hari R, Komissarov A, Wee WY, Yurchenko AA, Kliver S,
733 Tamazian G, Antunes A. 2016. Pangolin genomes and the evolution of mammalian
734 scales and immunity. *Genome Res.* 26:1312–1322.
735 Christin P-A, Weinreich DM, Besnard G. 2010. Causes and evolutionary significance of
736 genetic convergence. *Trends Genet.* 26:400–405.
737 Cole TL, Zhou C, Fang M, Pan H, Ksepka DT, Fiddaman SR, Emerling CA, Thomas DB, Bi
738 X, Fang Q. 2022. Genomic insights into the secondary aquatic transition of penguins.
739 *Nat. Commun.* 13:3912.
740 Comte N, Morel B, Hasić D, Guéguen L, Boussau B, Daubin V, Penel S, Scornavacca C,
741 Gouy M, Stamatakis A, et al. 2020. Treerecs: an integrated phylogenetic tool, from
742 sequences to reconciliations. *Bioinformatics* 36:4822–4824.
743 Conway Morris S. 1999. The crucible of creation: The Burgess Shale and the rise of animals.
744 Oxford, New York: Oxford University Press
745 Delsuc F, Metcalf JL, Wegener Parfrey L, Song SJ, González A, Knight R. 2014.
746 Convergence of gut microbiomes in myrmecophagous mammals. *Mol. Ecol.* 23:1301–
747 1317.
748 Delsuc F, Scally M, Madsen O, Stanhope MJ, de Jong WW, Catzeflis FM, Springer MS,
749 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.

Dudchenko O, Batra SS, Omer AD, Nyquist SK, Hoeger M, Durand NC, Shamim MS, Machol I, Lander ES, Aiden AP. 2017. De novo assembly of the *Aedes aegypti* genome using Hi-C yields chromosome-length scaffolds. *Science* 356:92–95.

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

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

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

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.

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

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

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

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

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

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

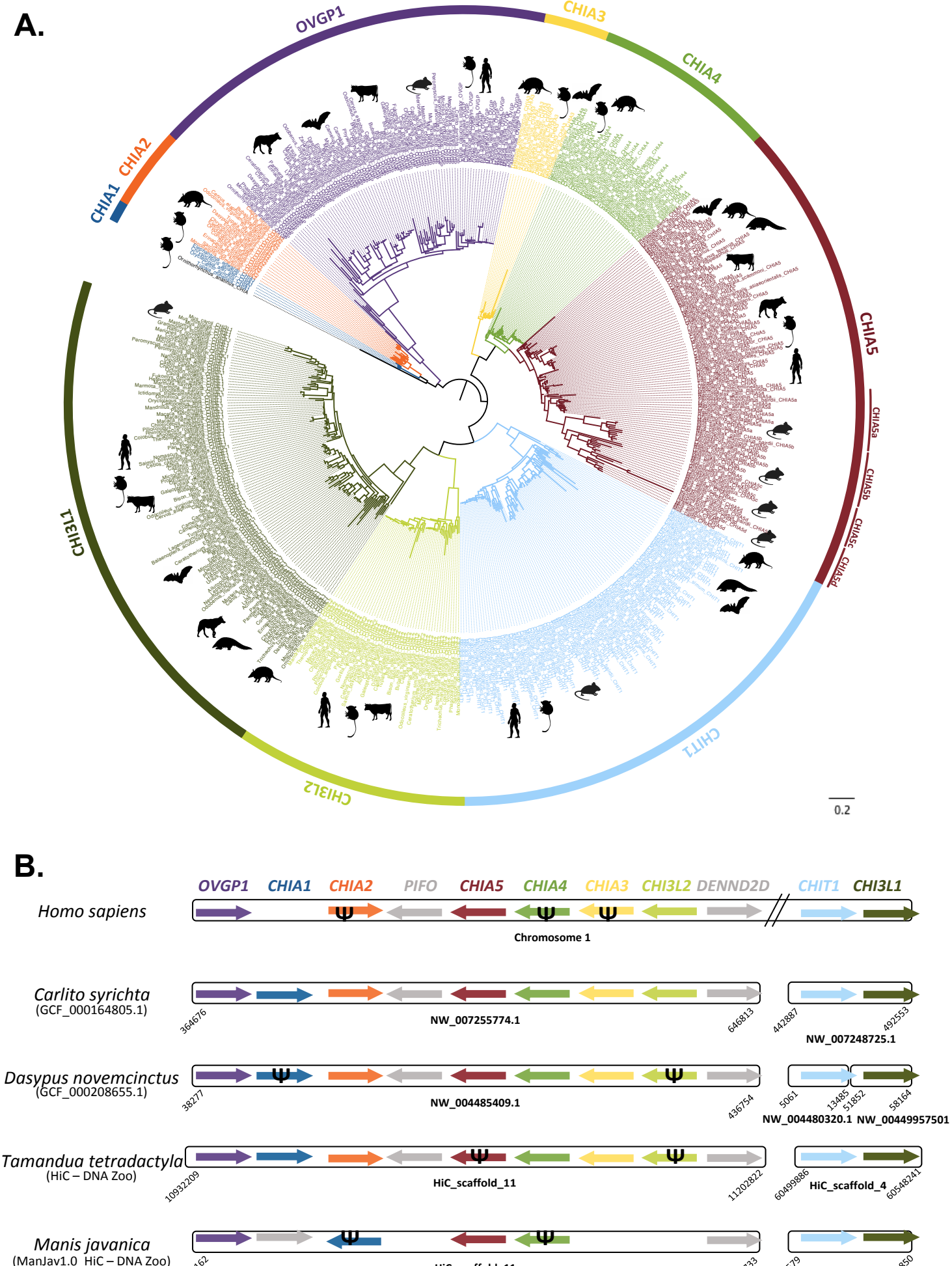
784 Gould SJ. 1990. Wonderful life: The Burgess Shale and the nature of history. WW Norton &
785 Company
786 Gould SJ. 2002. The Structure of Evolutionary Theory. Harvard University Press
787 Gouy M, Guindon S, Gascuel O. 2010. SeaView version 4: A multiplatform graphical user
788 interface for sequence alignment and phylogenetic tree building. *Mol. Biol. Evol.*
789 27:221–224.
790 Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I, Adiconis X, Fan L,
791 Raychowdhury R, Zeng Q, et al. 2011. Trinity: reconstructing a full-length
792 transcriptome without a genome from RNA-Seq data. *Nat. Biotechnol.* 29:644–652.
793 Hamid R, Khan MA, Ahmad M, Ahmad MM, Abdin MZ, Musarrat J, Javed S. 2013.
794 Chitinases: An update. *J. Pharm. Bioallied Sci.* 5:21–29.
795 Hussain M, Wilson JB. 2013. New paralogues and revised time line in the expansion of the
796 vertebrate GH18 family. *J. Mol. Evol.* 76:240–260.
797 Jacob F. 1977. Evolution and tinkering. *Science* 196:1161–1166.
798 Janiak MC, Chaney ME, Tosi AJ. 2018. Evolution of acidic mammalian chitinase genes
799 (CHIA) is related to body mass and insectivory in Primates. *Mol. Biol. Evol.* 35:607–
800 622.
801 Jeuniaux C. 1961. Chitinase: An addition to the list of hydrolases in the digestive tract of
802 vertebrates. *Nature* 192:135–136.
803 Jeuniaux C. 1966. [111] Chitinases. In: Methods in enzymology. Vol. 8. Elsevier. p. 644–
804 650.
805 Jeuniaux C. 1971. On some biochemical aspects of regressive evolution in animals. In:
806 Biochemical evolution and the origin of life. E. Schoffeniels. p. 304–313.
807 Jeuniaux C, Cornelius C. 1997. Distribution and activity of chitinolytic enzymes in the
808 digestive tract of birds and mammals. In: First international conference on
809 Chitin/Chitosan.
810 Katoh K, Standley DM. 2013. MAFFT multiple sequence alignment software Version 7:
811 Improvements in performance and usability. *Mol. Biol. Evol.* 30:772–780.
812 Kearse M, Moir R, Wilson A, Stones-Havas S, Cheung M, Sturrock S, Buxton S, Cooper A,
813 Markowitz S, Duran C, et al. 2012. Geneious Basic: An integrated and extendable
814 desktop software platform for the organization and analysis of sequence data.
815 *Bioinformatics* 28:1647–1649.
816 Koepfli K-P, Jenks SM, Eizirik E, Zahirpour T, Van Valkenburgh B, Wayne RK. 2006.
817 Molecular systematics of the Hyaenidae: relationships of a relictual lineage resolved

