

Transcriptomic Data Reveal Divergent Paths of Chitinase Evolution Underlying Dietary Convergence in Anteaters and Pangolins

Rémi Allio ^{1,2,*}, Sophie Teullet ^{1,†}, Dave Lutgen ^{1,3,4,†}, Amandine Magdeleine¹, Rachid Koual¹, Marie-Ka Tilak ¹, Benoit de Thoisy ^{5,6}, Christopher A. Emerling ^{1,7}, Tristan Lefébure ⁸, Frédéric Delsuc ^{1,*}

¹ISEM, CNRS, IRD, Univ. Montpellier, Montpellier, France

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

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

⁴Swiss Ornithological Institute, Sempach, Switzerland

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

⁶Kwata NGO, Cayenne, French Guiana, France

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

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

*Corresponding authors: E-mails: remi.allio@inrae.fr; frederic.delsuc@umontpellier.fr.

†These authors contributed equally.

Accepted: January 05, 2025

Abstract

Ant-eating mammals represent a textbook example of convergent evolution. Among them, anteaters and pangolins exhibit the most extreme convergent phenotypes with complete tooth loss, elongated skulls, protruding tongues, and hypertrophied salivary glands producing large amounts of saliva. However, comparative genomic analyses have shown that anteaters and pangolins differ in their chitinase acidic 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 (*CHIA1-4*), Asian pangolins (*Manis* spp.) have only one functional paralog (*CHIA5*). Here, we performed a comparative transcriptomic analysis of salivary glands in 33 placental species, including 16 novel transcriptomes from ant-eating species and close relatives. Our results suggest that salivary glands play an important role in adaptation to an insect-based diet, as expression of different *CHIA* paralogs is observed in insectivorous species. Furthermore, convergently evolved pangolins and anteaters express different chitinases in their digestive tracts. In the Malayan pangolin, *CHIA5* is overexpressed in all major digestive organs, whereas in the southern tamandua, all four functional paralogs are expressed, at very high levels for *CHIA1* and *CHIA2* in the pancreas and for *CHIA3* and *CHIA4* in the salivary glands, stomach, liver, and pancreas. Overall, our results demonstrate that divergent molecular mechanisms within the chitinase acidic gene family underlie convergent adaptation to the ant-eating diet in pangolins and anteaters. This study highlights the role of historical contingency and molecular tinkering of the chitin digestive enzyme toolkit in this classic example of convergent evolution.

Key words: chitinases, convergent evolution, myrmecophagy, mammals, salivary glands, transcriptomics.

© The Author(s) 2025. Published by Oxford University Press on behalf of Society for Molecular Biology and Evolution.

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

Significance

This study shows that, despite numerous convergent morphological ant-eating adaptations, pangolins and anteaters have distinct molecular mechanisms for digesting chitin, a key component of their highly specialized diets. This divergence is reflected in their distinctive chitinase gene repertoires and expression patterns across digestive organs. These findings illustrate how historical contingency has shaped gene family evolution through molecular tinkering in these two convergent ant-eating lineages.

Introduction

Convergent evolution provides a fascinating window into the mechanisms by which similar environmental pressures shape the phenotypes of phylogenetically distant taxa. Indeed, despite the enormous diversity of organisms on Earth and the many potential ways to adapt to similar conditions, the strong deterministic force of natural selection has led to numerous instances of recurrent phenotypic adaptation (Losos 2011; McGhee 2011; Losos 2018). Although classical models of convergence at the molecular level often assume identical mutations in the same genes across species (Arendt and Reznick 2008), emerging evidence from comparative genomics and transcriptomics suggests that the recruitment of the same or similar genes and pathways may also lead to similar phenotypes across divergent lineages. For instance, convergent electric fish, which have evolved independently at least six times, provide a good illustration of the complexity of the selective process that follows from the interaction of contingency, constraints, and convergence (Zakon et al. 2006). In this case, the same genes have been independently recruited and differentially expressed in novel electric organs due to developmental constraints and their function subsequently adjusted by natural selection involving convergent amino acid substitutions in functionally important domains (Galant et al. 2014; Liu et al. 2019; Wang and Yang 2021). This suggests an important role for evolutionary constraints imposed by existing genomic architectures and developmental pathways, leading to the repeated use of similar genetic material in the origin of evolutionary novelties (Shubin et al. 2009). In this context, historical contingency often leads to evolutionary tinkering as natural selection works from available material (Jacob 1977). Thus, both historical contingency and deterministic evolution appear to have influenced the evolution of current biodiversity, and one of the key questions is to assess the relative influence of these two evolutionary processes (Blount et al. 2018).

As intuited by Jacob (1977), molecular tinkering appears to be particularly common and has indeed shaped the evolutionary history of a number of gene families (McGlothlin et al. 2016; Pillai et al. 2020; Xie et al. 2021). The particular evolutionary dynamics observed in gene families can lead to both evolutionary opportunities due to gene duplications

paired with the acquisition of a new function and also evolutionary constraints due to ancestral loss of function. A good example resides in the evolution of chitinase genes in placental mammals, which belong to the large Glycosyl Hydrolase Family 18 (GH18) gene family (Bussink et al. 2007; Funkhouser and Aronson 2007). Recent studies have shown that chitinase genes may play an important digestive function in insectivorous species (Emerling et al. 2018; Janiak et al. 2018; Wang et al. 2020; Cheng et al. 2023). Indeed, while the placental ancestor possessed five functional chitinase acidic (*CHIA*) paralogs, the evolution of this gene family was subsequently shaped through multiple pseudogenization events associated with dietary adaptation during the placental radiation (Emerling et al. 2018). Interestingly, the widespread gene loss observed in carnivorous and herbivorous lineages in particular resulted in a global positive correlation between the number of functional *CHIA* paralogs and the percentage of invertebrates in the diet across placentals. Indeed, mammals with a low proportion of insects in their diet present none or only 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; Fig. 1).

Myrmecophagous mammals, with more than 90% of their diet consisting of social insects (Redford 1987), have convergently evolved dietary adaptations 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). With regard to their chitinase gene repertoire, they are generally grouped with the most insectivorous species (Fig. 1). Specifically, the southern tamandua (*Tamandua tetradactyla*) and the armadillo (*Oryzomys*) indeed possess four (*CHIA1-4*) and five (*CHIA1-5*) functional paralogs, respectively. However, pangolins appear as a striking exception. Despite their strict myrmecophagous diet and many associated convergent features shared with other myrmecophagous species (anteaters in particular), the two investigated species (*Manis javanica* and *Manis pentadactyla*) possess only one functional *CHIA* paralog (*CHIA5*). The presence of the sole *CHIA5* in

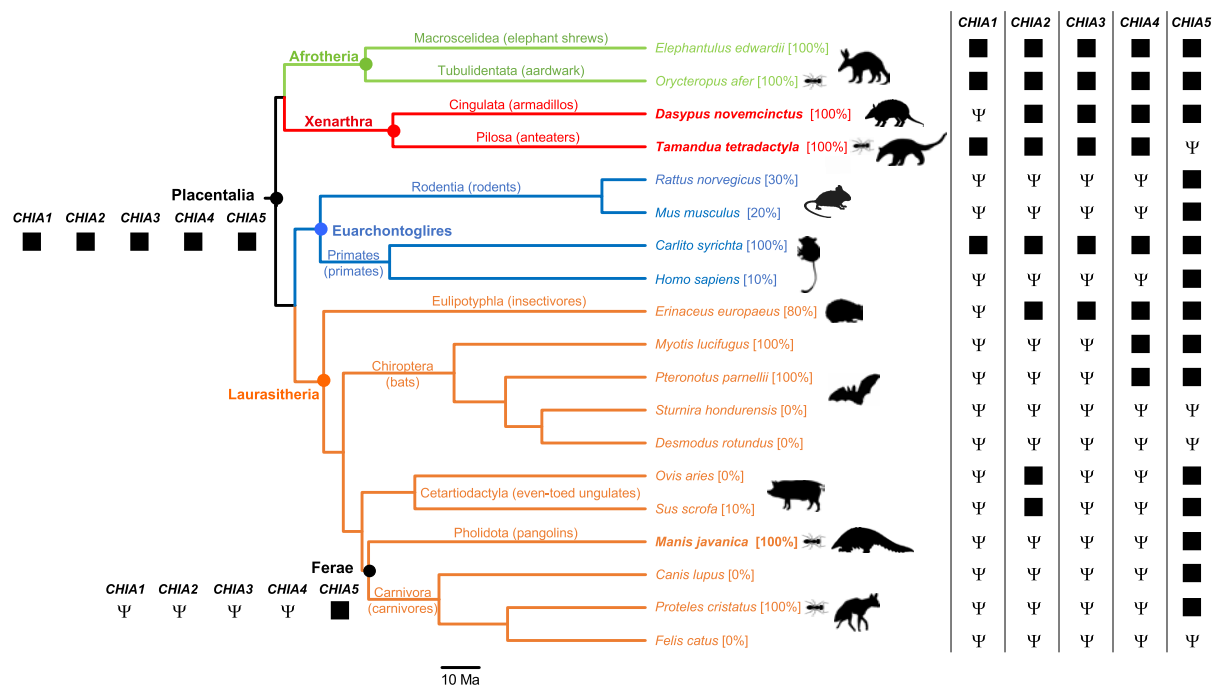


Fig. 1. Dated placental mammal phylogeny including representative species of the four major clades (Afrotheria, Xenarthra, Euarchontoglires, and Laurasiatheria) for which *CHIA* gene repertoires have been previously characterized. Numbers between brackets represent percentages of invertebrates in the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

pangolins was hypothesized to be the consequence of historical contingency on the evolution of the chitinase family with the probable loss of *CHIA1-4* functionality in the most recent common ancestor of Pholidota and Carnivora (Ferae; Emerling et al. 2018; Fig. 1). It has indeed recently been confirmed that a noninsect-based diet has caused structural and functional changes in the *CHIA* gene repertoire resulting in multiple losses of function in Carnivora with only few species including insects in their diet retaining a fully functional *CHIA5* gene (Tabata et al. 2022). These recent results, combined with the apparent importance of chitinase paralogs in insect digestion, have prompted questions regarding how pangolins succeed in digesting chitin with only one functional paralog.

