EVOLUTIONARY BIOLOGY

Chitinase genes (*CHIA*s) provide genomic footprints of a post-Cretaceous dietary radiation in placental mammals

Christopher A. Emerling,[1,2]* Frédéric Delsuc,[2] Michael W. Nachman[1]

The end-Cretaceous extinction led to a massive faunal turnover, with placental mammals radiating in the wake of nonavian dinosaurs. Fossils indicate that Cretaceous stem placentals were generally insectivorous, whereas their earliest Cenozoic descendants occupied a variety of dietary niches. It is hypothesized that this dietary radiation resulted from the opening of niche space, following the extinction of dinosaurian carnivores and herbivores. We provide the first genomic evidence for the occurrence and timing of this dietary radiation in placental mammals. By comparing the genomes of 107 placental mammals, we robustly infer that chitinase genes (*CHIA*s), encoding enzymes capable of digesting insect exoskeletal chitin, were present as five functional copies in the ancestor of all placental mammals, and the number of functional *CHIA*s in the genomes of extant species positively correlates with the percentage of invertebrates in their diets. The diverse repertoire of *CHIA*s in early placental mammals corroborates fossil evidence of insectivory in Cretaceous eutherians, with descendant lineages repeatedly losing *CHIA*s beginning at the Cretaceous/Paleogene (K/Pg) boundary as they radiated into noninsectivorous niches. Furthermore, the timing of gene loss suggests that interordinal diversification of placental mammals in the Cretaceous predates the dietary radiation in the early Cenozoic, helping to reconcile a long-standing debate between molecular timetrees and the fossil record. Our results demonstrate that placental mammal genomes, including humans, retain a molecular record of the post-K/Pg placental adaptive radiation in the form of numerous chitinase pseudogenes.

INTRODUCTION

The traditional model of mammalian evolution posits that after originating in the Triassic-Jurassic, Mesozoic mammals were typically small, nocturnal, generalized terrestrial insectivores, limited in taxonomic and ecological diversity due to predation by and/or competition with dinosaurs. After the extinction of nonavian dinosaurs at the Cretaceous/Paleogene (K/Pg) boundary, Cenozoic mammals then radiated into an enormous range of ecologically distinct niches (*1*, *2*). Recent research suggests that this model is incomplete, with some fossil evidence for Mesozoic diversification and niche expansion in mammals (*3*–*6*). However, most Mesozoic eutherians (placental mammals + stem taxa) were small, likely terrestrial-scansorial, and had jaw and dental morphologies consistent with insectivory (*1*, *4*, *7*). Furthermore, although there is some evidence of eutherian niche diversification in the Late Cretaceous (*1*, *8*, *9*), the fossil record unequivocally points to an unprecedented adaptive radiation of placental mammals after the K/Pg boundary (*1*, *8*–*11*).

Here, we test the hypothesis that this landmark evolutionary radiation left a molecular signature in placental mammal genomes, specifically in regard to diet. Jeuniaux (*12*) hypothesized that mammals inherited gastrointestinal (GI) chitinases from a vertebrate ancestor and subsequently lost them as they adapted to noninsectivorous diets. Molecular studies have isolated an acidic mammalian chitinase (CHIA) expressed in the mammalian GI tract (*13*, *14*) that is stable in the presence of GI proteases (*15*), suggesting a role in chitin digestion. In the absence of dietary chitin, selection on chitinase genes may have been relaxed, allowing for the accumulation of formerly deleterious mutations (for example, frameshift and nonsense mutations), resulting in unitary pseudogenes and gene deletions. We tested Jeuniaux's (*12*) hypothesis by analyzing patterns of acidic mammalian chitinase gene (*CHIA*) loss in placental mammals relative to the K/Pg boundary. We provide evidence that *CHIA*s retain a genomic signal of the early Cenozoic radiation of placental mammals, via the repeated, convergent decay of *CHIA*s in noninsectivorous mammals after the K/Pg boundary.

[1]Museum of Vertebrate Zoology, University of California, Berkeley, CA 94720, USA. [2]Institut des Sciences de l'Évolution de Montpellier (ISEM), Université de Montpellier, CNRS, IRD, EPHE, Montpellier, France.
*Corresponding author. Email: christopher.emerling@umontpellier.fr

RESULTS AND DISCUSSION

Five *CHIA* paralogs in the last common ancestor of placental mammals

The original study describing *CHIA* reported only a single gene (*13*), but we found numerous placental *CHIA* paralogs, similar to recent analyses (*16*, *17*). Phylogenetic analyses cluster these into five major clades (Fig. 1; figs. S1 to S3; and table S1), which we refer to as *CHIA1*, *CHIA2*, *CHIA3*, *CHIA4*, and *CHIA5*. All four superorders of placental mammals (Xenarthra, Afrotheria, Euarchontoglires, and Laurasiatheria) are represented by orthologs in all five *CHIA* clades, implying that the last common ancestor of placental mammals had five *CHIA*s. There is a consistent syntenic orientation of the paralogs, with multiple instances of all five being localized to the same genomic contig/scaffold (fig. S4 and table S2). Each putatively functional *CHIA* paralog is of nearly identical length, consists of 11 coding exons, and has canonical splice donor and acceptor sites, a catalytic domain in exon 5 (fig. S5) (*18*), and a chitin-binding domain in exon 11 (fig. S6) (*19*), suggesting that all five *CHIA*s encode enzymes with chitinolytic function. Furthermore, transcriptome analyses indicate that all five *CHIA*s can be expressed in placental mammal GI tracts (table S3) and, therefore, each potentially participates in the breakdown of ingested chitin.

CHIA and loss dietary shifts to carnivory and herbivory

Although we inferred gene duplications early in mammalian history (fig. S1), the only duplications in crown Placentalia resulted in paralogs likely to be nonfunctional or lacking chitinolytic function (tables S1 and S4 and figs. S2 and S3), with at least some of these duplicates appearing to have neofunctionalized into genes with immune-related functions (fig. S3) (*20*). By contrast, we found that *CHIA* loss, via pseudogenization