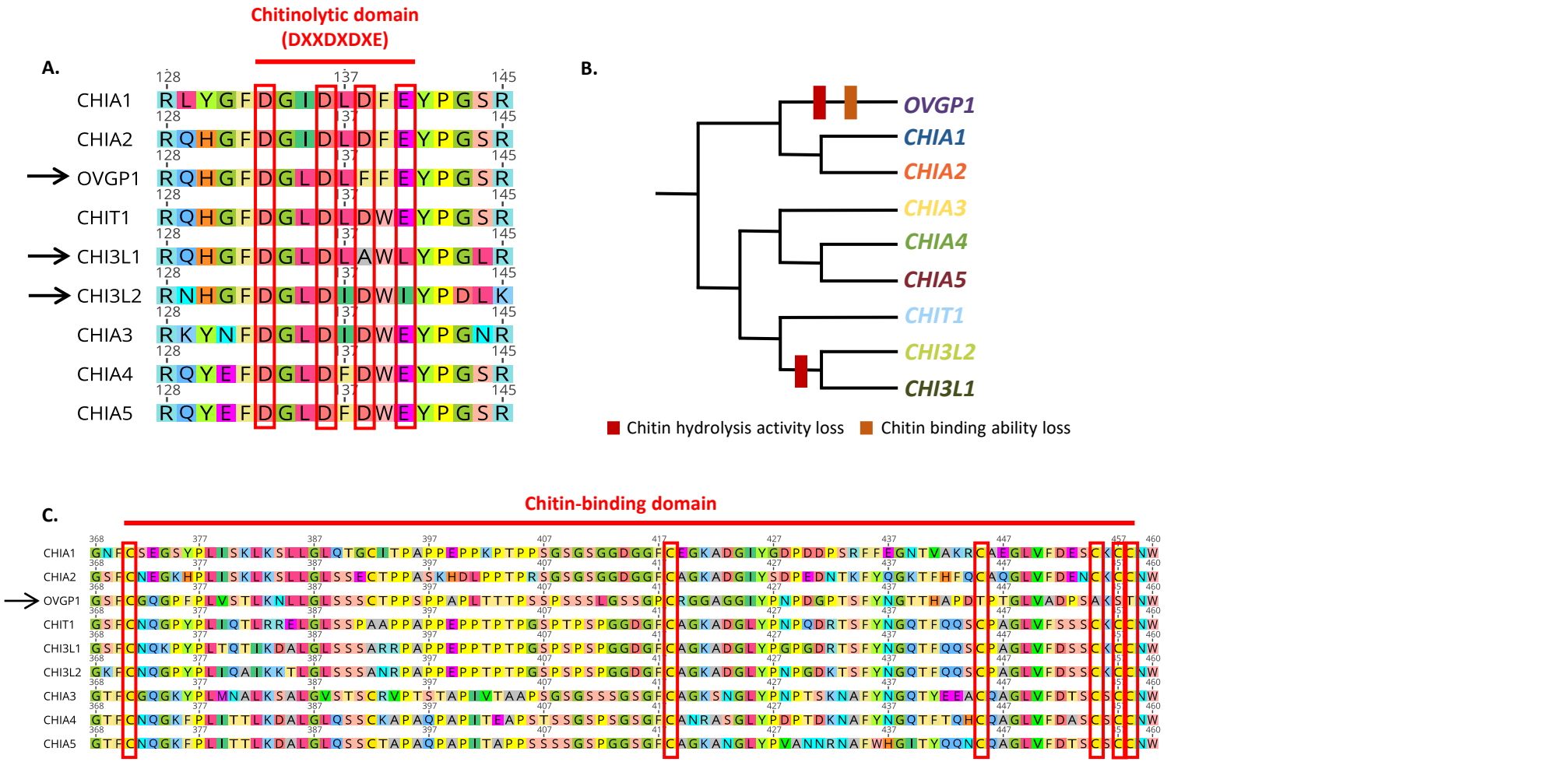
818 by a molecular supermatrix. *Mol. Phylogenet. Evol.* 38:603–620.
819 Kozlov AM, Darriba D, Flouri T, Morel B, Stamatakis A. 2019. RAxML-NG: a fast, scalable
820 and user-friendly tool for maximum likelihood phylogenetic inference. *Bioinformatics*
821 35:4453–4455.
822 Laheri S, Ashary N, Bhatt P, Modi D. 2018. Oviductal glycoprotein 1 (OVGP1) is expressed
823 by endometrial epithelium that regulates receptivity and trophoblast adhesion. *J.*
824 *Assist. Reprod. Genet.* 35:1419–1429.
825 Le SQ, Gascuel O. 2008. An improved general amino acid replacement matrix. *Mol. Biol.*
826 *Evol.* 25:1307–1320.
827 Lee CG, Da Silva CA, Dela Cruz CS, Ahangari F, Ma B, Kang M-J, He C-H, Takyar S, Elias
828 JA. 2011. Role of chitin and Chitinase/Chitinase-like proteins in inflammation, tissue
829 remodeling, and injury. *Annu. Rev. Physiol.* 73:479–501.
830 Losos JB. 2011. Convergence, adaptation, and constraint. *Evol. Int. J. Org. Evol.* 65:1827–
831 1840.
832 Losos JB. 2018. Improbable destinies: Fate, chance, and the future of evolution. Penguin
833 Love MI, Huber W, Anders S. 2014. Moderated estimation of fold change and dispersion for
834 RNA-seq data with DESeq2. *Genome Biol.* 15:550.
835 Ma J-E, Jiang H-Y, Li L-M, Zhang X-J, Li G-Y, Li H-M, Jin X-J, Chen J-P. 2018b. The fecal
836 metagenomics of Malayan pangolins identifies an extensive adaptation to
837 myrmecophagy. *Front. Microbiol.* 9:2793.
838 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
839 sequencing of the full-length transcriptome of the Sunda pangolin (*Manis javanica*).
840 *Gene* 692:208–216.
841 Ma J-E, Li L-M, Jiang H-Y, Zhang X-J, Li J, Li G-Y, Chen J-P. 2018a. Acidic mammalian
842 chitinase gene is highly expressed in the special oxyntic glands of *Manis javanica*.
843 *FEBS Open Bio* 8:1247–1255.
844 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.
845 Transcriptomic analysis identifies genes and pathways related to myrmecophagy in
846 the Malayan pangolin (*Manis javanica*). *PeerJ* 5:e4140.
847 MacManes M. 2014. On the optimal trimming of high-throughput mRNA sequence data.
848 *Front. Genet.* 5:13.
849 McGhee GR. 2011. Convergent evolution: Limited forms most beautiful. MIT Press
850 McNab BK. 1984. Physiological convergence amongst ant-eating and termite-eating
851 mammals. *J. Zool.* 203:485–510.