One possible evolutionary solution to inheriting a depleted gene family resides in the modification of gene expression patterns in the remaining functional paralogs. Indeed, *CHIA5* was recently found to be highly expressed in the main digestive organs of the Malayan pangolin (Ma et al. 2017, 2019; Cheng et al. 2023) suggesting that pangolins might compensate for their reduced chitinase repertoire by an increased ubiquitous expression of their only remaining functional *CHIA5* paralog in multiple organs. While this result is very encouraging, it lacks a general

comparison with *CHIA* paralog expression in other mammals and more specifically with other myrmecophagous mammals that present more functional *CHIA* paralogs. If gene expression indeed plays a compensatory role, one can expect that *CHIA5* expression in pangolins would be comparatively higher and more ubiquitous among digestive organs than the expression of the other *CHIA* paralogs in convergent myrmecophagous species.

To further explore *CHIA* paralog expression in mammals and more particularly in convergent myrmecophagous species, we adopted a 3-fold approach. First, with the aim of identifying all functional paralogs and better understanding their function in chitin digestion, we reconstructed the first detailed evolutionary history of the chitinase-like gene family in mammals based on phylogenetic analyses of publicly available genomic and transcriptomic data. In a second step, we generated a large comparative data set of salivary gland transcriptomes encompassing 33 mammalian species from various lineages with diverse diets (herbivores, carnivores, frugivores, insectivores, and omnivores), enabling for the first time the comparison of *CHIA* expression across mammalian species. The objective here was to determine whether insectivores and myrmecophagous species indeed exhibit differential chitinase paralog expression in their

salivary glands compared to mammals with other diets. In a third step, we focused on two convergent myrmecophagous species (the southern tamandua and the Malayan pangolin) and an insectivorous species (the nine-banded armadillo) for which we were able to assemble and generate transcriptomes of several digestive and nondigestive tissues, to compare the use of their chitinase gene repertoire expression across different organs. The objective of this final step was to determine whether variations in genomic chitinase repertoires were associated with distinct expression patterns in digestive tissues or whether these patterns were independent of the functional gene repertoire. Overall, by leveraging species diversity on the one hand and organ diversity on the other, our results shed light on the molecular underpinnings of convergent evolution in ant-eating mammals by revealing that divergent paths of chitinase gene family evolution underlie dietary convergence between anteaters and pangolins.

Results

Mammalian Chitinase Gene Family Evolution

In order to gain further insights into the evolution and potential function of chitinase-related genes in mammalian genomes, we performed the first detailed phylogenetic reconstruction of the chitinase-like gene family based on functional paralogs using a gene tree/species tree reconciliation approach. The reconciled maximum likelihood tree of mammalian chitinase genes is presented in Fig. 2a. Our analyses showed that this gene family is constituted by nine paralogs whose evolution is notably characterized by gene loss 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/CHIBL1-2* and *CHIA3-5* groups of paralogs. Within the *CHIT1/CHIBL* clade, two consecutive duplications gave rise to *CHIT1* and then *CHIBL1* and *CHIBL2*. 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 recent and specific to placentals. This scenario of chitinase gene evolution is consistent with our new synteny analysis showing physical proximity of *CHIA1-2* and *OVGP1* on one hand and *CHIA3-5* on the other hand (Fig. 2b), which implies that chitinase genes evolved by successive tandem duplications.

Comparison of Ancestral Sequences

The ancestral amino acid sequences of the nine chitinase paralogs were reconstructed from the reconciled mammalian

gene tree and compared to gain further insight into the potential function of the enzymes they encode (Fig. 3; complete ancestral sequences and associated probabilities available from Zenodo). The alignment of predicted amino acid sequences locates the chitinolytic domain between positions 133 and 140 with the preserved pattern DXXDXDXE. The ancestral sequences of *CHIBL1* and *CHIBL2*, as all contemporary protein sequences of the corresponding genes, have a mutated chitinolytic domain with the absence of a glutamic acid at position 140 (Fig. 3a), which is the active proton-donor site necessary for chitin hydrolysis (Olland et al. 2009; Hamid et al. 2013). This indicates that the ability to degrade chitin has likely been lost before the duplication leading to *CHIBL1* and *CHIBL2* (Fig. 3b). The ancestral sequence of *OVGP1* also presents a mutated chitinolytic site although the glutamic acid at position 140 is present (Fig. 3a). The evolution of the different chitinases therefore seems to be related to changes in their active site. The six cysteine residues allowing the binding to chitin are found at positions 371, 418, 445, 455, 457, and 458 (Fig. 3c). The absence of one of these cysteines prevents binding to chitin (Tjoelker et al. 2000) as this is the case in the ancestral *OVGP1* protein where the last four cysteine residues are changed (Fig. 3c). The other ancestral sequences present the six conserved cysteine residues and thus can bind to chitin (Fig. 3c).

Chitinase Gene Expression in Mammalian Salivary Glands

To test the hypothesis that salivary glands play an important functional role in the digestion of ants and termites in ant-eating mammals, we analyzed the gene expression profiles of the nine chitinase paralogs revealed by the gene family tree reconstruction in 40 salivary gland transcriptomes representing 33 species (Fig. 4). *CHIA1* was expressed only in the elephant shrew (*Elephantulus myurus*; 23.22 normalized counts [NC]). *CHIA2* was expressed only in the wild boar (*Sus scrofa*; 48.84 NC). *CHIA3* was expressed in the two insectivorous California leaf-nosed bat individuals (*Macrotus californicus*; 367.70 and 35.03 NC) and in all three southern tamandua individuals (*T. tetradactyla*; 48.66, 41.52, and 15.14 NC). *CHIA4* was also highly expressed in all three southern tamanduas (565.61, 214.83, and 180.26 NC) and in the two California leaf-nosed bats (*M. californicus*; 17,224.06, and 16,880.24 NC), but also in the giant anteater (*Myrmecophaga tridactyla*; 50.74 NC). Expression of *CHIA5* was at least an order of magnitude higher in the two Malayan pangolin individuals (*M. javanica*; 196,778.69 and 729.18 NC) and Thomas's nectar bat (*Hsunycteris thomasi*; 7,301.82 NC) than in the three other species in which we detected expression of this gene: the domestic mouse (*Mus musculus*; 40.15 NC), common genet (*Genetta genetta*; 132.64 NC),

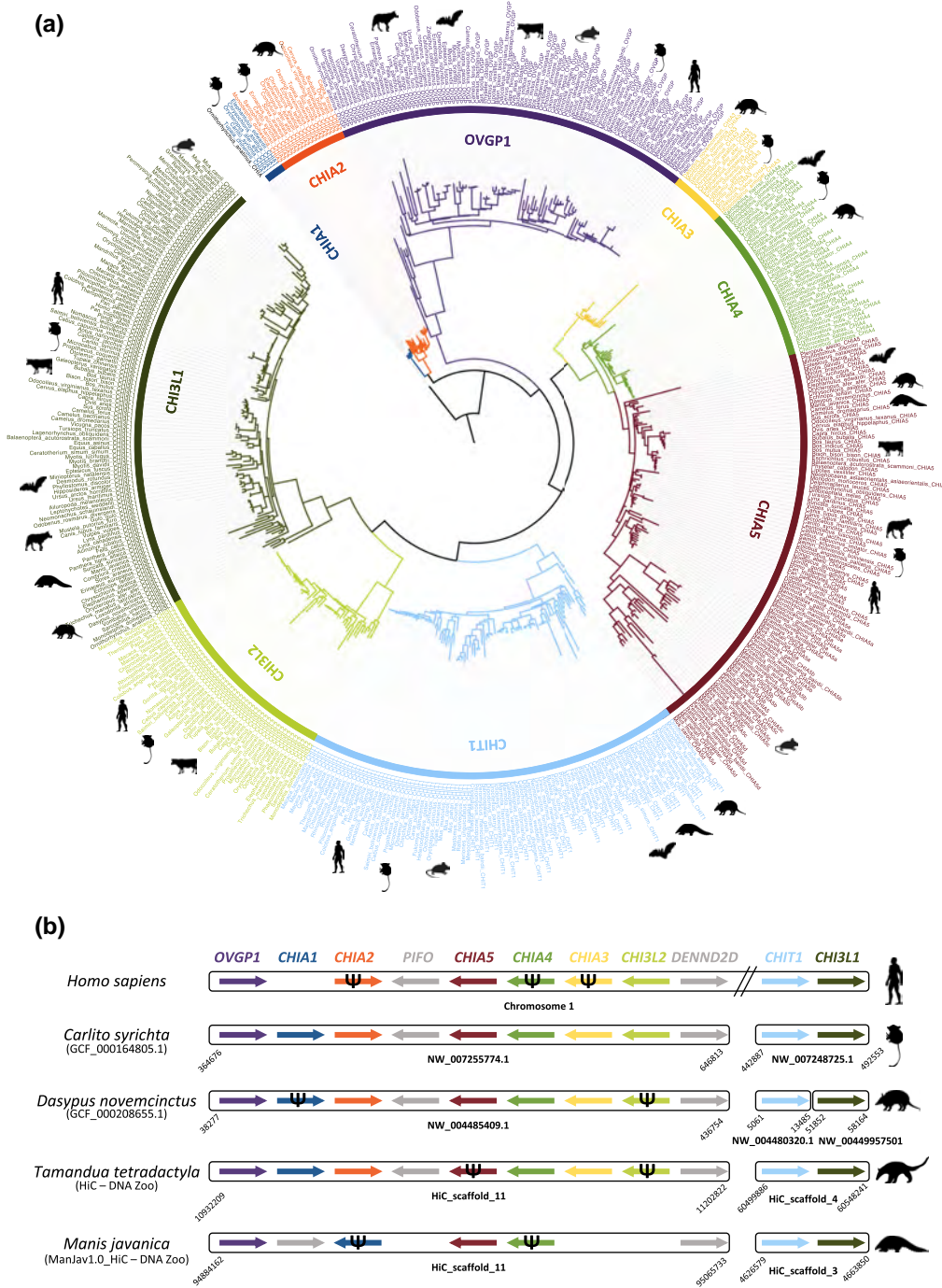


Fig. 2. a) Mammalian chitinase-like gene family tree reconstructed using a maximum likelihood gene-tree/species-tree reconciliation approach on protein sequences. The nine chitinase paralogs are indicated on the outer circle. Scale bar represents the mean number of amino acid substitutions per site. b) Syntenic analysis of the nine chitinase paralogs in humans (*H. sapiens*), tarsier (*C. syrichta*), nine-banded armadillo (*D. novemcinctus*), and the two main focal convergent ant-eating species: the southern tamandua (*T. tetradactyla*) and the Malayan pangolin (*M. javanica*). Assembly names and accession numbers are indicated below species names. Boxes represent different contigs with their most upstream and downstream BLAST hit positions to chitinase genes (colored arrows). Genes PIFO and DENND2D (gray arrows) are not chitinase paralogs but were used in the synteny analysis. 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 nonsignificant BLAST hits were not represented and are probably not functional or absent. Silhouettes were obtained from www.phylopic.org.