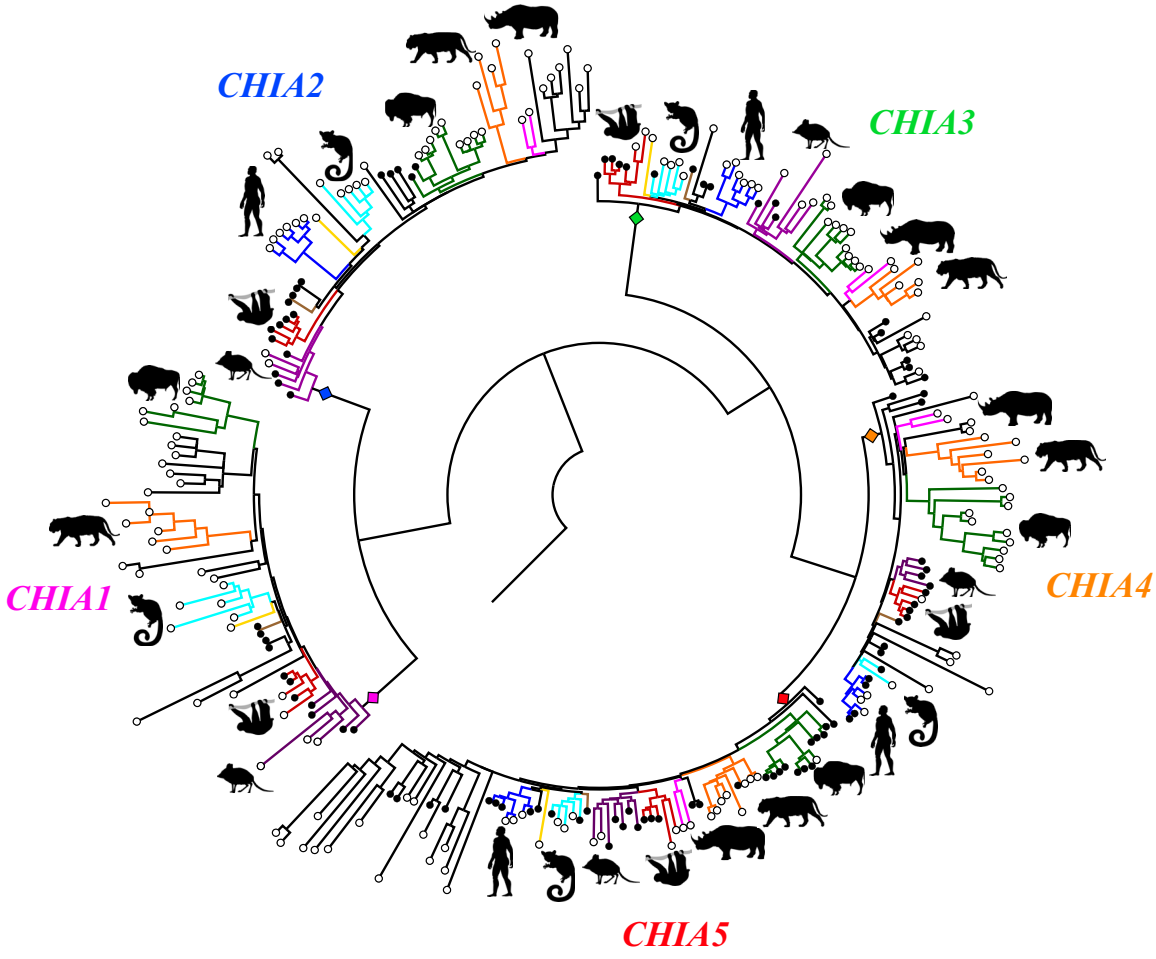

Fig. 1. Placental mammal *CHIA* gene tree, simplified from fig. S3. Closed circles indicate functional *CHIA*s, and open circles indicate pseudogenic *CHIA*s and/or *CHIA*s lacking a chitinolytic and/or chitin-binding domain. Colored branches correspond to a subset of placental mammal clades: green, Cetartiodactyla; orange, Carnivora; pink, Perissodactyla; red, Xenarthra; purple, Afrotheria; brown, Scandentia; cyan, Strepsirrhini; blue, Anthropoidea; yellow, Dermoptera. Silhouettes and licenses here and throughout are from phylopic.org.

and whole-gene deletion, has occurred very frequently, particularly in carnivorous and herbivorous lineages (table S1). Strikingly, species with five to three *CHIA* copies have, on average, a diet that consists almost entirely of invertebrates (five *CHIA*s, 88.3% diet invertebrates on average; four *CHIA*s, 85%; three *CHIA*s, 80%), species with two copies on average consume a moderate amount of invertebrates (63.3%), and species with one or no *CHIA*s average minimal invertebrate consumption (one *CHIA*, 21.9%; no *CHIA*s, 7%). Phylogenetic generalized least-squares (PGLS) regression analyses using *CHIA*s derived from both gene models ($n = 73$, $r^2 = 0.3358$, $P = 7.824 \times 10^{-8}$) and genomic contigs ($n = 70$, $r^2 = 0.4213$, $P = 1.219 \times 10^{-9}$) support a positive relationship between the percentage of invertebrates in the diet and the number of functional *CHIA*s in the genome (Fig. 2), a finding corroborated by a recent *CHIA* analysis focused on Primates (*17*). The association of high numbers of functional *CHIA*s with insectivory, and the inference of five *CHIA*s in the last common ancestor of placental mammals, suggests that this ancestor and most crown Cretaceous placental lineages were probably also insectivorous. If correct, then such a diet would likely have imposed constraints on body size and other life history characters in these lineages and should therefore help infer traits in the earliest placental mammals (*21*, *22*).

Whereas the earliest placental mammals had five *CHIA*s, inactivating mutations and gene deletions shared between multiple species point to losses of functional *CHIA*s during the origins of modern placental clades (Fig. 3 and table S4). Whereas some *CHIA*s were lost in individual species or families, other losses occurred before interordinal (Paenungulata and Ostentoria) and intraordinal (Chiroptera, Carnivora, Perissodactyla, Cetartiodactyla, and Lagomorpha) diversification. Significantly, all or nearly all *CHIA* genes were lost during the origin and diversification of noninsectivorous lineages, including the herbivorous sloths (Folivora), hyraxes, elephants, and sirenians (Paenungulata), Old World fruit bats (Pteropodidae), horses, rhinoceroses, and tapirs (Perissodactyla), camels, swine, and ruminants (Cetartiodactyla), colugos (Dermoptera), lemurs (Lemuriformes), monkeys and apes (Anthropoidea), rabbits and pikas (Lagomorpha), rodents (Rodentia), and the largely carnivorous false vampire bats (Megadermatidae), whales (Cetacea), and dogs, cats, and kin (Carnivora) (Fig. 1 and tables S1, S4, and S5).

CHIA loss and the post-K/Pg placental mammal radiation

The presence of *CHIA* pseudogenes in various species allowed us to more precisely date the timing of gene inactivations relative to the K/Pg boundary. We used a method (*23*) based on the accumulation

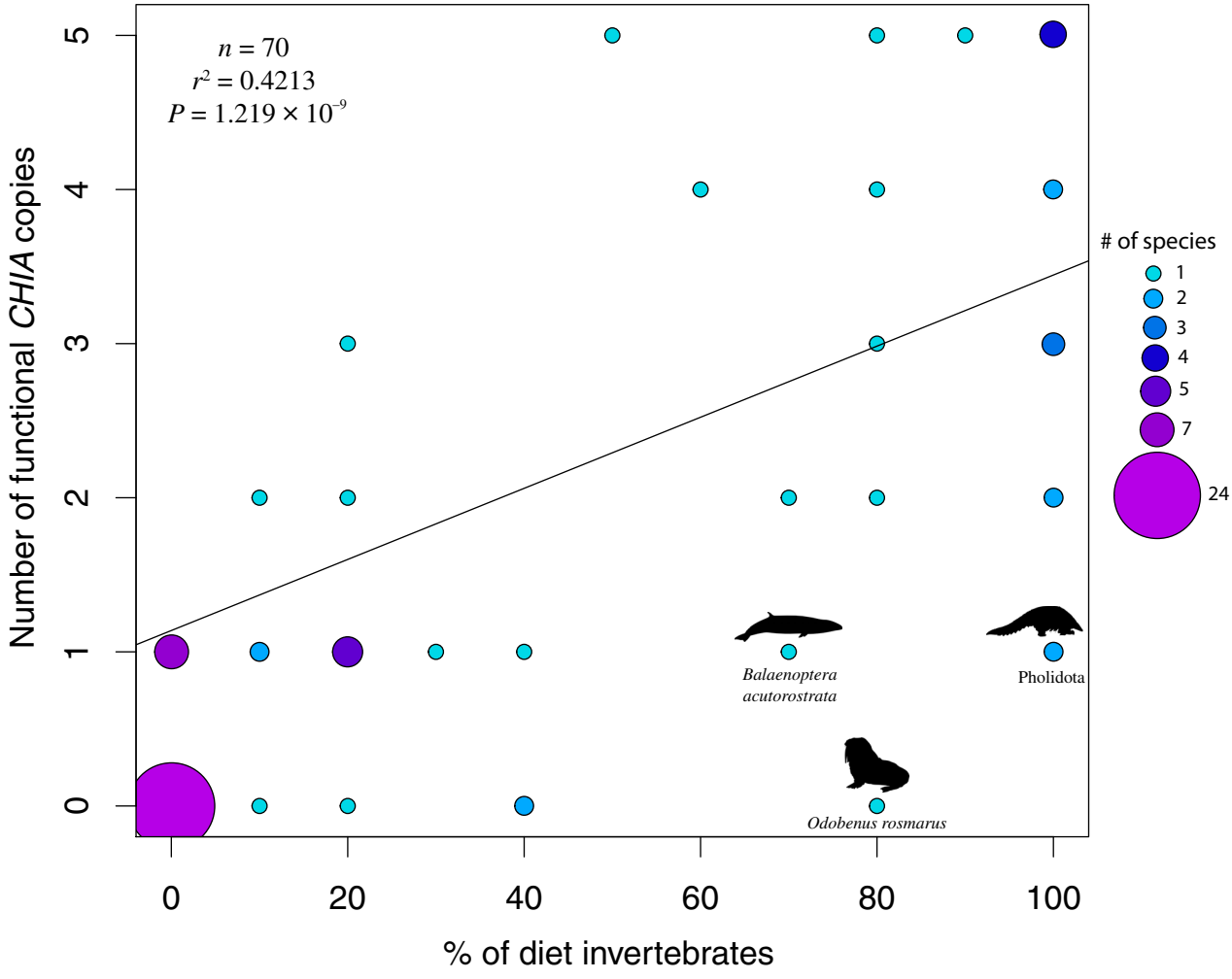

Fig. 2. PGLS regression of the number of functional contig-derived *CHIA*s versus the percent of the diet consisting of invertebrates. See text for discussion of highlighted species.

of nonsynonymous mutations after selection is relaxed. Using divergence times between lineages estimated with a molecular clock, shared inactivating mutations provide minimum dates for pseudogenization. We then analyzed the ratio of nonsynonymous to synonymous nucleotide substitutions (dN/dS) to model the timing of the transition from purifying to relaxed selection (table S6).