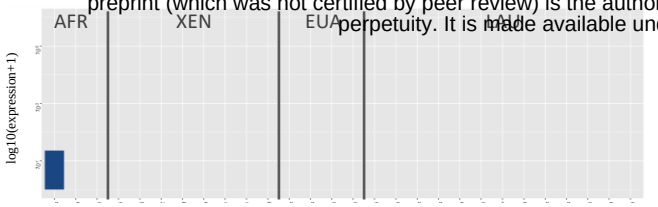
852 Meredith RW, Janečka JE, Gatesy J, Ryder OA, Fisher CA, Teeling EC, Goodbla A, Eizirik
853 E, Simão TLL, Stadler T, et al. 2011. Impacts of the Cretaceous terrestrial revolution
854 and KPg extinction on mammal diversification. *Science* 334:521–524.
855 Morel B, Kozlov AM, Stamatakis A, Szöllősi GJ. 2020. GeneRax: A tool for species-tree-
856 aware maximum likelihood-based gene family tree inference under gene duplication,
857 transfer, and loss. *Mol. Biol. Evol.* 37:2763–2774.
858 Nguyen L-T, Schmidt HA, von Haeseler A, Minh BQ. 2015. IQ-TREE: a fast and effective
859 stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol. Biol. Evol.*
860 32:268–274.
861 Nguyen NTT, Vincens P, Dufayard JF, Roest Crollius H, Louis A. 2022. Genomicus in 2022:
862 Comparative tools for thousands of genomes and reconstructed ancestors. *Nucleic*
863 *Acids Res.* 50:D1025–D1031.
864 Novacek MJ. 1992. Mammalian phylogeny: Shaking the tree. *Nature* 356:121–125.
865 O'Leary MA, Bloch JI, Flynn JJ, Gaudin TJ, Giallombardo A, Giannini NP, Goldberg SL,
866 Kraatz BP, Luo Z-X, Meng J, et al. 2013. The Placental Mammal Ancestor and the
867 Post–K-Pg Radiation of Placentals. *Science* 339:662–667.
868 Olland AM, Strand J, Presman E, Czerwinski R, Joseph-McCarthy D, Krykbaev R,
869 Schlingmann G, Chopra R, Lin L, Fleming M, et al. 2009. Triad of polar residues
870 implicated in pH specificity of acidic mammalian chitinase. *Protein Sci.* 18:569–578.
871 Phillips CJ, Phillips CD, Goecks J, Lessa EP, Sotero-Caio CG, Tandler B, Gannon MR,
872 Baker RJ. 2014. Dietary and flight energetic adaptations in a salivary gland
873 transcriptome of an insectivorous bat. *PLOS ONE* 9:e83512.
874 Pillai AS, Chandler SA, Liu Y, Signore AV, Cortez-Romero CR, Benesch JLP, Laganowsky
875 A, Storz JF, Hochberg GKA, Thornton JW. 2020. Origin of complexity in
876 haemoglobin evolution. *Nature* 581:480–485.
877 Ratnasingham S, Hebert PDN. 2007. bold: The barcode of life data system
878 (http://www.barcodinglife.org). *Mol. Ecol. Notes* 7:355–364.
879 Recklies AD, White C, Ling H. 2002. The chitinase 3-like protein human cartilage
880 glycoprotein 39 (HC-gp39) stimulates proliferation of human connective-tissue cells
881 and activates both extracellular signal-regulated kinase- and protein kinase B-
882 mediated signalling pathways. *Biochem. J.* 365:119–126.
883 Redford KH. 1987. Ants and termites as food. In: Genoways HH, editor. Current
884 Mammalogy. Boston, MA: Springer US. p. 349–399.
885 Reiss KZ. 2001. Using phylogenies to study convergence: the case of the ant-eating