Downloaded from <https://academic.oup.com/gbe/article/17/2/evaf002/7945845> by guest on 22 February 2025

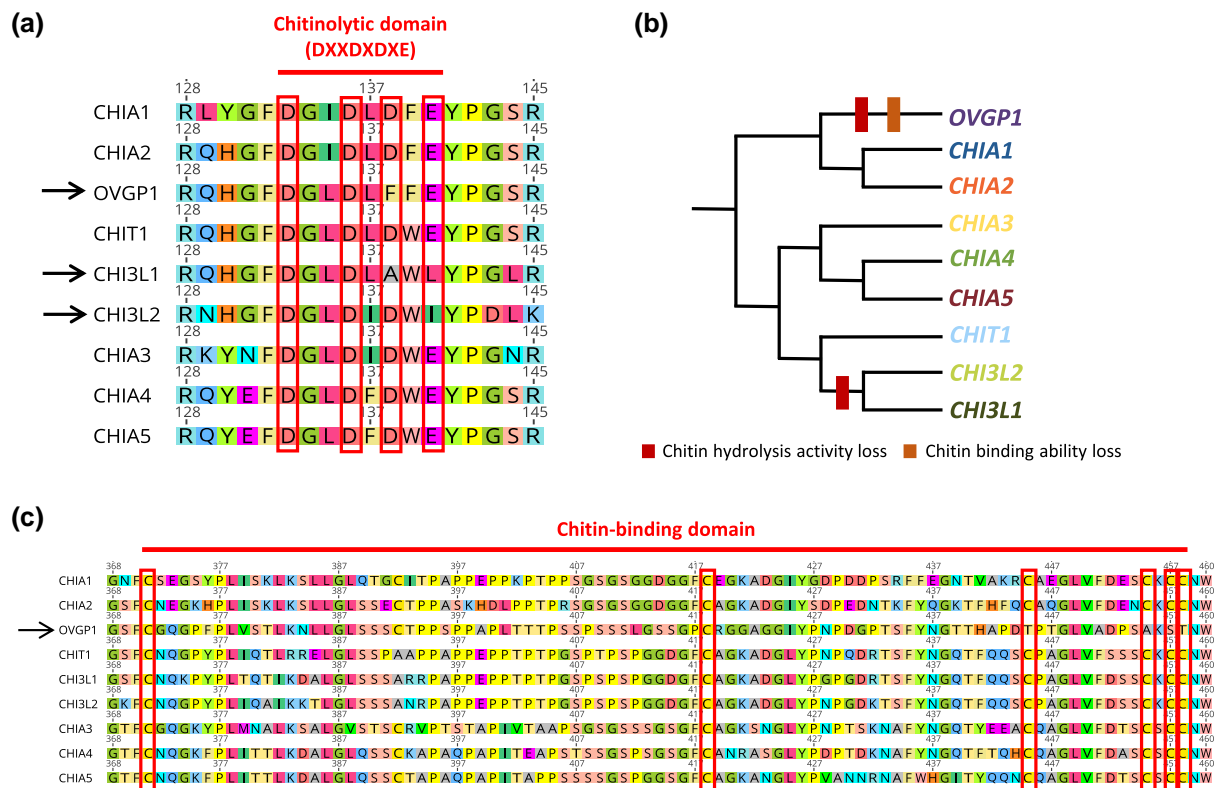


Fig. 3. Comparison of predicted ancestral protein sequences of the nine mammalian chitinase paralogs. a) Conserved amino acid residues of the canonical chitinolytic domain active site (DXXDXDXE). Arrows indicate paralogs in which changes occurred in the active site. b) Summary of the evolution of chitinase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

and wild boar (*S. scrofa*; 152.20 NC). *CHIT1* was expressed in many species (12 out of 40 samples) with values ranging from 46.76 NC in a single southern tamandua (*T. tetradactyla*) individual to 115,739.25 NC in the short-tailed shrew tenrec (*Microgale brevicaudata*). *CHI3L1* was expressed in most species (24 out of 40 samples) with values ranging from 61.68 NC in the giant anteater (*M. tridactyla*) to 1,297.01 NC in a Malayan pangolin (*M. javanica*) individual. *CHI3L2* was expressed in human (*Homo sapiens*; 1334.07 NC), wild boar (*S. scrofa*; 246.41 NC), elephant shrew (*E. myurus*; 94.65 NC), and common tenrec (*Tenrec ecaudatus*; 68.62 NC). *OVGP1* was only found expressed at very low levels in domestic dog (*Canis lupus familiaris*; 6.80 NC), human (*H. sapiens*; 15.33 NC), one of the two Malayan pangolins (*M. javanica*; 4.99 NC), and wild boar (*S. scrofa*; 17.84 NC). Finally, the southern aardwolf (*P. cristatus*), Norway rat (*Rattus norvegicus*), Parnell’s mustached bat (*Pteronotus parnellii*), and six phyllostomid bat species (*Carollia sowelli*, *Centurio senex*, *Glossophaga commissarisi*, *Sturnira hondurensis*, *Trachops cirrhosus*, and *Uroderma bilobatum*) did not appear to express any of the nine chitinase gene paralogs in any of our salivary gland samples.

Chitinase Gene Expression in Digestive and Nondigestive 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; *Dasypus novemcinctus*) and two of the main convergent myrmecophagous species (the southern anteater, *T. tetradactyla*, and the Malayan pangolin, *M. javanica*). This analysis revealed marked differences in expression level of these genes among the three species and among their digestive and nondigestive organs (Fig. 5). In the nine-banded armadillo (*D. novemcinctus*), although only *CHIA1* is pseudogenized and consequently 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 (51.90 NC). 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 (377,324.73 and 735,264.20 NC) and salivary glands (196,778.69 and 729.18 NC) and at milder levels in the tongue (121.24 NC), liver (254.79

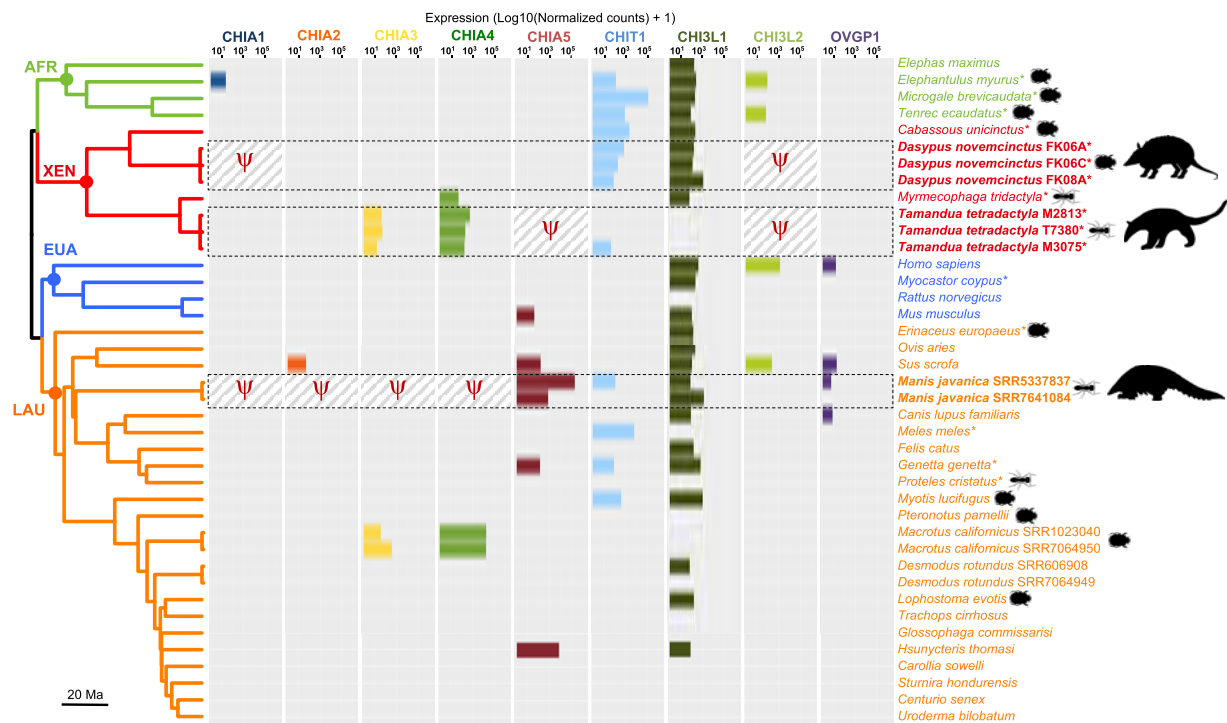


Fig. 4. Expression of the nine chitinase paralogs in 40 mammalian salivary gland transcriptomes. The 33 species are presented in their phylogenetic context covering the four major placental clades: Afrotheria (AFR), Xenarthra (XEN), Euarchontoglires (EUA), and Laurasiatheria (LAU). The chronogram was extracted from www.timetree.org (Kumar et al. 2022). Nonfunctional pseudogenes are only indicated for the three focal species (in bold) using a Ψ symbol: nine-banded armadillo (*D. novemcinctus*), southern tamandua (*T. tetradactyla*), and Malayan pangolin (*M. javanica*). Expression level is represented as log10 (normalized counts + 1). Asterisks indicate the 16 new transcripts produced in this study. Myrmecophagous and insectivorous species are indicated by ant and beetle silhouettes, respectively. Silhouettes were obtained from www.phylopic.org.