Estimates of the timing of *CHIA* inactivations provide support for a dietary radiation of placental mammals after the K/Pg mass extinction 66 million years (Ma) ago (Fig. 4 and table S7). Assuming a densely calibrated molecular timetree with broad taxonomic representation (*24*), the earliest *CHIA* pseudogene arose just before the K/Pg boundary (67.5 Ma ago), with 6 having occurred during the Paleocene (64.2 to 57 Ma ago) and 11 in the Eocene (53.4 to 37.4 Ma ago). Although we were unable to precisely estimate some inactivation dates due to assumption violations of the codon evolution models (see below), shared inactivating mutations (Fig. 3 and table S4) provide minimum dates for gene loss in these cases. Therefore, we estimate that another two *CHIA* inactivations occurred at the very end of the Cretaceous (67 to 66 Ma ago), six occurred in the Paleocene (63 to 56 Ma ago), and eight in the Eocene (55.5 to 34 Ma ago; Fig. 4).

There remain a handful of additional gene losses that may have occurred in the Cretaceous, but dating these events is less precise. For example, multiple *CHIA*s in rodents are extremely fragmentary pseudogenes or appear to have been completely deleted (table S1 and S4 and fig. S4); one parsimonious solution is that there were multiple *CHIA* losses early in rodent history, as early as the Cretaceous. Alternatively, these *CHIA*s may have been lost repeatedly, paralleling the patterns we found across placental mammals (Fig. 4). Additional sequencing of rodent genomes belonging to long phylogenetic branches may provide further clarity on this issue. Another example concerns a possible shared 1–base pair (bp) deletion between cetartiodactyls and perissodactyls in exon 10 of *CHIA3*. If this is a synapomorphy, then such a mutation would be dated to the Cretaceous (*24*–*26*) and would help to resolve the laurasiatherian polytomy. Nonetheless, 1-bp deletions are common, the positioning of this deletion is uncertain, and this exon is missing in the earliest diverging cetartiodactyl and ostentorian lineages, indicating that this mutation plausibly occurred convergently. Hence, various uncertainties concerning the timing of specific gene deletions and the precise dating of phylogenies and gene inactivations (Fig. 4) allow for the possibility of a few Cretaceous *CHIA* losses. However, the available data suggest that placental mammals began reducing their number of *CHIA* copies and, therefore, their consumption of insects, at and following the K/Pg boundary (Fig. 4), a conclusion broadly consonant with both the fossil record (table S5) and inferences made

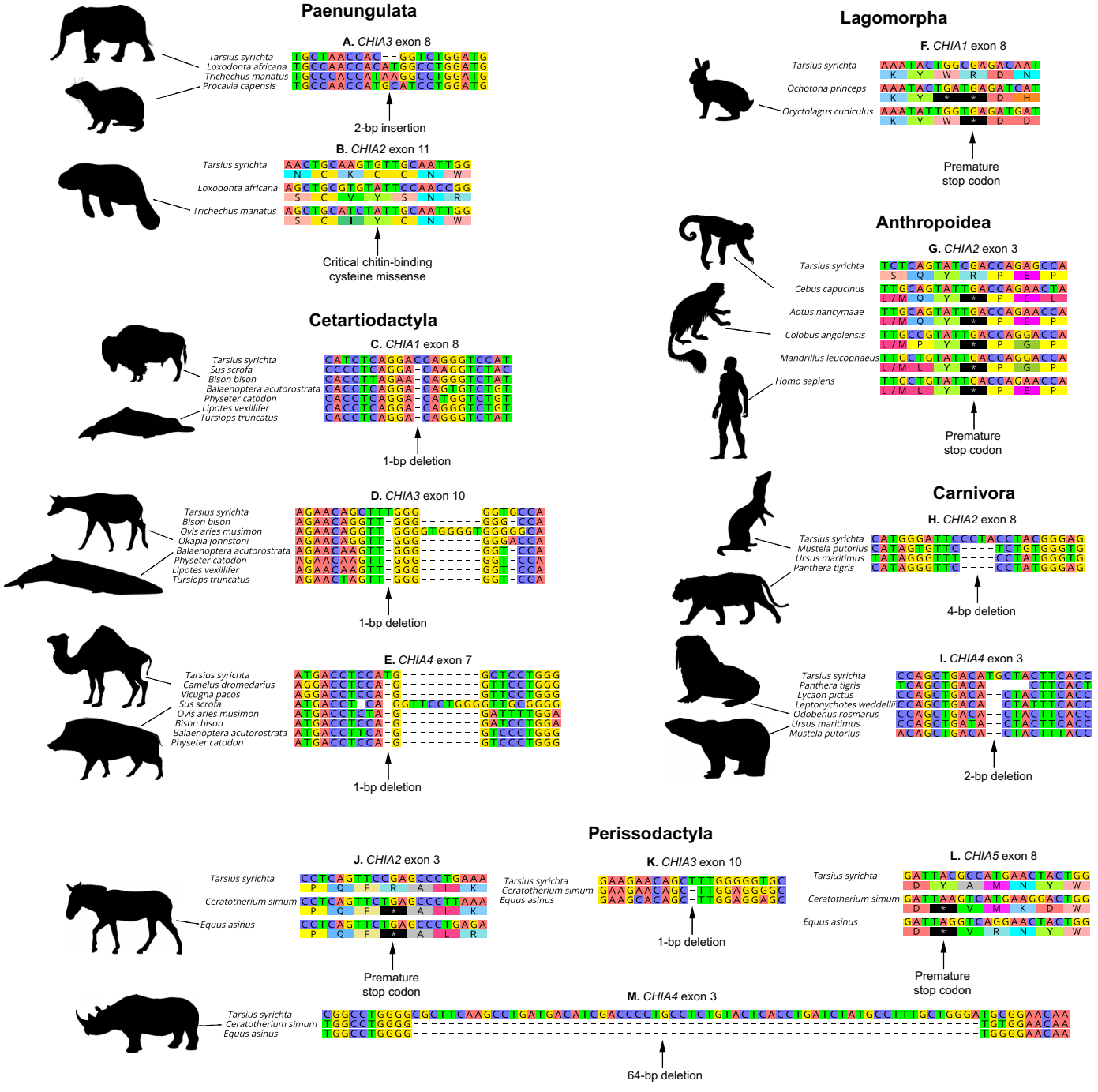

Fig. 3. Examples of *CHIA* shared inactivating mutations. Each alignment has a *Tarsius syrichta* ortholog to provide a functional outgroup comparison. See Fig. 4 for gene inactivations shown in (**A**) to (**M**) mapped onto the placental mammal phylogeny.

from correlations between the rate of genomic evolution and life-history traits (*27*).

Our results may also help reconcile seemingly incongruent results between molecular timetrees and the fossil record (*25*, *26*, *28*). Divergence time estimates based on relaxed molecular clocks routinely provide evidence of Cretaceous interordinal diversification (*25*, *26*), but analyses of the fossil record have yielded no unambiguous Cretaceous crown placental mammals (*22*, *29*). Our results suggest that Cretaceous placental diversification was uncoupled from phenotypic divergence in the early Cenozoic because diversification of placental mammals substantially predates (25.8 Ma) the repeated loss of chitinase genes beginning near the K/Pg boundary (Fig. 4).