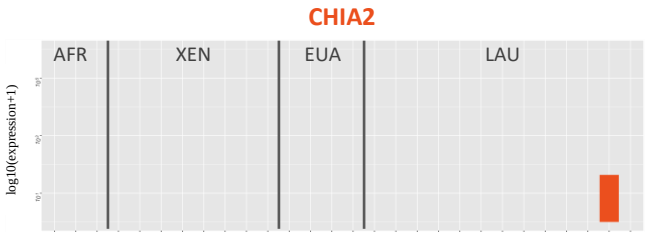
886 mammals. *Am. Zool.* 41:507–525.
887 Saint-Dizier M, Marnier C, Tahir MZ, Grimard B, Thoumire S, Chastant-Maillard S,
888 Reynaud K. 2014. OVGP1 is expressed in the canine oviduct at the time and place of
889 oocyte maturation and fertilization. *Mol. Reprod. Dev.* 81:972–982.
890 Salgaonkar N, Prakash D, Nawani NN, Kapadnis BP. 2015. Comparative studies on ability of
891 N-acetylated chitooligosaccharides to scavenge reactive oxygen species and protect
892 DNA from oxidative damage. *Indian J. Biotechnol.* 14:186–192.
893 Sanders JG, Beichman AC, Roman J, Scott JJ, Emerson D, McCarthy JJ, Girguis PR. 2015.
894 Baleen whales host a unique gut microbiome with similarities to both carnivores and
895 herbivores. *Nat. Commun.* 6:8285.
896 Slater GSC, Birney E. 2005. Automated generation of heuristics for biological sequence
897 comparison. *BMC Bioinformatics* 6:31.
898 Smith SA, Robbins LW, Steiert JG. 1998. Isolation and characterization of a chitinase from
899 the nine-banded armadillo, *Dasypus novemcinctus*. *J. Mammal.* 79:486–491.
900 Soneson C, Love MI, Robinson MD. 2016. Differential analyses for RNA-seq: Transcript-
901 level estimates improve gene-level inferences. *F1000 Res.* 4:1521.
902 Springer MS, Meredith RW, Teeling EC, Murphy WJ. 2013. Technical Comment on "The
903 Placental Mammal Ancestor and the Post–K-Pg Radiation of Placentals." *Science*
904 341:613–613.
905 Stamatakis A. 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of
906 large phylogenies. *Bioinformatics* 30:1312–1313.
907 Strobel S, Roswag A, Becker NI, Trenczek TE, Encarnaçao JA. 2013. Insectivorous bats
908 digest chitin in the stomach using acidic mammalian chitinase. *PloS One* 8:e72770.
909 Tabata E, Itoigawa A, Koinuma T, Tayama H, Kashimura A, Sakaguchi M, Matoska V,
910 Bauer PO, Oyama F. 2022. Noninsect-based diet leads to structural and functional
911 changes of Acidic Chitinase in Carnivora. *Mol. Biol. Evol.* 39:msab331.
912 Tjoelker LW, Gosting L, Frey S, Hunter CL, Le Trong H, Steiner B, Brammer H, Gray PW.
913 2000. Structural and functional definition of the human chitinase chitin-binding
914 domain. *J. Biol. Chem.* 275:514–520.
915 Tucker R. 1958. Taxonomy of the salivary glands of vertebrates. *Syst. Biol.* 7:74–83.
916 Vandewege MW, Sotero-Caio CG, Phillips CD. 2020. Positive selection and gene expression
917 analyses from salivary glands reveal discrete adaptations within the ecologically
918 diverse bat family Phyllostomidae. *Genome Biol. Evol.* 12:1419–1428.
919 Wang K, Tian S, Galindo-González J, Dávalos LM, Zhang Y, Zhao H. 2020. Molecular