NC on average when expressed), pancreas (168.64 and 39.33 NC), large intestine (238.45 and 79.32 NC), and small intestine (847.51 and 13.72 NC), but also in skin (178.95 NC) and spleen (12.06 NC) samples. Conversely, in the southern tamandua (*T. tetradactyla*), only *CHIA5* is pseudogenized and accordingly not expressed (Fig. 5). *CHIA1* was found highly expressed in the pancreas (64,443.05 NC) and weakly expressed in testes (22.74 and 14.73 NC), and *CHIA2* also had very high expression in the pancreas (1,589,834.39 NC) and low expression in testes (36.51 and 34.52 NC) and lungs (8.22 NC). *CHIA3* was also expressed in the pancreas (359.03 NC), testes (241.79 and 35.42 NC), tongue (39.53 and 12.44 NC), salivary glands (48.66, 41.52, and 15.14 NC), and liver (32.40 NC). Finally, *CHIA4* was expressed in the testes (19.48 and 14.59 NC), spleen (109.97 and 73.31 NC), lungs (340.84 NC), salivary glands (565.61, 214.83, and 180.26 NC), and glandular stomach (116.11 NC). More globally, *CHIT1* was expressed in all tissues in *M. javanica*, in the testes, tongue, salivary glands, and small intestine in *T. tetradactyla*, and in the cerebellum, lungs, salivary glands, and liver in *D. novemcinctus*. *CHI3L1* was found to be expressed in the majority of digestive and nondigestive

tissues in all three species. *CHI3L2* is nonfunctional or even absent in the genome of these three species and was consequently not expressed. *OVGP1* was only weakly expressed in the lungs and salivary glands of *M. javanica* (2.22 and 4.99 NC, respectively).

Discussion

Evolution of Chitinase Paralogs Toward Different Functions

Chitinases have long been suggested to play an important role in insect digestion within mammals (Jeuniaux 1961, 1966, 1971; Jeuniaux and Cornelius 1978). After the initial discovery of a single chitinase gene (Boot et al. 2001), comparative genomics and phylogenetics have revealed a gene family (GH18) in which chitinases and chitinase-like proteins may work together to facilitate chitin digestion in the digestive tracts of mammals. The first phylogenetic analyses of this gene family have revealed a dynamic evolutionary history marked by gene duplication and loss following a typical birth-and-death model and a high degree of synteny among mammals (Bussink et al. 2007; Funkhouser and

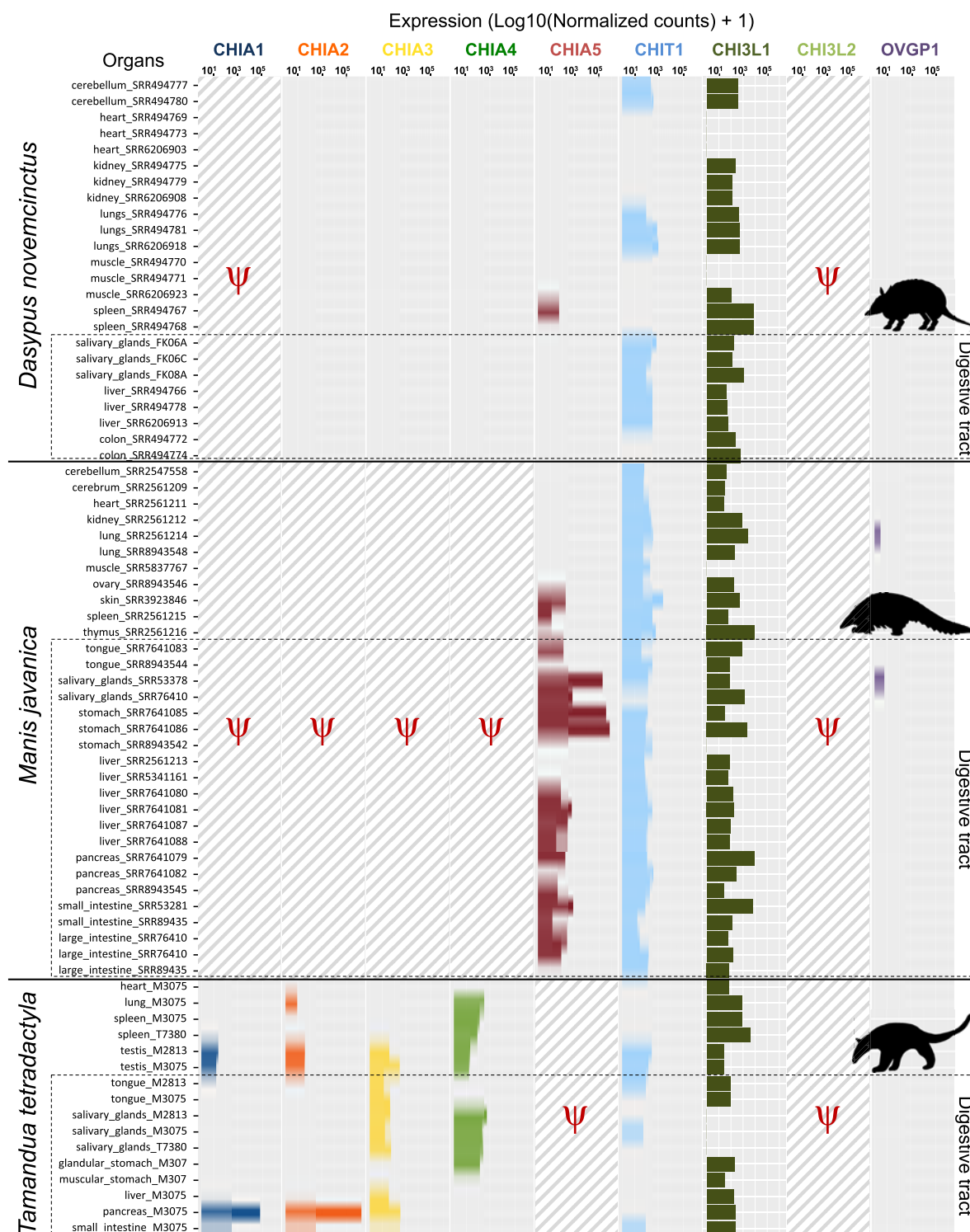


Fig. 5. Expression of the nine chitinase paralogs in 72 transcriptomes from different organs of the three focal species: the nine-banded armadillo (*D. novemcinctus*), the Malayan pangolin (*M. javanica*), and the southern tamandua (*T. tetradactyla*). Nonfunctional pseudogenes are represented by a Ψ symbol and hatched background. Boxes indicate organs of the digestive tract. Expression level is represented as log10 (normalized counts + 1). Silhouettes were obtained from www.phylopic.org.

Aronson 2007; Hussain and Wilson 2013). Our new comprehensive maximum likelihood phylogenetic analyses recovered nine functional paralogous chitinase gene sequences in mammalian genomes (Fig. 2a). In addition to the five previously characterized *CHIA* paralogs (Emerling et al. 2018; Janiak et al. 2018), we were able to include an additional gene (*OVGP1*), previously identified by Hussain and Wilson (2013), which is most closely related to the previously characterized *CHIA1* and *CHIA2* genes. In placentals, *OVGP1* plays a role in fertilization and embryonic development (Buhi 2002; Saint-Dizier et al. 2014; Algarra et al. 2016; Laheri et al. 2018). However, other aliases for *OVGP1* include *Mucin 9* and *CHIT5* suggesting 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 that also includes *CHIA3-5* and *CHI3L2* (Fig. 2b). Marsupial *CHIA4* sequences were located at the base of the *CHIA4-5* clade suggesting that this duplication might be recent and specific to placentals. The physical proximity of *CHIA3-5* on the same chromosomal fragment implies that these three *CHIA* genes evolved through successive tandem duplications. However, evidence of concerted evolution by gene conversion potentially misleading phylogenetic reconstruction between the two more recent duplicates (*CHIA4* and *CHIA5*), at least in some taxa (Emerling et al. 2018), suggests that further analyses are needed to fully unravel the origin and function of these two specific paralogs.

Comparison of the ancestral amino acid sequences of the nine chitinase paralogs revealed differences in their ability to bind and degrade chitin (Fig. 3), suggesting that these paralogs have evolved toward different functional specializations. The evolution of chitinase-like proteins was accompanied by a loss of enzymatic activity for chitin hydrolysis, which occurred several times independently (Bussink et al. 2007; Funkhouser and Aronson 2007; Hussain and Wilson 2013; Fig. 3b). *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 suggested to play a role in the inflammatory response (Lee et al. 2011) and are expressed in nondigestive tissues, in agreement with our comparative transcriptomic results. Thus, it has been proposed that the expansion of the chitinase gene family is related to the emergence of the innate and adaptive immune systems in vertebrates (Funkhouser and Aronson 2007).

The evolution of the different *CHIA1-5* genes has involved changes in their catalytic sites, which have consequences for 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). Experimental testing of the chitin degrading activity of ancestral reconstructions of each of the five *CHIA* enzymes, on different substrates and at different pH of enzymes, would help determine if there are differences in organ specificity of each enzyme. Furthermore, studying the potential molecular binding properties of these enzymes to other substrates would shed additional light on their functional roles. For example, changing 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 in 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. 2023), may help to decipher their respective roles in mammalian digestion (see below).