Functional *CHIA* copy number and diet mismatches

Although the patterns of *CHIA* loss and retention largely reflect patterns of dietary evolution, there exist discrepancies that suggest that more research is needed. First, some *CHIA* copy number and diet mismatches are likely an artifact of imprecise dietary coding. We used

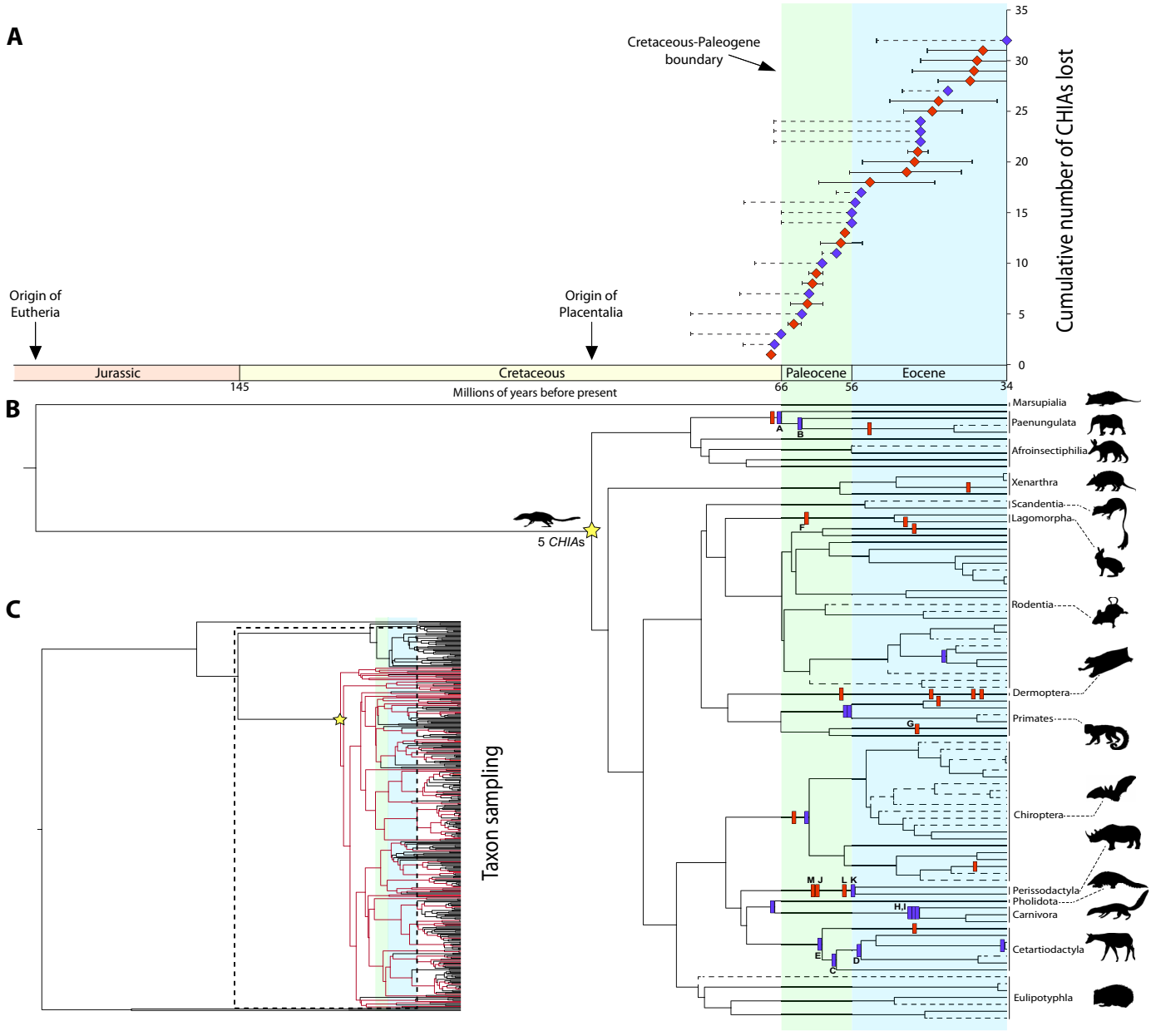

Fig. 4. Patterns of *CHIA* loss through time. (**A**) The cumulative number of *CHIA* losses in placental mammals relative to the origin of eutherians, placental mammals, and the K/Pg boundary. Red symbols indicate mean estimates of gene inactivation dates based on four model assumptions, and purple symbols indicate minimum pseudogenization dates based on shared inactivating mutations. Solid bars indicate the range of estimates (table S7), and dashed bars indicate potential range of dates based on branch lengths. (**B**) The phylogenetic positioning of the *CHIA* inactivation estimates shown in (A). Letters correspond to gene losses shown in Fig. 3. (**C**) Timetree (*24*) indicating lineages sampled for *CHIA* analyses (red branches). The dashed box indicates zoomed-in portion of phylogeny shown in (B).

"percent invertebrates in the diet" as a proxy for the amount of chitin in the diet, due to the tendency of invertebrate prey of placental mammals being insects, but some mammals eat copious amounts of invertebrates that likely have minimal amounts of chitin. For instance, the walrus (*Odobenus rosmarus*) has no functional *CHIA*s, and 80% of its diet is composed of invertebrates (Fig. 2 and table S1); however, most of these prey are bivalves (table S1). Future studies may benefit from directly calculating the amount of chitin in the diet to determine whether this leads to a more precise prediction of *CHIA* copy number.

Shared inactivating mutations suggest that other mismatches between diet and *CHIA* copy number likely result from historical contingencies, with species that inherited few *CHIA* copies being constrained by the apparent rarity of *CHIA* duplications. For instance, we found that the baleen whale *Balaenoptera acutorostrata* has only a single *CHIA* despite consuming copious amounts of chitin-rich crustaceans (Fig. 2). On the basis of shared inactivating mutations, four *CHIA*s were pseudogenized in the ancestors that baleen whales share with herbivorous hoofed mammals and cephalopod- and fish-eating toothed whales (Fig. 3 and table S4), suggesting that historical dietary shifts predating the origin of filter feeding have limited baleen whales to a single *CHIA* (table S5). Similarly, the myrmecophagous (ant-/termite-eating) pangolins (Pholidota) have a single *CHIA* (Fig. 2), in contrast to the

convergently myrmecophagous Southern tamandua (*Tamandua tetradactyla*, Vermilingua) and aardvark (*Orycteropus afer*, Tubulidentata), which have four and five functional *CHIA*s, respectively. We found evidence of *CHIA* loss in the common ancestor of pangolins and placental carnivores (table S4), which, paired with fossil evidence (table S5), suggests that pangolins may have descended from carnivorous ancestors before reverting to a strictly insectivorous diet.

Regardless of potential historical contingencies, it is unclear why secondary *CHIA* duplications appear to be so rare in placental mammals and limited to instances where chitinolytic ability has been lost (fig. S3), because this would seem to be a simple mechanism to increase chitinase activity when adapting to a chitin-rich diet. These chitinophagous, CHIA-limited mammals may be compensating via alternative mechanisms, such as increasing the expression of another *CHIA* or using chitin-digesting bacteria (*30*, *31*), chitinolytic lysozymes (*32*), or an additional mammalian chitinase, chitotriosidase (CHIT1) (*33*). *CHIT1* is generally expressed in immunity-related tissues (*33*, *34*), but we found evidence that pangolin salivary glands express *CHIT1* alongside *CHIA5* (table S3), possibly suggesting co-option for digestive function.

Other mismatches seem more difficult to explain. For instance, the shrews (Soricidae) have historically been considered to be insectivores that have retained their diets from an ancestral placental mammal stock (*35*), yet the species represented in this study (*Sorex araneus*) only has two intact *CHIA*s (table S1). This contrasts with other members of this hypothetical ancestral stock, such as tree shrews (Scandentia; five *CHIA*s), elephant shrews (Macroscelidea; five *CHIA*s), hedgehogs (Erinaceidae; four *CHIA*s), tenrecs (Tenrecidae; four *CHIA*s), golden moles (Chrysochloridae; five *CHIA*s), and true moles (Talpidae; three *CHIA*s), which all have three to five intact *CHIA*s. Although classically considered insectivores, it is possible that this designation is too coarse, because soricids frequently feed on vertebrate prey or scavenge (39.8% clade average) (*36*), and such a dietary shift may have led to some *CHIA* losses. Alternatively, some gene absences may be due to assembly errors because we cannot confirm *CHIA3*'s deletion in *S. araneus* with synteny data. Analyses of additional genomes may help clarify this and other apparent deviations (table S5).

Finally, note that we are operating under the assumption that having five *CHIA*s increases the expression of chitinases in the GI tract, allowing for greater efficiency of chitin digestion. The relationship between functional *CHIA* copy number and diet suggests that this hypothesis is quite plausible, and our gene expression analyses (table S3) provide evidence that multiple CHIAs can be expressed in the same GI organ, with two *CHIA*s being expressed in the pancreas of an insectivorous tree shrew (*Tupaia chinensis*) and three in a salivary gland of an insectivorous bat (*Macrotus californicus*). However, it remains possible that some *CHIA*s are expressed in and specialized for different tissues and may have other functions outside of the GI tract. For instance, *CHIA5* is expressed in both the GI tract and the lungs of mouse and human, indicating a possible defensive role against chitinous pathogens (*13*). Notably, only *CHIA1* and *CHIA2* appear to be robustly expressed in the tree shrew pancreas, suggesting that *CHIA3*, *CHIA4*, and *CHIA5* may be better optimized for other tissues. Furthermore, certain *CHIA*s are retained more frequently than others, with nearly half of the species surveyed having *CHIA5* versus approximately 12% having *CHIA1*, indicating a possible functional bias in *CHIA* loss. Future comparative studies could test for tissue-specific expression and differences in optimal pH activity and stability to understand the distinctions between these paralogs, which could lead to a better understanding of some of the mismatches in *CHIA* copy number and diet.