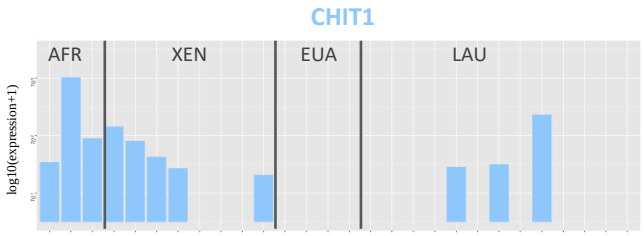
920 adaptation and convergent evolution of frugivory in Old World and neotropical fruit
921 bats. *Mol. Ecol.* 29:4366–4381.
922 Waterhouse RM, Seppey M, Simão FA, Manni M, Ioannidis P, Klioutchnikov G, Kriventseva
923 EV, Zdobnov EM. 2018. BUSCO applications from quality assessments to gene
924 prediction and phylogenomics. *Mol. Biol. Evol.* 35:543–548.
925 Westbury MV, Le Duc D, Duchêne DA, Krishnan A, Prost S, Rutschmann S, Grau JH, Dalén
926 L, Weyrich A, Norén K, et al. 2021. Ecological specialization and evolutionary
927 reticulation in extant Hyaenidae. *Mol. Biol. Evol.* 38:3884–3897.
928 Wolf JBW. 2013. Principles of transcriptome analysis and gene expression quantification: an
929 RNA-seq tutorial. *Mol. Ecol. Resour.* 13:559–572.
930 Xie VC, Pu J, Metzger BP, Thornton JW, Dickinson BC. 2021. Contingency and chance
931 erase necessity in the experimental evolution of ancestral proteins. *eLife* 10:e67336.
932 Yates AD, Achuthan P, Akanni W, Allen James, Allen Jamie, Alvarez-Jarreta J, Amode MR,
933 Armean IM, Azov AG, Bennett R, et al. 2020. Ensembl 2020. *Nucleic Acids Res.*
934 48:D682–D688.
935 Yusoff AM, Tan TK, Hari R, Koepfli K-P, Wee WY, Antunes A, Sitam FT, Rovie-Ryan JJ,
936 Karuppannan KV, Wong GJ. 2016. De novo sequencing, assembly and analysis of
937 eight different transcriptomes from the Malayan pangolin. *Sci. Rep.* 6:1–11.
938 Zhang F, Xu N, Yu Y, Wu S, Li S, Wang W. 2019. Expression profile of the digestive
939 enzymes of *Manis javanica* reveals its adaptation to diet specialization. *ACS Omega*
940 4:19925–19933.

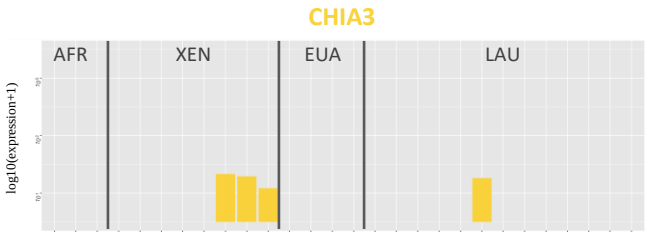

A.
OVGP1
CHIA3
CHIA4
CHIA1
CHIA2
CHIA5
CHIA5a
CHIA5b
CHIA5c
CHIA5d
CHI3L1
CHI3L2
CHIT1
0.2
B.
OVGP1 CHIA1 CHIA2 PIFO CHIA5 CHIA4 CHIA3 CHI3L2 DENND2D CHIT1 CHI3L1
Homo sapiens
Chromosome 1
Carlito syrichta
(GCF_000164805.1)
364676
NW_007255774.1
646813
442887
492553
NW_007248725.1
Dasypus novemcinctus
(GCF_000208655.1)
38277
NW_004485409.1
436754
5061
13485
51852
58164
NW_004480320.1
NW_00449957501
Tamandua tetradactyla
(HiC – DNA Zoo)
10932209
HiC_scaffold_11
11202822
60499886
HiC_scaffold_4
60548241
Manis javanica
(ManJav1.0_HiC – DNA Zoo)
94884162
HiC_scaffold_11
95065733
4626579
HiC_scaffold_3
4663850

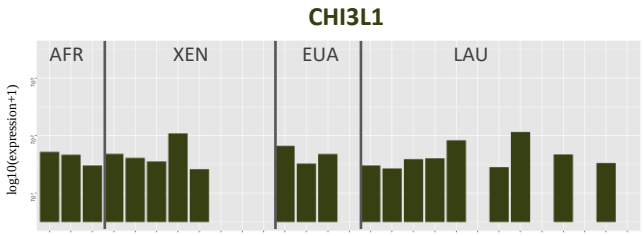

Chitinolytic domain
(DXXDXDXE)
A.
CHIA1
CHIA2
OVGP1
CHIT1
CHI3L1
CHI3L2
CHIA3
CHIA4
CHIA5
B.
OVGP1
CHIA1
CHIA2
CHIA3
CHIA4
CHIA5
CHIT1
CHI3L2
CHI3L1
Chitin hydrolysis activity loss
Chitin binding ability loss
Chitin-binding domain
C.
CHIA1
CHIA2
OVGP1
CHIT1
CHI3L1
CHI3L2
CHIA3
CHIA4
CHIA5