Impact of Historical Contingency and Molecular Tinkering on Chitinase Evolution and Expression

In the specific case of adaptation to myrmecophagy, comparative genomic and transcriptomic analyses of these chitinase genes, particularly those encoding chitinolytic enzymes (*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. 2023). On the 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 at least in some of its descendants (Fig. 6). In the southern tamandua (*T. tetradactyla*), the inactivating mutations of *CHIA5* were 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). In our study, this gene was not found to be expressed in the salivary glands of the giant anteater. On the other hand, *CHIA5* is functional in insectivorous carnivores (Carnivora) and pangolins (Pholidota), whereas *CHIA1-4* are pseudogenized (Emerling et al. 2018; Tabata et al. 2022). Similar inactivating mutations have been observed in the *CHIA1* gene in carnivores and pangolins and dated to at least 67 Ma, 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

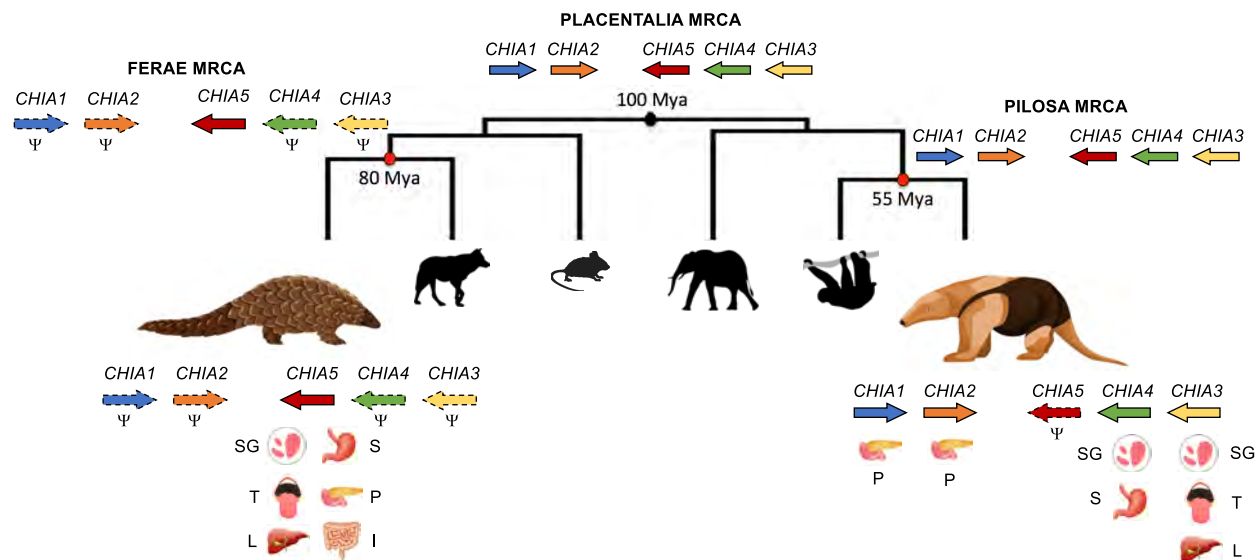


Fig. 6. Summary figure presenting the evolution and expression of chitinase acidic (*CHIA*) paralogous genes in the convergently evolved Malayan pangolin (*M. javanica*) and southern tamandua (*T. tetradactyla*) in their phylogenetic context. Reconstructed *CHIA* gene repertoires are indicated for the two myrmecophagous species and for the most recent common ancestor (MRCA) of placentals, pangolins + carnivores (Ferae) and anteaters + sloths (Pilosa). Nonfunctional pseudogenes are represented by the Ψ symbol and dashed line contour. Organ icons indicate expression of the corresponding gene in different digestive organs. SG, salivary glands; S, stomach; T, tongue; P, pancreas; L, liver; I, intestine. Silhouettes were obtained from www.phylopic.org and www.vecteezy.com.

functional *CHIA* gene (*CHIA5*), likely due to a historical contingency related to their common inheritance with carnivores (Fig. 6). These analyses have thus revealed contrasting pseudogenization events between convergent myrmecophagous species, with lesser anteaters (genus *Tamandua*) retaining functional orthologs for four out of the five chitin-degrading *CHIA* genes (*CHIA1-4*), while the Malayan pangolin (*M. javanica*) inherited only the fifth one (*CHIA5*) (Emerling et al. 2018). This peculiar evolutionary history raised the question whether the Malayan pangolin might compensate 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 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 spanning a diversity of species with different diets revealed that, among ant-eating mammals, the Malayan pangolin (*M. javanica*), the southern tamandua (*T. tetradactyla*), and the giant anteater (*M. tridactyla*) all express one or more chitin-degrading genes in their salivary glands. More specifically, we found that *CHIA1* and *CHIA2* were almost never expressed in mammalian salivary glands. By contrast, *CHIA4* was found to be expressed in the giant anteater (*M. tridactyla*) and expression of both *CHIA3* and *CHIA4* was observed in the three southern tamandua

(*T. tetradactyla*) individuals surveyed. Moreover, we were able to confirm the hypothesis implying an overexpression of the only functional *CHIA* gene possessed by the Malayan pangolin. Indeed, salivary gland expression profiles of *CHIA5* in *M. javanica* were much higher than in the four other species (Thomas’s nectar bat, mouse, genet, and wild boar) in which we detected expression of this gene and also substantially higher than the expression of any other chitin-degrading *CHIA* in the 32 other mammalian species considered. Finally, apart from anteaters, *CHIA3* and *CHIA4* were found to be highly expressed only in the two individuals of the insectivorous California leaf-nosed bat (*M. californicus*), but not in any of the other 11 examined bat species, including insectivorous species such as *Myotis myotis*, *P. parnellii*, and *Lophostoma evotis*. A possible explanation is that these genes have been pseudogenized in many of these bat species, which would be concordant with the findings of comparative genomic studies reporting widespread pseudogenizations of *CHIA* paralogs across multiple bat species (Emerling et al. 2018), with complete loss of *CHIA1-5* function in noninsectivorous old world fruit bats, most frugivorous bats, and the sanguivorous common vampire bat (Wang et al. 2020). However, although *CHIA4* and *CHIA5* appear to be functional in the insectivorous little brown myotis (*Myotis lucifugus*; Emerling et al. 2018; Wang et al. 2020), we did not observe expression of these genes in the salivary gland transcriptome we analyzed. Also, *CHIA5* was found to be highly

expressed in Thomas's nectar bat (*H. thomasi*). Although this bat species feeds mostly on nectar and fruits, its diet also includes a substantial part of insects suggesting that *CHIA5* might play a role in chitin digestion in the oral cavity, as a result of salivary gland secretion. Transcriptomic analyses of additional digestive tissues besides salivary glands in bats (Vandeweghe 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). Overall, our chitinase gene expression results therefore support a primary role for salivary glands in prey digestion through the use of distinct *CHIA* paralogs (*CHIA3*, *CHIA4*, and *CHIA5*) in different insect-eating placental mammal species.

Our differential expression comparison of the distinct chitinase paralogs across different organs further highlights the importance of *CHIA5* for Malayan pangolin digestive physiology by confirming its ubiquitous expression in all major tissues of the digestive tract (tongue, salivary glands, stomach, pancreas, liver, and large and small intestines) (Ma et al. 2017, 2019; Cheng et al. 2023; Fig. 6). More specifically, *CHIA5* was found to be 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 of the stomach (Ma et al. 2018; Cheng et al. 2023), 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. 2023) and all functional *CHIAs* were found expressed in its digestive tract but not in the same tissues (Fig. 6). *CHIA1* and *CHIA2* were particularly highly expressed in the pancreas whereas *CHIA3* and *CHIA4* were expressed across several other organs of the digestive tract including tongue, salivary glands, stomach, and liver. *CHIA1-4* were also expressed in other nondigestive organs (testes, lungs, and spleen), but their coexpression in the salivary glands of the three southern tamandua individuals sampled here strongly suggests that they play a crucial role in chitin digestion in this myrmecophagous species. Conversely, in the less specialized 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 diverse tissues of the individuals studied here, including salivary glands, and *CHIA5* was found only weakly expressed in one spleen sample. 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). We could not confirm this result, given that the liver and colon were the only additional digestive organs besides salivary glands

represented in our data set for this species. However, the comparison with the two myrmecophagous species seems to fit well with its less specialized insectivorous diet and actually further underscores 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 in part by using paralogs of different chitinase genes to digest chitin (Fig. 6), probably due to phylogenetic constraints leading to the loss of *CHIA1*, *CHIA2*, *CHIA3*, and *CHIA4* in the most recent common ancestor of Ferae (Carnivora and Pholidota; 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 with 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 in convergently evolved anteaters and pangolins. Working from different starting materials (i.e. different *CHIA* paralogs), natural selection led pangolins and anteaters to follow different paths in their convergent adaptation to the myrmecophagous diet.

Insights from Paralogous Gene Expression in Comparative Transcriptomic Studies

Conducting comparative transcriptomic studies between phylogenetically distant species is challenging (Dunn et al. 2013; Roux et al. 2015). Confounding factors include, but are not limited to, inconsistencies in tissue sampling methodology, timing of sample collection, and differences in tissue preservation methods. Any of these factors could significantly affect the quality of transcriptome assembly and quantification of gene expression profiles, which could affect underlying expression patterns related to phylogenetic or dietary similarities between species. Recognizing these challenges, we applied state-of-the-art assembly and annotation methods and performed rigorous normalization of read counts to mitigate some of the interspecific variability. Moreover, despite the interest in looking at overall expression patterns to identify the main effect associated with gene expression variation, exploratory comparative transcriptomic analyses also have some limitations. Indeed, when comparing the overall gene expression

pattern of different species, the first step is to identify comparable elements. These comparable elements can be restricted to single-copy orthologs or extended to homologous gene families containing different paralogs. However, some biases may be introduced during this step (see Li et al. 2023 for a review). On the one hand, focusing only on orthologous genes completely neglects the effects of paralogous gene expression. On the other hand, working at the scale of large homologous families (orthogroups) often leads to summarizing the expression of multiple orthologous genes into a single expression value. In our case, for example, following the orthogroup detection and summarizing the expression for each orthogroup would have led to a single expression value for the entire chitinase gene family (found as a single orthogroup). By contrast, thanks to our detailed investigation of the evolution of this gene family, phylogenetic and expression analyses of the chitinase orthogroups revealed interesting patterns that would have been missed by the global approach (i.e. effect of contingency bypassed by the relative expression of chitinase family genes). In particular, this approach highlighted differences in gene expression between closely related paralogs (i.e. *CHIAs*) in the digestive organs of the southern tamandua and the Malayan pangolin, which was crucial for our understanding of the molecular mechanisms involved in this case of convergent dietary adaptation. This result underscores the importance of using both genome- and transcriptome-wide analyses to identify novel candidate genes influencing specific traits and more targeted approaches based on existing knowledge. The latter is essential to deepen our understanding of the underlying mechanisms observed in specific cases, such as those of convergent evolution linked to historical contingency, as explored in this study.