CONCLUSIONS

Together, these results suggest that chitinase genes (*CHIA*s) provide a genomic signal of the post-K/Pg dietary radiation of placental mammals. Consistent with the fossil record, the patterns of *CHIA* evolution indicate that the earliest placental mammals were highly insectivorous and that descendant lineages adapted to noninsectivorous dietary niches near and after the K/Pg boundary. Consequently, many placental mammals, including humans (*16*, *17*), retain "genomic fossils" in the form of *CHIA* pseudogenes, providing a molecular record of their insectivorous past.

MATERIALS AND METHODS

Collection of annotated sequences

We quantified the total number of acidic mammalian chitinase genes (*CHIA*s) in the genomes of 107 placental mammal species (table S1), using both annotated gene models (88 species, 54 families, and 17 orders) and direct queries of genomic contigs (74 species, 62 families, and 19 orders; fig. S7), with 55 species overlapping between the two data sets. All genomes examined were available in National Center for Biotechnology Information (NCBI)'s nucleotide collection and whole-genome shotgun (WGS) contig database except three assemblies generated using the Discovar de novo protocol (*Tamandua tetradactyla*, *Chaetophractus vellerosus*, and *Tolypeutes matacus*; https://software.broadinstitute.org/software/discovar/blog/) and made available to us by the Broad Institute. The latter assemblies were imported into Geneious v. 9.1.8 (*37*) for analysis. For annotated gene models, we obtained sequences derived from NCBI's Eukaryotic Genome Annotation (EGA) pipeline by BLASTing (discontiguous megablast) a human *CHIA* mRNA reference sequence (NM_201652) against the NCBI nucleotide collection and performing separate Basic Local Alignment Search Tool (BLAST) searches against all the placental taxonomic orders. We downloaded all relevant EGA gene models, which are denoted by an "XM_" accession number prefix and are annotated as CHIA or any annotation suggesting homology to this protein (for example, "acidic mammalian chitinase-like"). We considered the gene absent/nonfunctional under any of the following conditions: (i) we retrieved negative BLAST results; (ii) the gene model was annotated as a "low-quality protein," a designation given if construction of the model required correction for premature stop codons and/or frameshift insertions or deletions (indels) found in the reference genomic contigs; and (iii) the sequence lacked a canonical catalytic domain (DXXDXDXE) (*18*).

Gene tree analyses

We discovered numerous examples of species with potential *CHIA* paralogs, which necessitated comparison through phylogenetic analyses to estimate their evolutionary history. We obtained additional gene models from nonplacental mammal vertebrate taxa and a *CHIT1* (chitotriosidase) outgroup (NM_003465), imported all of the sequences into Geneious, aligned them using MUltiple Sequence Comparison by Log-Expectation (MUSCLE) (*38*), manually adjusted the alignments, and removed sequences and sequence positions determined by eye to have dubious homology (data set S1). We then estimated a gene tree using Randomized Axelerated Maximum Likelihood (RAxML) v. 8.2.10 (*39*) in Cyberinfrastructure for Phylogenetic Research (CIPRES) (RAxML-HPC2 on Extreme Science and Engineering Discovery Environment) (*40*) with 500 bootstrap replications but otherwise default settings (-m GTRCAT). Placental *CHIA* sequences clustered in five distinct clades,

which we designated *CHIA1*, *CHIA2*, *CHIA3*, *CHIA4*, and *CHIA5*. On the basis of these clade assignments, sequences from all five *CHIA* subclades were realigned separately (that is, all *CHIA1* sequences, all *CHIA2* sequences, etc.), and then each subclade alignment was aligned successively using the Geneious Translation align tool implementing the MUSCLE algorithm. We then performed a RAxML analysis on this finalized alignment (data set S1) in CIPRES using the same parameters (fig. S1).

Assembly of contig-derived sequences

Given the inference of five distinct clades of *CHIA* within placental mammals, we directly queried the genomic contigs (WGS) of a subset of species to verify the presence of distinct *CHIA* genes (data set S2), determine their synteny, and validate the presence of inactivating mutations. Some genome assemblies present in WGS had not been annotated by EGA at the time of our study, so we examined additional genomes to broaden taxonomic diversity. We directly queried at least one member of each mammalian family for which there was a genome assembly available and, in some cases, examined additional species from a family if we were unable to recover a particular *CHIA* paralog or to check for inactivating mutations shared by confamilials. On the basis of the predicted gene models, very few mammals appeared to have all five predicted *CHIA* genes. We chose one of these species, the Philippine tarsier (*T. syrichta*), as the reference genome from which to obtain and compare all five *CHIA* genes (fig. S1). We BLASTed (megablast) the *T. syrichta* gene models against its WGS assembly and confirmed that all five *T. syrichta CHIA* sequences map to separate genomic regions (table S1). Each of these sequences is characterized by a functional catalytic domain in exon 5 (DXXDXDXE; fig. S5) (*18*), six critical cysteines in the chitin-binding domain (fig. S6) (*19*), intact exon-intron boundaries and the absence of inactivating mutations. Each gene model has 11 coding exons, with the exception of *CHIA2*, which has 10. After comparisons with gene models from other species, it became apparent that exon 2 was erroneously excluded from the *T. syrichta* gene model, despite being present in the assembly.

We designed a set of five in silico probes from the *T. syrichta* contigs, one for each *CHIA* gene, to capture *CHIA* sequences in the genome assemblies of other placental mammals. These probes encompassed the exons, introns, and flanking sequences of each of the predicted genes, allowing for greater confidence in homology and synteny between the respective paralogs. We BLASTed the probes (discontiguous megablast) against the WGS and performed local BLAST searches on the three Discovar de novo assemblies (discontiguous megablast). In instances of negative BLAST results (that is, no hits for *CHIA* sequences), we designed and BLASTed probes based on more closely related taxa. When we obtained BLAST hits, we downloaded the entire contiguous sequence that encompassed all the hits with the highest homology, ignoring hits that were clearly repetitive elements. As we downloaded each sequence, we recorded the accession number of each contig, the position of the BLAST hits within each contig, and the position of flanking genes to determine synteny (tables S1, S2, and S8 and fig. S4). We imported each captured sequence into Geneious, aligned the imported sequence to the *T. syrichta* probe reference using MUSCLE, and successively aligned additional sequences derived from the same ordinal and/or superordinal taxon, frequently using alignments of outgroup taxa to anchor the new alignments (data set S2). We examined each of the gene sequences for putative inactivating mutations, including missense mutations to the conserved catalytic domain or chitin-binding domain residues, start codon mutations, frameshift indels, premature stop codons, and splice site mutations (table S4). If a sequence had any of these mutations, it was assumed to be a pseudogene, unless its only mutation(s) is (are) (i) a start codon mutation but it had an alternative start upstream of the first exon or (ii) a GT→GC splice donor mutation because GC is a relatively common acceptable splice donor variant (*41*). If a sequence had a substantial amount of missing sequence (five or more exons) and its closest available relative(s) had a pseudogenic ortholog, then we assumed that it is a pseudogene.

After finalizing this data set, we removed all introns and flanking sequences, aligned each sequence to its corresponding orthologs (for example, all *CHIA1*s, all *CHIA2*s, etc.) using MUSCLE in Geneious, and then aligned each individual paralog alignment to each other successively using Geneious' Translation align with the MUSCLE algorithm. Using this master alignment, we aligned (MUSCLE in Geneious) the following published *CHIA* mRNAs: *Macaca fascicularis* (NM_001284548, stomach), *Homo sapiens* (NM_201653, stomach/lung), *Mus musculus* (NM_023186, salivary gland), *Rattus norvegicus* (NM_207586, stomach), *Bos taurus* (NM_174699, fetal liver), and *Sus scrofa* (NM_001258377, lung). Next, we aligned (MUSCLE in Geneious) the following mRNAs assembled from RNA sequencing experiments (see details below): *T. chinensis CHIA1* (pancreas), *T. chinensis CHIA2* (pancreas), *M. californicus CHIA3* (submandibular gland), and *M. californicus CHIA4* (submandibular gland). Finally, we aligned (Translation align, MUSCLE in Geneious) several chitinase-like mRNA sequences expressed in *M. musculus*: *Chil6*/*BYm*/*Chil8* (NM_178412), *Chil4*/*Ym2*/*Chi3l4* (NM_145126), *Chil3*/*Ym1*/*Chi3l3* (NM_009892), and *Chil5*/*Bclp*/*Chi3l7* (NM_001080816), as well as a *CHIT1* outgroup (NM_003465). We examined the master alignment by eye, adjusted it manually, removed sequence positions of dubious homology, and removed sequences that were likely to be phylogenetically uninformative because of missing data and/or minimal or no overlap with related species (for example, one exon only). We then executed a RAxML analysis through CIPRES on this contig-derived sequence alignment (data set S1) using the methods described above.