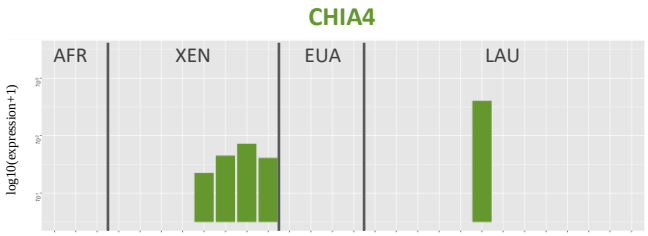


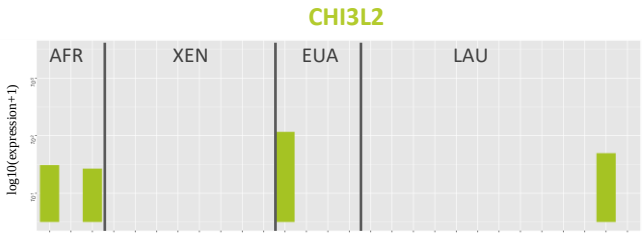


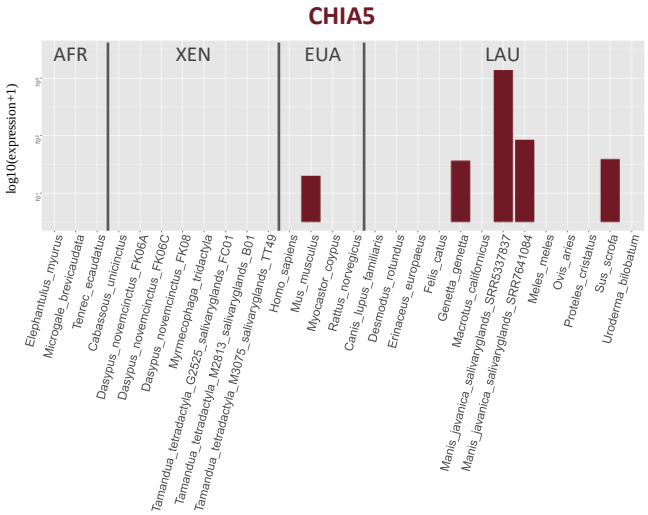


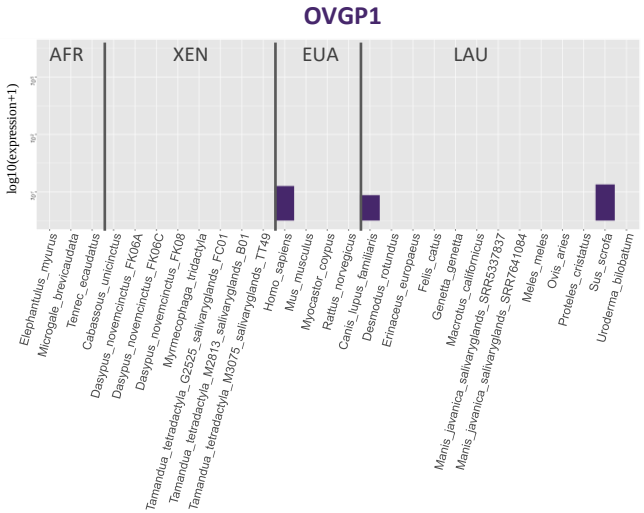




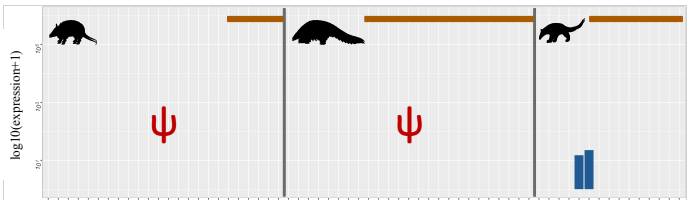




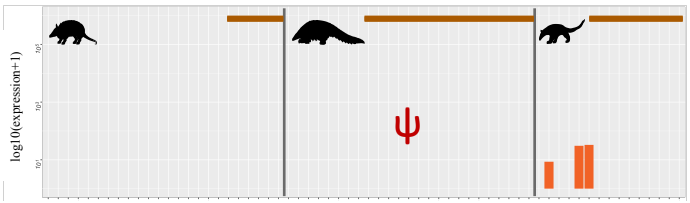




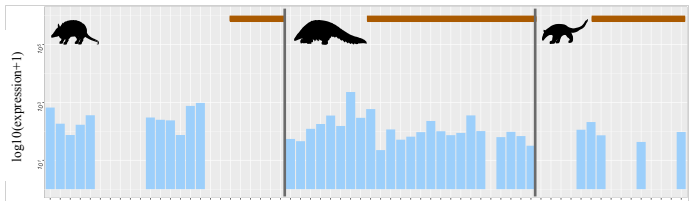
CHIA1



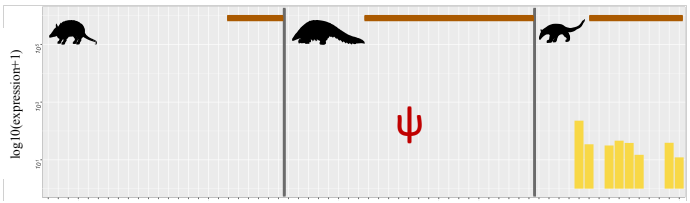
CHIA2



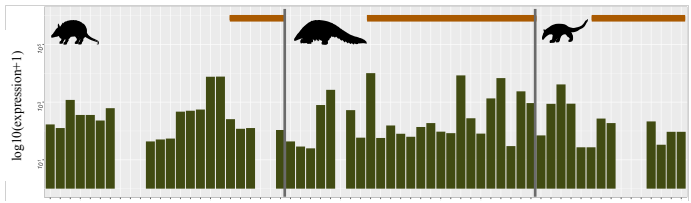
CHIT1



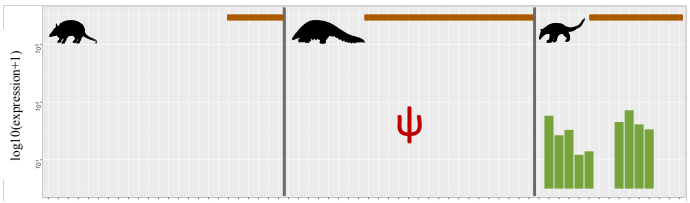
CHIA3



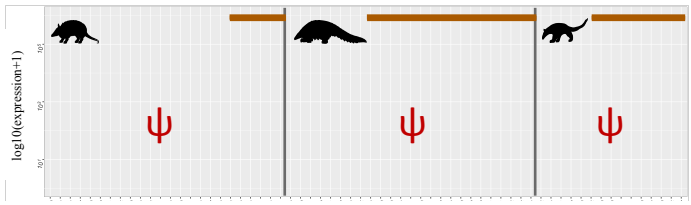
CHI3L1



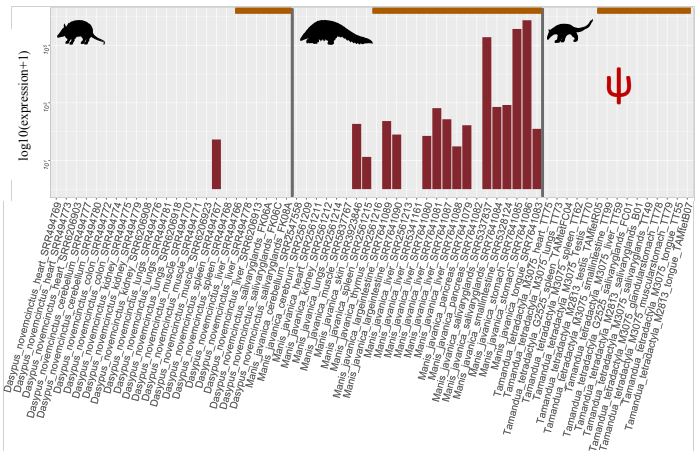
CHIA4



CHI3L2



CHIA5



OVGP1

Sample name	Species	Tissue type	Individual name	Sex	Source	Country of origin	Study
CABuniCU04	*Cabassous unicinctus*	Salivary gland (submandibular)	M2757	Male	JAGUARS	French Guiana	This study
SRR5889344	*Canis lupus familiaris*	Salivary gland	NA	NA	SRA NCBI	USA	Broad Institute, unpublished
DASnovFK06A	*Dasypus novemcinctus*	Salivary gland (submandibular)	FK06	NA	ISEM	USA	This study
DASnovFK06C	*Dasypus novemcinctus*	Salivary gland (submandibular)	FK06	NA	ISEM	USA	This study
DASnovFK08A	*Dasypus novemcinctus*	Salivary gland (submandibular)	FK08	NA	ISEM	USA	This study
SRR494766	*Dasypus novemcinctus*	Liver	0986	Male	SRA NCBI	USA	Broad Institute, unpublished
SRR494767	*Dasypus novemcinctus*	Spleen	0986	Male	SRA NCBI	USA	Broad Institute, unpublished
SRR494768	*Dasypus novemcinctus*	Spleen	0986	Male	SRA NCBI	USA	Broad Institute, unpublished
SRR494769	*Dasypus novemcinctus*	Heart	0986	Male	SRA NCBI	USA	Broad Institute, unpublished
SRR494770	*Dasypus novemcinctus*	Muscle	0986	Male	SRA NCBI	USA	Broad Institute, unpublished
SRR494771	*Dasypus novemcinctus*	Muscle	0986	Male	SRA NCBI	USA	Broad Institute, unpublished
SRR494772	*Dasypus novemcinctus*	Colon	0986	Male	SRA NCBI	USA	Broad Institute, unpublished
SRR494773	*Dasypus novemcinctus*	Heart	0986	Male	SRA NCBI	USA	Broad Institute, unpublished
SRR494774	*Dasypus novemcinctus*	Colon	0986	Male	SRA NCBI	USA	Broad Institute, unpublished
SRR494775	*Dasypus novemcinctus*	Kidney	0986	Male	SRA NCBI	USA	Broad Institute, unpublished
SRR494776	*Dasypus novemcinctus*	Lung	0986	Male	SRA NCBI	USA	Broad Institute, unpublished
SRR494777	*Dasypus novemcinctus*	Cerebellum	0986	Male	SRA NCBI	USA	Broad Institute, unpublished
SRR494778	*Dasypus novemcinctus*	Liver	0986	Male	SRA NCBI	USA	Broad Institute, unpublished
SRR494779	*Dasypus novemcinctus*	Kidney	0986	Male	SRA NCBI	USA	Broad Institute, unpublished
SRR494780	*Dasypus novemcinctus*	Cerebellum	0986	Male	SRA NCBI	USA	Broad Institute, unpublished
SRR494781	*Dasypus novemcinctus*	Lung	0986	Male	SRA NCBI	USA	Broad Institute, unpublished
SRR6206903	*Dasypus novemcinctus*	Heart	NA	NA	SRA NCBI	USA	Chen *et al,* 2019
SRR6206908	*Dasypus novemcinctus*	Kidney	NA	NA	SRA NCBI	USA	Chen *et al,* 2019
SRR6206913	*Dasypus novemcinctus*	Liver	NA	NA	SRA NCBI	USA	Chen *et al,* 2019