Materials and Methods

Chitinase Gene Family Tree Reconstruction

Reconstruction of Chitinase Gene Family Evolution

Mammalian sequences similar to the protein sequence of the human *CHIA* chitinase acidic gene (NP_970615.2) were searched in the National Center for Biotechnology Information (NCBI) nonredundant protein database using BLASTP ($E < 10$). The protein sequences identified by BLASTP ($n = 1,476$) were then aligned using MAFFT v7.450 (Katoh and Standley 2013) with the following parameters (–auto –op 1.53 –ep 0.123 –aamatrix BLOSUM62). 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). From the reconstructed tree, the sequences were filtered according to the following criteria: (i) fast-evolving sequences with a BLAST E -value greater than 0 and not belonging to the

chitinase family were excluded; (ii) in cases of multiple isoforms, only the longest was retained; (iii) sequences whose length represented less than at least 50% of the total alignment length were removed; (iv) in case of identical sequences of different lengths from the same species, the longest was kept; and (v) sequences labeled as “hypothetical protein” and “predicted: low quality protein” were discarded. This procedure resulted in a data set containing 528 mammalian sequences that were realigned using MAFFT with the following parameters (–auto –op 1.53 –ep 0.123 –aamatrix BLOSUM62). This alignment contained 581 amino acid positions and 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 data set 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 with the following parameters (–auto –op 1.53 –ep 0.123 –aamatrix BLOSUM62), the phylogeny was inferred with RAxML under the GTR + G4 model, 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 files 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 with MAFFT with the following parameters (–auto –op 1.53 –ep 0.123 –aamatrix BLOSUM62) (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 (*H. 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 were downloaded from the 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 no significant BLAST hits 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 16 individuals representing 12 placental mammal species (supplementary table S1, Supplementary Material online). Total RNA was extracted from individual salivary gland tissue samples using the RNeasy extraction kit (Qiagen, Germany). Then, RNA-seq library construction and Illumina sequencing on a HiSeq 2500 system using paired-end 2 × 125 bp reads were conducted by the Montpellier GenomiX platform (MGX) resulting in 16 newly produced salivary gland transcriptomes. This sampling was completed with the 26 mammalian salivary gland transcriptomes available as paired-end Illumina sequencing reads in the Short Read Archive (SRA) of the NCBI as of December 15, 2022 representing an additional 21 species (supplementary table S1, Supplementary Material online). This taxon sampling includes representatives from all major mammal superorders Afrotheria ($n = 4$), Xenarthra ($n = 4$), Euarchontoglires ($n = 4$), and Laurasiatheria ($n = 21$) and covers six different diet categories: carnivory ($n = 4$), frugivory and herbivory ($n = 8$), insectivory ($n = 9$), myrmecophagy ($n = 5$), and omnivory ($n = 7$) (supplementary table S1, Supplementary Material online). Four of the five lineages in which myrmecophagous mammals evolved are represented: southern aardwolf (*P. cristatus*, Carnivora),

Malayan pangolin (*M. javanica*, Pholidota), southern naked-tailed armadillo (*Cabassous unicinctus*, Cingulata), giant anteater (*M. tridactyla*, Pilosa), and southern tamandua (*T. tetradactyla*, Pilosa). Species replicates in the form of different individuals were included for the southern tamandua (*T. tetradactyla*; $n = 3$), the nine-banded armadillo (*D. novemcinctus*; $n = 3$), the Malayan pangolin (*M. javanica*; $n = 2$), the vampire bat (*Desmodus rotundus*; $n = 2$), and the California leaf-nosed bat (*M. californicus*; $n = 2$). We unfortunately were not able to obtain fresh salivary gland samples from the aardvark (*O. afer*, Tubulidentata), the only missing myrmecophagous lineage in our sampling.

Transcriptomes from Additional Organs

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

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). Reads were allowed a minimum of 40% of bases with a PHRED score at least 15 ($-\text{qualified_quality_phred} \geq 15$), 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 cleaned paired-end reads ($-\text{seqType fq} -\text{left R1.fastq} -\text{right R2.fastq}$; result files available from Zenodo). For one individual vampire bat (*D. rotundus*), three salivary gland transcriptomes (SRR606902, SRR606908, and SRR606911) were combined to obtain a better assembly. For each of the 104 transcriptome assemblies, completeness was assessed by the presence of Benchmark Universal Single Copy Orthologs (BUSCO v5) based on a predefined data set (mammalia_odb10) of 9,226 single-copy orthologs conserved in over 90% of mammalian species (Manni et al. 2021). This pipeline was run through the gVolante web server (Nishimura et al. 2017) to evaluate the percentage of complete, duplicated, fragmented, and missing single-copy orthologs within each transcriptome (supplementary table S2, Supplementary Material online).

Transcriptome Annotation and Orthogroup Inference

The 104 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 FastTree (Price et al. 2010) for gene tree reconstructions. For expression analyses, orthogroups containing more than 20 copies for at least one species were discarded, resulting in the selection of 13,392 orthogroups for further analyses.

Gene Expression Analyzes

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 pseudoalignments 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 the reads aligned on the assembled transcriptome; Wolf 2013). Counts (raw number of mapped reads) and the transcripts per kilobase million are reported (result files available from Zenodo). Based on the previously inferred orthogroups, orthogroup-level abundance estimates were imported and summarized using tximport (Soneson et al. 2016). To minimize sequencing depth variation across samples and gene outlier effect (a few highly and differentially expressed genes may have strong and global influence on every gene read count), orthogroup-level raw reads counts were normalized using the median of the ratios of observed counts using DESeq2 (Love et al. 2014).

Chitinase Expression in Salivary Glands

The chitinase orthogroup was extracted from the orthogroups inferred by OrthoFinder2 using BLASTX with the reference chitinase database previously created. The 476 amino acid sequences composing this orthogroup were assigned to the nine chitinase orthologs (*CHIA1-5*, *CHIT1*, *CH13L1*, *CH13L2*, and *OVGP1*) using the maximum

likelihood Evolutionary Placement Algorithm implemented in RAxML-EPA (Berger et al. 2011) with the reference chitinase sequence alignment and reconciled phylogenetic tree previously inferred using GeneRax (result files available from Zenodo). This allowed excluding three additional contaminant sequences and dividing the chitinase orthogroup into nine suborthogroups corresponding to each chitinase paralog. To take advantage of the transcriptome-wide expression information for the expression standardization, these new orthogroups were included in the previous orthogroup-level abundance matrix estimates and the same normalization approach using DESeq2 was conducted. Finally, gene-level abundance estimates for all chitinase paralogs were extracted and compared on a log₁₀ scale.

Supplementary Material

Supplementary material online is available at *Genome Biology and Evolution* online.

Acknowledgments

We would like to thank Hugues Parrinello (Montpellier GenomiX platform) for advice on RNA-seq, Mariana Escobar Rodríguez and Gautier Debaecker for help with transcriptome assembly and annotation, and Marie Sémon for providing useful advice on RNA-seq 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. We also thank the four anonymous referees for their helpful comments. Computational analyses benefited from the Montpellier Bioinformatics Biodiversity (MBB) platform.

Funding

The JAGUARS collection is supported through grants from the Collectivité Territoriale de Guyane, from the European Union, and from Direction Générale des Territoires et de la Mer/Préfet de la Région Guyane attributed to Kwata NGO. 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 2024-313 of the Institut des Sciences de l'Evolution de Montpellier.

Data Availability

Raw RNA-seq Illumina reads have been submitted to the Short Read Archive (SRA) of the National Center for

Biotechnology Information (NCBI) and are available under BioProject number PRJNA909065. Transcriptome assemblies, phylogenetic data sets, corresponding trees, and other [supplementary materials](#) are available from [zenodo.org](#) (<https://doi.org/10.5281/zenodo.7355329>).