The resulting tree (fig. S2) largely recapitulated the results from the gene model–derived analyses (fig. S1) with the major exception of *CHIA4* being nested within *CHIA5*. Upon closer inspection, it became apparent that the chiropteran *CHIA4* and *CHIA5* sequences seemed to be influencing this aberrant topology. Specifically, chiropteran *CHIA4* is paraphyletic at the base of the *CHIA4* clade, and chiropteran *CHIA5* sequences form the sister group to the entire *CHIA4* clade. Reexamining the gene model–derived phylogeny (fig. S1) shows that the chiropteran gene models do not include any *CHIA5* sequences but seem to imply multiple recent gene duplications among the chiropteran *CHIA4*s. However, *CHIA4* and *CHIA5* sequences can readily be distinguished as occupying distinct positions on the same contig for multiple, distantly related species [*Rousettus aegyptiacus* (Pteropodidae), *Hipposideros armiger* (Hipposideridae), *Rhinolophus sinicus* (Rhinolophidae), *Miniopterus natalensis* (Miniopteridae), and *Eptesicus fuscus* (Vespertilionidae); table S2 and fig. S2].

Upon directly examining the sequences, we found evidence that gene conversion may be to blame for these discordant results (table S9). Specifically, when comparing chiropteran *CHIA4* and *CHIA5* sequences of the same species to a reference *CHIA4* sequence (*T. syrichta*), we found that most exons were more similar between the *CHIA4* and *CHIA5* paralogs than the chiropteran *CHIA4*s were to the orthologous *T. syrichta CHIA4* sequence. We compared the introns for two species and found a similar pattern. To give a detailed example, *E. fuscus CHIA4* exons 1, 2, 4, and 11 had higher similarity to *T. syrichta CHIA4* (84.6 to

96.7%; average, 91.9%) than to *E. fuscus CHIA5* (66.2 to 89.5%; average, 74.8%), whereas *E. fuscus CHIA4* exons 3 and 5 to 10 had higher similarity to *E. fuscus CHIA5* (95.5 to 100%; average, 97.7%) than to *T. syrichta CHIA4* (90.1 to 95.2%; average, 92.3%). Similarly, *E. fuscus CHIA4* introns 2 to 4, 7, and 10 were more similar to *T. syrichta CHIA4* (40.8 to 67%; average, 59.3%) than to *E. fuscus CHIA5* (11.2 to 36.7%; average, 25%), whereas *E. fuscus CHIA4* introns 5, 6, 8, and 9 had higher identity to *E. fuscus CHIA5* (76.7 to 96.5%; average, 89.7%) than to *T. syrichta CHIA4* (49.8 to 71.6 %; average, 60.8%). Further examination of another 10 mammals that retain functional *CHIA4* and *CHIA5* indicates that gene conversion for these two paralogs is likely phylogenetically widespread. This is perhaps an unsurprising conclusion given that *CHIA4* and *CHIA5* are in tandem (fig. S4), appear to be derived from a relatively recent (placental mammal-specific?) gene duplication (fig. S1), and that gene conversion is prone to occurring in tandemly duplicated genes that share high homology (*42*). We reanalyzed the contig-derived data set after removing all chiropteran *CHIA4* and *CHIA5* sequences (data set S1), and the resulting phylogeny (Fig. 1 and fig. S3) resolves *CHIA4* and *CHIA5* as separate clades.

Gene models versus contig-derived sequences

Of the 311 genes for which we directly compared both EGA gene models and sequences derived from genomic contigs (WGS), 95.5% (297) were in agreement about functionality, that is, both predicted a functional gene or both predicted a nonfunctional/absent gene, although many absent gene models could be found as pseudogenes in the contigs. The exceptions included sequences predicted to be of low quality but no inactivating mutations could be found in the contig sequence (two), instances where EGA did not predict a sequence to be present in a genome but we confirmed that the gene is present and intact (nine), and cases in which EGA predicted a functional gene but we found evidence of pseudogenization (three). Although the two data sets are highly congruent, the errors in the gene models led us to base all our results and discussion in the paper and the Supplementary Materials on the contig-derived sequences, with two exceptions: one of two PGLS regression analyses referred to in Results and fig. S1.

Regression analyses

We tested whether the number of intact *CHIA*s in a genome correlates with the amount of insect prey consumed by a species. We derived the latter metric from EltonTraits v. 1.0 (*36*), which provides the estimated percentage of invertebrate prey in the diet, a proxy for the amount of insects consumed. We performed PGLS regression analyses with the caper package in R (*43*, *44*), testing the hypothesis separately with the gene model (EGA) and contig-derived (WGS) data sets. We assumed the timetree from Emerling *et al.* (*24*) and implemented phylogenetic corrections using maximum-likelihood estimates of Pagel's λ. We made taxonomic substitutions when species in our data set were not represented in the Emerling *et al.* (*24*) phylogeny. Specifically, we assumed that species in the same taxon (for example, genus and family) could be interchanged if there were no other congeners or confamilials in the phylogeny. For instance, Emerling *et al.* (*24*) have only a single hipposiderid in their tree, *Hipposideros commersoni*, and our chitinase data set included *H. armiger*, which we deemed to be an acceptable substitution. Species were removed from the analyses if there were no taxonomic equivalents.

Estimation of gene inactivation dates

To estimate the temporal distribution of *CHIA* loss during placental history, we used the dN/dS ratio method of Meredith *et al.* (*23*) to date the inactivation of the different *CHIA* paralogs. This method is derived from the assumption that genes, on average, are under purifying selection throughout evolutionary history until they become inactivated. Once inactivated, these pseudogenes are expected to undergo relaxed selection on average, and the dN/dS ratio on the transitional ("mixed") branch (that is, transitioning from purifying to relaxed selection) provides a signal of the timing of this shift in selection pressures. This can then be used to calculate the timing of gene inactivation, assuming divergence times obtained from molecular dating analyses, with shared inactivating mutations (table S4) providing a minimum branch on which pseudogenization occurred. We performed separate analyses for each *CHIA* paralog (data set S1) using PAML (phylogenetic analysis by maximum likelihood) v. 4.8 (*45*), implementing two separate codon model assumptions (F1X4 and F3X4) and assuming the topology and divergence times of Emerling *et al.* (*24*). We removed all chiropteran *CHIA4* and *CHIA5* sequences due to their disproportionate influence on the gene tree topology (fig. S2), and we assumed that any other *CHIA* gene conversion has occurred at an equal rate on average during the evolution of placental mammals.

When performing dN/dS ratio analyses on a particular *CHIA* paralog, we allowed for estimation of ω (dN/dS ratio) on each of the mixed category branches (purifying to relaxed selection), each individual clade of pseudogene (relaxed selection) branches (that is, each set of descendant branches that post-dates a mixed branch), and the remaining branches were grouped together as functional (purifying selection) category branches [table S6; see the study of Meredith *et al.* (*23*) for additional details]. After this initial analysis, we successively fixed each set of pseudogene branches to ω = 1 to test whether the dN/dS estimates are statistically distinguishable from an assumption of relaxed selection. Furthermore, in instances where mixed branches were estimated to have a dN/dS ratio > 1, we similarly compared them to a null model where the branch was fixed at ω = 1. Finally, we implemented a model whereby all pseudogene category branches were fixed at ω = 1 (table S6). The mixed category branches estimated during this final analysis were then entered into the calculations of Meredith *et al.* (*23*) to obtain point estimates of gene inactivation (table S7). We discarded estimates where the mixed branch ω was greater than 1 and/or the pseudogene branch dN/dS ratios significantly deviated from a model where ω was fixed at 1, due to the violation of the assumption that pseudogenization involves a transition from purifying to relaxed selection.