SRR6206918	*Dasypus novemcinctus*	Lung	NA	NA	SRA NCBI	USA	Chen *et al,* 2019
SRR6206923	*Dasypus novemcinctus*	Muscle	NA	NA	SRA NCBI	USA	Chen *et al,* 2019
SRR606902	*Desmodus rotundus*	Salivary gland	NA	NA	SRA NCBI	Brazil	National Institute of Allergy and Infectious Diseases, unpublished
SRR606908	*Desmodus rotundus*	Salivary gland	NA	NA	SRA NCBI	Brazil	National Institute of Allergy and Infectious Diseases, unpublished
SRR606911	*Desmodus rotundus*	Salivary gland	NA	NA	SRA NCBI	Brazil	National Institute of Allergy and Infectious Diseases, unpublished
ELEmyuNA02	*Elephantulus myurus*	Salivary gland (submandibular)	TDR	NA	ISEM	South Africa	This study
ERIeurRA02	*Erinaceus europaeus*	Salivary gland (submandibular)	RA03	NA	ISEM	France	This study
SRR3218717	*Felis catus*	Salivary gland	NA	Female	SRA NCBI	USA	Visser *et al*, 2019
GENgenRA01	*Genetta genetta*	Salivary gland (submandibular)	RA02	NA	ISEM	France	This study
SRR1957200	*Homo sapiens*	Salivary gland	NA	NA	SRA NCBI	USA	Duff *et al,* 2015
SRR1023040	*Macrotus californicus*	Salivary gland (submandibular)	NA	Male	SRA NCBI	USA	Texas Tech University; unpublished
SRR2547558	*Manis javanica*	Cerebellum	NA	Female	SRA NCBI	Malysia	Yusoff *et al,* 2016
SRR2561209	*Manis javanica*	Brain	NA	Female	SRA NCBI	Malysia	Yusoff *et al,* 2016
SRR2561211	*Manis javanica*	Heart	NA	Female	SRA NCBI	Malysia	Yusoff *et al,* 2016
SRR2561212	*Manis javanica*	Kidney	NA	Female	SRA NCBI	Malysia	Yusoff *et al,* 2016
SRR2561213	*Manis javanica*	Liver	NA	Female	SRA NCBI	Malysia	Yusoff *et al,* 2016
SRR2561214	*Manis javanica*	Lung	NA	Female	SRA NCBI	Malysia	Yusoff *et al,* 2016
SRR2561215	*Manis javanica*	Spleen	NA	Female	SRA NCBI	Malysia	Yusoff *et al,* 2016
SRR2561216	*Manis javanica*	Thymus	NA	Female	SRA NCBI	Malysia	Yusoff *et al,* 2016
SRR3923846	*Manis javanica*	Skin	NA	Female	SRA NCBI	Malysia	Yusoff *et al,* 2016
SRR5337837	*Manis javanica*	Salivary gland	NA	Female	SRA NCBI	China	Ma *et al,* 2017
SRR5341161	*Manis javanica*	Liver	NA	Female	SRA NCBI	China	Ma *et al,* 2017
SRR5328124	*Manis javanica*	Small intestine	NA	Female	SRA NCBI	China	Ma *et al,* 2017
SRR5837767	*Manis javanica*	Muscle	NA	NA	SRA NCBI	China	Jiangsu Normal University; unpublished
SRR7641079	*Manis javanica*	Pancreas	NA	Female	SRA NCBI	China	Ma *et al,* 2019
SRR7641080	*Manis javanica*	Liver	NA	Female	SRA NCBI	China	Ma *et al,* 2019
SRR7641081	*Manis javanica*	Liver	NA	Female	SRA NCBI	China	Ma *et al,* 2019
SRR7641082	*Manis javanica*	Pancreas	NA	Female	SRA NCBI	China	Ma *et al,* 2019
SRR7641083	*Manis javanica*	Tongue	NA	Female	SRA NCBI	China	Ma *et al,* 2019
SRR7641084	*Manis javanica*	Salivary gland	NA	Female	SRA NCBI	China	Ma *et al,* 2019
SRR7641085	*Manis javanica*	Stomach	NA	Female	SRA NCBI	China	Ma *et al,* 2019
SRR7641086	*Manis javanica*	Stomach	NA	Female	SRA NCBI	China	Ma *et al,* 2019