Literature Cited

- Algarra B, Han L, Soriano-Úbeda C, Avilés M, Coy P, Jovine L, Jiménez-Movilla M. The C-terminal region of OVGP1 remodels the zona pellucida and modifies fertility parameters. *Sci Rep*. 2016;6(1):32556. <https://doi.org/10.1038/srep32556>.
- Arendt J, Reznick D. Convergence and parallelism reconsidered: what have we learned about the genetics of adaptation? *Trends Ecol Evol*. 2008;23(1):26–32. <https://doi.org/10.1016/j.tree.2007.09.011>.
- Areshkov PO, Avdieiev SS, Balynska OV, Leroith D, Kavsan VM. 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*. 2011;8(1):39–48. <https://doi.org/10.7150/ijbs.8.39>.
- Berger SA, Krompass D, Stamatakis A. Performance, accuracy, and web server for evolutionary placement of short sequence reads under maximum likelihood. *Syst Biol*. 2011;60(3):291–302. <https://doi.org/10.1093/sysbio/syr010>.
- Blount ZD, Lenski RE, Losos JB. Contingency and determinism in evolution: replaying life's tape. *Science*. 2018;362(6415):eaam5979. <https://doi.org/10.1126/science.aam5979>.
- Boot RG, Blommaert EF, Swart E, Ghauharali-van der Vlugt K, Bijl N, Moe C, Place A, Aerts JM. Identification of a novel acidic mammalian chitinase distinct from chitotriosidase. *J Biol Chem*. 2001;276(9):6770–6778. <https://doi.org/10.1074/jbc.M009886200>.
- Bray NL, Pimentel H, Melsted P, Pachter L. Near-optimal probabilistic RNA-seq quantification. *Nat Biotechnol*. 2016;34(5):525–527. <https://doi.org/10.1038/nbt.3519>.
- Buhi WC. Characterization and biological roles of oviduct-specific, oestrogen-dependent glycoprotein. *Reproduction*. 2002;123(3):355–362. <https://doi.org/10.1530/rep.0.1230355>.
- Bussink AP, Speijer D, Aerts JM, Boot RG. Evolution of mammalian chitinase(-like) members of family 18 Glycosyl Hydrolases. *Genetics*. 2007;177(2):959–970. <https://doi.org/10.1534/genetics.107.075846>.
- Chen A-S, Taguchi T, Sakai K, Kikuchi K, Wang M-W, Miwa I. Antioxidant activities of chitobiose and chitotriose. *Biol Pharm Bull*. 2003;26(9):1326–1330. <https://doi.org/10.1248/bpb.26.1326>.
- Chen S, Zhou Y, Chen Y, Gu J. fastp: an ultra-fast all-in-one FASTQ pre-processor. *Bioinformatics*. 2018;34(17):i884–i890. <https://doi.org/10.1093/bioinformatics/bty560>.
- 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, Irwin DM, et al. Hologenomic insights into mammalian adaptations to myrmecophagy. *Natl Sci Rev*. 2023;10(4):nwac174. <https://doi.org/10.1093/nsr/nwac174>.
- Choo SW, Rayko M, Tan TK, Hari R, Komissarov A, Wee WY, Yurchenko AA, Kliver S, Tamazian G, Antunes A, et al. Pangolin genomes and the evolution of mammalian scales and immunity. *Genome Res*. 2016;26(10):1312–1322. <https://doi.org/10.1101/gr.203521.115>.
- Comte N, Morel B, Hasić D, Guéguen L, Boussau B, Daubin V, Penel S, Scornavacca C, Gouy M, Stamatakis A, et al. TreeRecs: an integrated phylogenetic tool, from sequences to reconciliations. *Bioinformatics*. 2020;36(18):4822–4824. <https://doi.org/10.1093/bioinformatics/btaa615>.
- Dudchenko O, Batra SS, Omer AD, Nyquist SK, Hoeger M, Durand NC, Shamim MS, Machol I, Lander ES, Aiden AP. De novo assembly of the *Aedes aegypti* genome using Hi-C yields chromosome-length scaffolds. *Science*. 2017;356(6333):92–95. <https://doi.org/10.1126/science.aal3327>.
- Dunn CW, Luo X, Wu Z. Phylogenetic analysis of gene expression. *Integr Comp Biol*. 2013;53(5):847–856. <https://doi.org/10.1093/icb/ict068>.
- Emerling CA, Delsuc F, Nachman MW. Chitinase genes (*CHIAs*) provide genomic footprints of a post-Cretaceous dietary radiation in placental mammals. *Sci Adv*. 2018;4(5):eaar6478. <https://doi.org/10.1126/sciadv.aar6478>.
- Emms DM, Kelly S. OrthoFinder: phylogenetic orthology inference for comparative genomics. *Genome Biol*. 2019;20(1):238. <https://doi.org/10.1186/s13059-019-1832-y>.
- Ferreira-Cardoso S, Delsuc F, Hautier L. Evolutionary tinkering of the mandibular canal linked to convergent regression of teeth in placental mammals. *Curr Biol*. 2019;29(3):468–475.e3. <https://doi.org/10.1016/j.cub.2018.12.023>.
- Ferreira-Cardoso S, Fabre P-H, de Thoisy B, Delsuc F, Hautier L. Comparative masticatory myology in anteaters and its implications for interpreting morphological convergence in myrmecophagous placentals. *PeerJ*. 2020;8:e9690. <https://doi.org/10.7717/peerj.9690>.
- Funkhouser JD, Aronson NN. Chitinase family GH18: evolutionary insights from the genomic history of a diverse protein family. *BMC Evol Biol*. 2007;7(1):96. <https://doi.org/10.1186/1471-2148-7-96>.
- Gallant JR, Traeger LL, Volkening JD, Moffett H, Chen PH, Novina CD, Phillips GN, Anand R, Wells GB, Pinch M. Genomic basis for the convergent evolution of electric organs. *Science*. 2014;344(6191):1522–1527. <https://doi.org/10.1126/science.1254432>.
- Gil F, Arencibia A, García V, Ramírez G, Vázquez JM. Anatomic and magnetic resonance imaging features of the salivary glands in the dog. *Anat Histol Embryol*. 2018;47(6):551–559. <https://doi.org/10.1111/ahe.12396>.
- Gordon-Thomson C, Kumari A, Tomkins L, Holford P, Djordjevic JT, Wright LC, Sorrell TC, Moore GPM. Chitotriosidase and gene therapy for fungal infections. *Cell Mol Life Sci*. 2009;66(6):1116–1125. <https://doi.org/10.1007/s00018-009-8765-7>.
- Gouy M, Guindon S, Gascuel O. SeaView version 4: a multiplatform graphical user interface for sequence alignment and phylogenetic tree building. *Mol Biol Evol*. 2010;27(2):221–224. <https://doi.org/10.1093/molbev/msp259>.
- Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I, Adiconis X, Fan L, Raychowdhury R, Zeng Q, et al. Trinity: reconstructing a full-length transcriptome without a genome from RNA-seq data. *Nat Biotechnol*. 2011;29(7):644–652. <https://doi.org/10.1038/nbt.1883>.
- Hamid R, Khan MA, Ahmad M, Ahmad MM, Abidin MZ, Musarrat J, Javed S. Chitinases: an update. *J Pharm Bioallied Sci*. 2013;5(1):21–29. <https://doi.org/10.4103/0975-7406.106559>.
- Hussain M, Wilson JB. New paralogues and revised time line in the expansion of the vertebrate GH18 family. *J Mol Evol*. 2013;76(4):240–260. <https://doi.org/10.1007/s00239-013-9553-4>.
- Jacob F. Evolution and tinkering. *Science*. 1977;196(4295):1161–1166. <https://doi.org/10.1126/science.860134>.
- Janiak MC, Chaney ME, Tosi AJ. Evolution of acidic mammalian chitinase genes (*CHIA*) is related to body mass and insectivory in primates. *Mol Biol Evol*. 2018;35(3):607–622. <https://doi.org/10.1093/molbev/msx312>.
- Jeuniaux C. Chitinase: an addition to the list of hydrolases in the digestive tract of vertebrates. *Nature*. 1961;192(4798):135–136. <https://doi.org/10.1038/192135a0>.