CHIA gene expression analyses

Because the presence of multiple *CHIA* paralogs in the genomes of placental mammals does not guarantee their expression in GI tissues, we queried RNA sequencing experiments to test whether all five *CHIA*s can be expressed in the GI tract. We examined GI tissue RNA libraries of three species that have (*T. chinensis*) or potentially have (*M. californicus* and *Desmodus rotundus*) three or more *CHIA*s in their genomes (tables S1 and S3). These libraries have been deposited into NCBI's Sequence Read Archive, so we queried them via NCBI's BLAST interface. We BLASTed (megablast) all five of the assembled *T. chinensis CHIA* paralogs against a pancreas library (*46*). *M. californicus* and *D. rotundus* are phyllostomid chiropterans with no available genome assemblies, so we BLASTed (discontiguous megablast) assembled *CHIA3* to *CHIA5* sequences from *Pteronotus parnellii*, given that mormoopids are the closest living relatives to phyllostomids. We BLASTed against the single available salivary gland transcriptome for *M. californicus* and five salivary gland transcriptomes for *D. rotundus* (*47*–*49*).

D. rotundus, in contrast to *M. californicus*, showed no evidence of robust transcription of *CHIA*s in its salivary glands, with BLAST hits ranging from zero to 13 for *CHIA3* to *CHIA5* (table S3). We tested whether an additional mammalian chitinase (*CHIT1*) and prolactin-induced protein (*PIP*), a salivary gland positive control, show evidence of gene expression in *M. californicus* and *D. rotundus*, by BLASTing (discontiguous megablast) the coding sequences of *Myotis brandtii* gene models for both *CHIT1* (XM_005881866.2) and *PIP* (XM_005862784.2) against all six RNA libraries. We found evidence of robust transcription of *CHIT1* in the *M. californicus* salivary gland (5963 hits), compared to practically no hits in *D. rotundus* (0 to 6 hits), but *PIP* returned a large amount of hits for both *M. californicus* (20,000 hits) and four of the five *D. rotundus* libraries (6317 to 12,547 hits; table S3). This suggests that *D. rotundus* does not express *CHIA*s in its salivary glands, whereas *M. californicus* does, an observation consistent with their respective sanguivorous and insectivorous diets.

Because of the surprisingly minimal number of *CHIA*s in the genomes of the highly insectivorous pangolins (table S1), possibly due to a historical contingency (table S5), we tested whether the immune chitinase gene *CHIT1* is expressed in the GI tract alongside *CHIA5*. In NCBI, we BLASTed (megablast) the assembled *Manis javanica CHIA5* and the coding sequence of an *M. javanica CHIT1* gene model (XM_017667770.1) against an *M. javanica* salivary gland RNA library (table S3) (*50*).

SUPPLEMENTARY MATERIALS

Supplementary material for this article is available at http://advances.sciencemag.org/cgi/content/full/4/5/eaar6478/DC1

fig. S1. Vertebrate RAxML *CHIA* gene tree based on *CHIA* gene models.
fig. S2. Placental mammal RAxML *CHIA* gene tree based on genomic contig–derived and mRNA sequences.
fig. S3. Placental mammal RAxML *CHIA* gene tree based on genomic contig–derived and mRNA sequences, after removing chiropteran *CHIA4* and *CHIA5* sequences.
fig. S4. Diagram demonstrating synteny of *CHIA* paralogs and nearby genes on their respective genomic scaffolds/contigs.
fig. S5. *T. syrichta CHIA1*, *CHIA2*, *CHIA3*, *CHIA4*, and *CHIA5* gene models showing conserved chitinolytic domain (DXXDXDXE).
fig. S6. *T. syrichta CHIA1*, *CHIA2*, *CHIA3*, *CHIA4*, and *CHIA5* gene models showing chitin-binding domain with six conserved cysteine residues.
fig. S7. Cladogram showing relationships of species in WGS data set and mapped gene losses.
table S1. *CHIA* summary.
table S2. Synteny of *CHIA* paralogs.
table S3. *CHIA* gene expression analyses.
table S4. Distribution of genetic lesions and missing data across *CHIA* exons.
table S5. Discussion of *CHIA* number, timing of loss, diets of extant taxa, and the fossil record.
table S6. dN/dS ratio models used for gene inactivation estimates.
table S7. *CHIA* inactivation date estimates calculated from dN/dS ratio models in table S6 and assuming divergence dates from the study of Emerling *et al.* (*24*).
table S8. Scaffold/contig coordinates of *OVGP1*, *PIFO*, and *DENND2D* for species in fig. S4.
table S9. Evidence of gene conversion between *CHIA4* and *CHIA5*.
data set S1. Alignments used for analyses.
data set S2. Alignments used for constructing *CHIA* genes from scaffolds/contigs.
References (*51*–*79*)