SRR7641087	*Manis javanica*	Liver	NA	Female	SRA NCBI	China	Ma *et al,* 2019
SRR7641088	*Manis javanica*	Liver	NA	Female	SRA NCBI	China	Ma *et al,* 2019
SRR7641089	*Manis javanica*	Large intestine	NA	Female	SRA NCBI	China	Ma *et al,* 2019
SRR7641090	*Manis javanica*	Large intestine	NA	Female	SRA NCBI	China	Ma *et al,* 2019
MELmelRA01	*Meles meles*	Salivary gland (submandibular)	RA01	Female	ISEM	France	This study
MICspMV01	*Microgale brevicaudata*	Salivary gland (submandibular)	MV03	NA	ISEM	Madagascar	This study
SRR5878900	*Mus musculus*	Salivary gland	NA	NA	SRA NCBI	USA	Metwalli *et al,* 2018
MYOcoyPH03	*Myocastor coypus*	Salivary gland (submandibular)	Myo2	NA	ISEM	France	This study
MYRtriCAY01	*Myrmecophaga tridactyla*	Salivary gland (submandibular)	M3023	Male	JAGUARS	French Guiana	This study
ERR2076303	*Ovis aries*	Salivary gland	NA	Female	SRA NCBI	USA	Clark *et al,* 2017
PROcri01_S29	*Proteles cristatus*	Salivary gland (submandibular)	TS307	NA	ISEM	South Africa	This study
PROcri01_S2	*Proteles cristatus*	Salivary gland (submandibular)	TS307	NA	ISEM	South Africa	This study
SRR3056926	*Rattus norvegicus*	Salivary gland	NA	NA	SRA NCBI	USA	Barasch *et al,* 2017
SRR5802558	*Sus scrofa*	Salivary gland	NA	Male	SRA NCBI	China	China Agricultural Univeristy ; unpublished
TAMtetFC01	*Tamandua tetradactyla*	Salivary gland (submandibular)	T7380	Female	ISEM	French Guiana	This study
TAMtetFC04	*Tamandua tetradactyla*	Spleen	T7380	Female	ISEM	French Guiana	This study
TAMtetR05	*Tamandua tetradactyla*	Testis	M2813	Male	JAGUARS	French Guiana	This study
TAMtetB01	*Tamandua tetradactyla*	Salivary gland (submandibular)	M2813	Male	JAGUARS	French Guiana	This study
TAMtetB07	*Tamandua tetradactyla*	Tongue	M2813	Male	JAGUARS	French Guiana	This study
TAMtetTT49	*Tamandua tetradactyla*	Salivary gland (submandibular)	M3075	Male	JAGUARS	French Guiana	This study
TAMtetTT55	*Tamandua tetradactyla*	Tongue	M3075	Male	JAGUARS	French Guiana	This study
TAMtetTT59	*Tamandua tetradactyla*	Liver	M3075	Male	JAGUARS	French Guiana	This study
TAMterTT62	*Tamandua tetradactyla*	Spleen	M3075	Male	JAGUARS	French Guiana	This study
TAMtetTT70	*Tamandua tetradactyla*	Testis	M3075	Male	JAGUARS	French Guiana	This study
TAMtetTT73	*Tamandua tetradactyla*	Lung	M3075	Male	JAGUARS	French Guiana	This study
TAMtetTT75	*Tamandua tetradactyla*	Heart	M3075	Male	JAGUARS	French Guiana	This study
TAMtetTT78	*Tamandua tetradactyla*	Glandular stomach	M3075	Male	JAGUARS	French Guiana	This study
TAMtetTT79	*Tamandua tetradactyla*	Muscular stomach	M3075	Male	JAGUARS	French Guiana	This study
TAMtetTT99	*Tamandua tetradactyla*	Small intestine	M3075	Male	JAGUARS	French Guiana	This study
SETsetMV01	*Tenrec ecaudatus*	Salivary gland (submandibular)	MV01	NA	ISEM	Madagascar	This study
SRR1663490	*Uroderma bilobatum*	Salivary gland (submandibular)	NA	Male	SRA NCBI	Uruguay	Feijoo *et al,* 2017