- Jeuniaux C. Chitinases. *Methods in enzymology*. Vol. 8. New York - London: Academic Press; 1966. p. 644–650.
- Jeuniaux C. On some biochemical aspects of regressive evolution in animals. In: Schoffeniels E, editor. *Biochemical evolution and the origin of life*. Amsterdam: North Holland Publishing Company; 1971. p. 304–313.
- Jeuniaux C, Cornelius C. Distribution and activity of chitinolytic enzymes in the digestive tract of birds and mammals. In: Muzzarelli RAA, Pariser ER, editors. *Proceedings of the first international conference on chitin/chitosan*. Boston: Massachusetts Institute of Technology; 1978. p. 542–549.
- Katoh K, Standley DM. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol Biol Evol*. 2013;30(4):772–780. <https://doi.org/10.1093/molbev/mst010>.
- Kozlov AM, Darriba D, Flouri T, Morel B, Stamatakis A. RAxML-NG: a fast, scalable and user-friendly tool for maximum likelihood phylogenetic inference. *Bioinformatics*. 2019;35(21):4453–4455. <https://doi.org/10.1093/bioinformatics/btz305>.
- Kumar S, Suleski M, Craig JM, Kasprowicz AE, Sanderford M, Li M, Stecher G, Hedges SB. TimeTree 5: an expanded resource for species divergence times. *Mol Biol Evol*. 2022;39(8):msac174. <https://doi.org/10.1093/molbev/msac174>.
- Laheri S, Ashary N, Bhatt P, Modi D. Oviductal glycoprotein 1 (OVGP1) is expressed by endometrial epithelium that regulates receptivity and trophoblast adhesion. *J Assist Reprod Genet*. 2018;35(8):1419–1429. <https://doi.org/10.1007/s10815-018-1231-4>.
- Le SQ, Gascuel O. An improved general amino acid replacement matrix. *Mol Biol Evol*. 2008;25(7):1307–1320. <https://doi.org/10.1093/molbev/msn067>.
- Lee CG, Da Silva CA, Dela Cruz CS, Ahangari F, Ma B, Kang M-J, He C-H, Takyar S, Elias JA. Role of chitin and chitinase/chitinase-like proteins in inflammation, tissue remodeling, and injury. *Annu Rev Physiol*. 2011;73(1):479–501. <https://doi.org/10.1146/annurev-physiol-012110-142250>.
- Li S, Nakayama H, Sinha NR. How to utilize comparative transcriptomics to dissect morphological diversity in plants. *Curr Opin Plant Biol*. 2023;76:102474. <https://doi.org/10.1016/j.cub.2023.102474>.
- Liu A, He F, Zhou J, Zou Y, Su Z, Gu X. Comparative transcriptome analyses reveal the role of conserved function in electric organ convergence across electric fishes. *Front Genet*. 2019;10:664. <https://doi.org/10.3389/fgene.2019.00664>.
- Losos JB. Convergence, adaptation, and constraint. *Evol Int J Org Evol*. 2011;65(7):1827–1840. <https://doi.org/10.1111/j.1558-5646.2011.01289.x>.
- Losos JB. *Improbable destinies: fate, chance, and the future of evolution*. New York: Riverhead Books; 2018.
- Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-Seq data with DESeq2. *Genome Biol*. 2014;15(12):550. <https://doi.org/10.1186/s13059-014-0550-8>.
- Ma J-E, Jiang H-Y, Li L-M, Zhang X-J, Li H-M, Li G-Y, Mo D-Y, Chen J-P. SMRT sequencing of the full-length transcriptome of the Sunda pangolin (*Manis javanica*). *Gene*. 2019;692:208–216. <https://doi.org/10.1016/j.gene.2019.01.008>.
- Ma J-E, Li L-M, Jiang H-Y, Zhang X-J, Li J, Li G-Y, Chen J-P. Acidic mammalian chitinase gene is highly expressed in the special oxyntic glands of *Manis javanica*. *FEBS Open Bio*. 2018;8(8):1247–1255. <https://doi.org/10.1002/2211-5463.12461>.
- 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. Transcriptomic analysis identifies genes and pathways related to myrmecophagy in the Malayan pangolin (*Manis javanica*). *PeerJ*. 2017;5:e4140. <https://doi.org/10.7717/peerj.4140>.
- Macmanes M. On the optimal trimming of high-throughput mRNA sequence data. *Front Genet*. 2014;5:13. <https://doi.org/10.3389/fgene.2014.00013>.
- Manni M, Berkeley MR, Seppey M, Simão FA, Zdobnov EM. BUSCO update: novel and streamlined workflows along with broader and deeper phylogenetic coverage for scoring of eukaryotic, prokaryotic, and viral genomes. *Mol Biol Evol*. 2021;38(10):4647–4654. <https://doi.org/10.1093/molbev/msab199>.
- McGhee GR. *Convergent evolution: limited forms most beautiful*. Boston: MIT Press; 2011.
- McGlothlin JW, Kobiela ME, Feldman CR, Castoe TA, Geffney SL, Hanifin CT, Toledo G, Vonk FJ, Richardson MK, Brodie ED, et al. Historical contingency in a multigene family facilitates adaptive evolution of toxin resistance. *Curr Biol*. 2016;26(12):1616–1621. <https://doi.org/10.1016/j.cub.2016.04.056>.
- Meredith RW, Janečka JE, Gatesy J, Ryder OA, Fisher CA, Teeling EC, Goodbla A, Eizirik E, Simão TLL, Stadler T, et al. Impacts of the Cretaceous terrestrial revolution and KPg extinction on mammal diversification. *Science*. 2011;334(6055):521–524. <https://doi.org/10.1126/science.1211028>.
- Morel B, Kozlov AM, Stamatakis A, Szöllösi GJ. GeneRax: a tool for species-tree-aware maximum likelihood-based gene family tree inference under gene duplication, transfer, and loss. *Mol Biol Evol*. 2020;37(9):2763–2774. <https://doi.org/10.1093/molbev/msaa141>.
- Nguyen NTT, Vincens P, Dufayard JF, Roest Crollius H, Louis A. Genomic in 2022: comparative tools for thousands of genomes and reconstructed ancestors. *Nucleic Acids Res*. 2022;50(D1):D1025–D1031. <https://doi.org/10.1093/nar/gkab1091>.
- Nishimura O, Hara Y, Kuraku S. Gvolante for standardizing completeness assessment of genome and transcriptome assemblies. *Bioinformatics*. 2017;33(22):3635–3637. <https://doi.org/10.1093/bioinformatics/btx445>.
- Olland AM, Strand J, Presman E, Czerwinski R, Joseph-McCarthy D, Krykbaev R, Schlingmann G, Chopra R, Lin L, Fleming M, et al. Triad of polar residues implicated in pH specificity of acidic mammalian chitinase. *Protein Sci*. 2009;18(3):569–578. <https://doi.org/10.1002/pro.63>.
- Pillai AS, Chandler SA, Liu Y, Signore AV, Cortez-Romero CR, Benesch JLP, Laganowsky A, Storz JF, Hochberg GKA, Thornton JW. Origin of complexity in haemoglobin evolution. *Nature*. 2020;581(7809):480–485. <https://doi.org/10.1038/s41586-020-2292-y>.
- Price MN, Dehal PS, Arkin AP. FastTree 2—approximately maximum-likelihood trees for large alignments. *PLoS One*. 2010;5(3):e9490. <https://doi.org/10.1371/journal.pone.0009490>.
- Ratnasingham S, Hebert PDN. Bold: the barcode of life data system (<http://www.barcodinglife.org>). *Mol Ecol Notes*. 2007;7(3):355–364. <https://doi.org/10.1111/j.1471-8286.2007.01678.x>.
- Recklies AD, White C, Ling H. The chitinase 3-like protein human cartilage glycoprotein 39 (HC-gp39) stimulates proliferation of human connective-tissue cells and activates both extracellular signal-regulated kinase- and protein kinase B-mediated signalling pathways. *Biochem J*. 2002;365(1):119–126. <https://doi.org/10.1042/bj20020075>.
- Redford KH. *Ants and termites as food*. In: Genoways HH, editor. *Current mammalogy*. Boston, MA: Springer US; 1987. p. 349–399.
- Reiss KZ. Using phylogenies to study convergence: the case of the ant-eating mammals. *Am Zool*. 2001;41:507–525. [https://doi.org/10.1668/0003-1569\(2001\)041\[0507:UPTSCT\]2.0.CO;2](https://doi.org/10.1668/0003-1569(2001)041[0507:UPTSCT]2.0.CO;2).
- Roux J, Rosikiewicz M, Robinson-Rechavi M. What to compare and how: comparative transcriptomics for Evo-Devo. *J Exp Zool Part B Mol Dev Evol*. 2015;324(4):372–382. <https://doi.org/10.1002/jez.b.22618>.
- Saint-Dizier M, Marnier C, Tahir MZ, Grimard B, Thoumire S, Chastant-Maillard S, Reynaud K. OVGP1 is expressed in the canine oviduct at the time and place of oocyte maturation and fertilization. *Mol Reprod Dev*. 2014;81(11):972–982. <https://doi.org/10.1002/mrd.22417>.

- Salgaonkar N, Prakash D, Nawani NN, Kapadnis BP. Comparative studies on ability of N-acetylated chitooligosaccharides to scavenge reactive oxygen species and protect DNA from oxidative damage. *Indian J Biotechnol.* 2015;14:186–192. <http://nopr.niscpr.res.in/handle/123456789/31805>.
- Shubin N, Tabin C, Carroll S. Deep homology and the origins of evolutionary novelty. *Nature.* 2009;457(7231):818–823. <https://doi.org/10.1038/nature07891>.
- Slater GSC, Birney E. Automated generation of heuristics for biological sequence comparison. *BMC Bioinformatics.* 2005;6(1):31. <https://doi.org/10.1186/1471-2105-6-31>.
- Smith SA, Robbins LW, Steiert JG. Isolation and characterization of a chitinase from the nine-banded armadillo, *Dasypus novemcinctus*. *J Mammal.* 1998;79(2):486–491. <https://doi.org/10.2307/1382979>.
- Soneson C, Love MI, Robinson MD. Differential analyses for RNA-seq: transcript-level estimates improve gene-level inferences. *F1000 Res.* 2016;4:1521. <https://doi.org/10.12688/f1000research.7563.2>.
- Stamatakis A. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics.* 2014;30(9):1312–1313. <https://doi.org/10.1093/bioinformatics/btu033>.
- Strobel S, Roswag A, Becker NI, Trenczek TE, Encarnação JA. Insectivorous bats digest chitin in the stomach using acidic mammalian chitinase. *PLoS One.* 2013;8(9):e72770. <https://doi.org/10.1371/journal.pone.0072770>.
- Tabata E, Itoigawa A, Koinuma T, Tayama H, Kashimura A, Sakaguchi M, Matoska V, Bauer PO, Oyama F. Noninsect-based diet leads to structural and functional changes of acidic chitinase in Carnivora. *Mol Biol Evol.* 2022;39(1):msab331. <https://doi.org/10.1093/molbev/msab331>.
- Tjoelker LW, Gosting L, Frey S, Hunter CL, Le Trong H, Steiner B, Brammer H, Gray PW. Structural and functional definition of the human chitinase chitin-binding domain. *J Biol Chem.* 2000;275(1):514–520. <https://doi.org/10.1074/jbc.275.1.514>.
- Vandeweghe MW, Sotero-Caio CG, Phillips CD. Positive selection and gene expression analyses from salivary glands reveal discrete adaptations within the ecologically diverse bat family Phyllostomidae. *Genome Biol Evol.* 2020;12(8):1419–1428. <https://doi.org/10.1093/gbe/evaa151>.
- Wang K, Tian S, Galindo-González J, Dávalos LM, Zhang Y, Zhao H. Molecular adaptation and convergent evolution of frugivory in Old World and neotropical fruit bats. *Mol Ecol.* 2020;29(22):4366–4381. <https://doi.org/10.1111/mec.15542>.
- Wang Y, Yang L. Genomic evidence for convergent molecular adaptation in electric fishes. *Genome Biol Evol.* 2021;13(3):evab038. <https://doi.org/10.1093/gbe/evab038>.
- Wolf JBW. Principles of transcriptome analysis and gene expression quantification: an RNA-seq tutorial. *Mol Ecol Resour.* 2013;13(4):559–572. <https://doi.org/10.1111/1755-0998.12109>.
- Xie VC, Pu J, Metzger BP, Thornton JW, Dickinson BC. Contingency and chance erase necessity in the experimental evolution of ancestral proteins. *eLife.* 2021;10:e67336. <https://doi.org/10.7554/eLife.67336>.
- Yates AD, Achuthan P, Akanni W, Allen J, Allen J, Alvarez-Jarreta J, Amode MR, Armean IM, Azov AG, Bennett R, et al. Ensembl 2020. *Nucleic Acids Res.* 2020;48(D1):D682–D688. <https://doi.org/10.1093/nar/gkz966>.
- Yusoff AM, Tan TK, Hari R, Koepfli K-P, Wee WY, Antunes A, Sitam FT, Rovie-Ryan JJ, Karuppannan KV, Wong GJ. De novo sequencing, assembly and analysis of eight different transcriptomes from the Malayan pangolin. *Sci Rep.* 2016;6(1):1–11. <https://doi.org/10.1038/s41598-016-0001-8>.
- Zakon HH, Lu Y, Zwickl DJ, Hillis DM. Sodium channel genes and the evolution of diversity in communication signals of electric fishes: convergent molecular evolution. *Proc Natl Acad Sci U S A.* 2006;103(10):3675–3680. <https://doi.org/10.1073/pnas.0600160103>.
- Zhang F, Xu N, Yu Y, Wu S, Li S, Wang W. Expression profile of the digestive enzymes of *Manis javanica* reveals its adaptation to diet specialization. *ACS Omega.* 2019;4(22):19925–19933. <https://doi.org/10.1021/acsomega.9b02845>.

Associate editor: Carina Mugal