REFERENCES AND NOTES

1. T. S. Kemp, *The Origin and Evolution of Mammals* (Oxford Univ. Press, 2005).
2. J. D. Archibald, *Extinction and Radiation: How the Fall of Dinosaurs Led to the Rise of Mammals* (JHU Press, 2011).
3. Z.-X. Luo, Transformation and diversification in early mammal evolution. *Nature* **450**, 1011–1019 (2007).
4. D. M. Grossnickle, P. D. Polly, Mammal disparity decreases during the Cretaceous angiosperm radiation. *Proc. Biol. Sci.* **280**, 20132110 (2013).
5. R. A. Close, M. Friedman, G. T. Lloyd, R. B. J. Benson, Evidence for a mid-Jurassic adaptive radiation in mammals. *Curr. Biol.* **25**, 2137–2142 (2015).
6. G. P. Wilson, A. R. Evans, I. J. Corfe, P. D. Smits, M. Fortelius, J. Jernvall, Adaptive radiation of multituberculate mammals before the extinction of dinosaurs. *Nature* **483**, 457–460 (2012).
7. Z.-X. Luo, C.-X. Yuan, Q.-J. Meng, Q. Ji, A Jurassic eutherian mammal and divergence of marsupials and placentals. *Nature* **476**, 442–445 (2011).
8. T. J. D. Halliday, A. Goswami, Eutherian morphological disparity across the end-Cretaceous mass extinction. *Biol. J. Linn. Soc.* **118**, 152–168 (2016).
9. D. M. Grossnickle, E. Newham, Therian mammals experience an ecomorphological radiation during the Late Cretaceous and selective extinction at the K–Pg boundary. *Proc. Biol. Sci.* **283**, 20160256 (2016).
10. F. A. Smith, A. G. Boyer, J. H. Brown, D. P. Costa, T. Dayan, S. K. Ernest, A. R. Evans, M. Fortelius, J. L. Gittleman, M. J. Hamilton, L. E. Harding, K. Lintulaakso, S. K. Lyons, C. McCain, J. G. Okie, J. J. Saarinen, R. M. Sibly, P. R. Stephens, J. Theodor, M. D. Uhen, The evolution of maximum body size of terrestrial mammals. *Science* **330**, 1216–1219 (2010).
11. M. S. Springer, C. A. Emerling, R. W. Meredith, J. E. Janečka, E. Eizirik, W. J. Murphy, Waking the undead: Implications of a soft explosive model for the timing of placental mammal diversification. *Mol. Phylogenet. Evol.* **106**, 86–102 (2017).
12. C. Jeuniaux, On some biochemical aspects of regressive evolution in animals, in *Biochemical Evolution and the Origin of Life*, E. Schoffeniels, Ed. (North-Holland Publishing Company, 1971), pp. 304–313.
13. R. G. Boot, E. F. C. Blommaart, E. Swart, K. Ghauharali-van der Vlugt, N. Bijl, C. Moe, A. Place, J. M. F. G. Aerts, Identification of a novel acidic mammalian chitinase distinct from chitotriosidase. *J. Biol. Chem.* **276**, 6770–6778 (2001).
14. S. Strobel, A. Roswag, N. I. Becker, T. E. Trenczek, J. A. Encarnação, Insectivorous bats digest chitin in the stomach using acidic mammalian chitinase. *PLOS ONE* **8**, e72770 (2013).
15. M. Ohno, M. Kimura, H. Miyazaki, K. Okawa, R. Onuki, C. Nemoto, E. Tabata, S. Wakita, A. Kashimura, M. Sakaguchi, Y. Sugahara, N. Nukina, P. O. Bauer, F. Oyama, Acidic mammalian chitinase is a proteases-resistant glycosidase in mouse digestive system. *Sci. Rep.* **6**, 37756 (2016).
16. M. Hussain, J. B. Wilson, New paralogues and revised time line in the expansion of the vertebrate GH18 family. *J. Mol. Evol.* **76**, 240–260 (2013).
17. M. C. Janiak, M. E. Chaney, A. J. Tosi, Evolution of acidic mammalian chitinase genes (*CHIA*) is related to body mass and insectivory in primates. *Mol. Biol. Evol.* **35**, 607–622 (2017).
18. A. M. Olland, J. Strand, E. Presman, R. Czerwinski, D. Joseph-McCarthy, R. Krykbaev, G. Schlingmann, R. Chopra, L. Lin, M. Fleming, R. Kriz, M. Stahl, W. Somers, L. Fitz, L. Mosyak, Triad of polar residues implicated in pH specificity of acidic mammalian chitinase. *Protein Sci.* **18**, 569–578 (2009).
19. L. W. Tjoelker, L. Gosting, S. Frey, C. L. Hunter, H. L. Trong, B. Steiner, H. Brammer, P. W. Gray, Structural and functional definition of the human chitinase chitin-binding domain. *J. Biol. Chem.* **275**, 514–520 (2000).
20. J. Nio, W. Fujimoto, A. Konno, Y. Kon, M. Owhashi, T. Iwanaga, Cellular expression of murine Ym1 and Ym2, chitinase family proteins, as revealed by in situ hybridization and immunohistochemistry. *Histochem. Cell Biol.* **121**, 473–482 (2004).
21. J. Romiguier, V. Ranwez, E. J. P. Douzery, N. Galtier, Genomic evidence for large, long-lived ancestors to placental mammals. *Mol. Biol. Evol.* **30**, 5–13 (2013).
22. M. A. O'Leary, J. I. Bloch, J. J. Flynn, T. J. Gaudin, A. Giallombardo, N. P. Giannini, S. L. Goldberg, B. P. Kraatz, Z.-X. Luo, J. Meng, X. Ni, M. J. Novacek, F. A. Perini, Z. S. Randall, G. W. Rougier, E. J. Sargis, M. T. Silcox, N. B. Simmons, M. Spaulding, P. M. Velazco, M. Weksler, J. R. Wible, A. L. Cirranello, The placental mammal ancestor and the post–K-Pg radiation of placentals. *Science* **339**, 662–667 (2013).
23. R. W. Meredith, J. Gatesy, W. J. Murphy, O. A. Ryder, M. S. Springer, Molecular decay of the tooth gene Enamelin (*ENAM*) mirrors the loss of enamel in the fossil record of placental mammals. *PLOS Genet.* **5**, e1000634 (2009).
24. C. A. Emerling, H. T. Huynh, M. A. Nguyen, R. W. Meredith, M. S. Springer, Spectral shifts of mammalian ultraviolet-sensitive pigments (short wavelength-sensitive opsin 1) are associated with eye length and photic niche evolution. *Proc. R. Soc. B* **282**, 20151817 (2015).
25. R. W. Meredith, J. E. Janečka, J. Gatesy, O. A. Ryder, C. A. Fisher, E. C. Teeling, A. Goodbla, E. Eizirik, T. L. Simão, T. Stadler, D. L. Rabosky, R. L. Honeycutt, J. J. Flynn, C. M. Ingram, C. Steiner, T. L. Williams, T. J. Robinson, A. Burk-Herrick, M. Westerman, N. A. Ayoub, M. S. Springer, W. J. Murphy, Impacts of the Cretaceous Terrestrial Revolution and KPg extinction on mammal diversification. *Science* **334**, 521–524 (2011).
26. M. dos Reis, J. Inoue, M. Hasegawa, R. J. Asher, P. C. Donoghue, Z. Yang, Phylogenomic datasets provide both precision and accuracy in estimating the timescale of placental mammal phylogeny. *Proc. Biol. Sci.* **279**, 3491–3500 (2012).
27. J. Wu, T. Yonezawa, H. Kishino, Rates of molecular evolution suggest natural history of life history traits and a post-K-Pg nocturnal bottleneck of placentals. *Curr. Biol.* **27**, 3025–3033.e5 (2017).
28. J. D. Archibald, D. H. Deutschman, Quantitative analysis of the timing of the origin and diversification of extant placental orders. *J. Mamm. Evol.* **8**, 107–124 (2001).
29. T. J. D. Halliday, P. Upchurch, A. Goswami, Resolving the relationships of Paleocene placental mammals. *Biol. Rev.* **92**, 521–550 (2017).
30. F. Delsuc, J. L. Metcalf, L. Wegener Parfrey, S. J. Song, A. González, R. Knight, Convergence of gut microbiomes in myrmecophagous mammals. *Mol. Ecol.* **23**, 1301–1317 (2014).
31. J. G. Sanders, A. C. Beichman, J. Roman, J. J. Scott, D. Emerson, J. J. McCarthy, P. R. Girguis, Baleen whales host a unique gut microbiome with similarities to both carnivores and herbivores. *Nat. Commun.* **6**, 8285 (2015).
32. Y. Liu, G. He, H. Xu, X. Han, G. Jones, S. J. Rossiter, S. Zhang, Adaptive functional diversification of lysozyme in insectivorous bats. *Mol. Biol. Evol.* **31**, 2829–2835 (2014).

33. R. G. Boot, G. H. Renkema, A. Strijland, A. J. Van Zonneveld, J. M. F. G. Aerts, Cloning of a cDNA encoding chitotriosidase, a human chitinase produced by macrophages. *J. Biol. Chem.* **270**, 26252–26256 (1995).
34. R. G. Boot, A. P. Bussink, M. Verhoek, P. A. de Boer, A. F. Moorman, J. M. Aerts, Marked differences in tissue-specific expression of chitinases in mouse and man. *J. Histochem. Cytochem.* **53**, 1283–1292 (2005).
35. W. D. Matthew, The carnivora and the insectivora of the bridger basin, middle eocene. *Mem. Am. Museum Nat. Hist.* **9**, 1199–203 (1909).
36. H. Wilman, J. Belmaker, J. Simpson, C. de la Rosa, M. M. Rivadeneira, W. Jetz, EltonTraits 1.0: Species-level foraging attributes of the world's birds and mammals. *Ecology* **95**, 2027 (2014).
37. M. Kearse, R. Moir, A. Wilson, S. Stones-Havas, M. Cheung, S. Sturrock, S. Buxton, A. Cooper, S. Markowitz, C. Duran, T. Thierer, B. Ashton, P. Meintjes, A. Drummond, Geneious Basic: An integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics* **28**, 1647–1649 (2012).
38. R. C. Edgar, MUSCLE: Multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* **32**, 1792–1797 (2004).
39. A. Stamatakis, RAxML version 8: A tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* **30**, 1312–1313 (2014).
40. M. A. Miller, W. Pfeiffer, T. Schwartz, Creating the CIPRES Science Gateway for inference of large phylogenetic trees, in *2010 Gateway Computing Environments Workshop, GCE 2010* (IEEE, 2010), pp. 1–8.
41. M. Burset, I. A. Seledtsov, V. V. Solovyev, Analysis of canonical and non-canonical splice sites in mammalian genomes. *Nucleic Acids Res.* **28**, 4364–4375 (2000).
42. J.-M. Chen, D. N. Cooper, N. Chuzhanova, C. Férec, G. P. Patrinos, Gene conversion: Mechanisms, evolution and human disease. *Nat. Rev. Genet.* **8**, 762–775 (2007).
43. A. Grafen, The phylogenetic regression. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **326**, 119–157 (1989).
44. D. Orme, R. Freckleton, G. Thomas, T. Petzoldt, S. Fritz, N. Isaac, W. Pearse, "caper: Comparative analyses of phylogenetics and evolution in R," 29 November 2013; https://cran.r-project.org/web/packages/caper/index.html.
45. Z. Yang, PAML 4: Phylogenetic analysis by maximum likelihood. *Mol. Biol. Evol.* **24**, 1586–1591 (2007).
46. Y. Fan, Z.-Y. Huang, C.-C. Cao, C.-S. Chen, Y.-X. Chen, D.-D. Fan, J. He, H. L. Hou, L. Hu, X.-T. Hu, X.-T. Jiang, R. Lai, Y.-S. Lang, B. Liang, S.-G. Liao, D. Mu, Y.-Y. Ma, Y. Y. Niu, X.-Q. Sun, J.-Q. Xia, J. Xiao, Z.-Q. Xiong, L. Xu, L. Yang, Y. Zhang, W. Zhao, X.-D. Zhao, Y.-T. Zheng, J.-M. Zhou, Y.-B. Zhu, G.-J. Zhang, J. Wang, Y.-G. Yao, Genome of the Chinese tree shrew. *Nat. Commun.* **4**, 1426 (2013).
47. D. H. W. Low, K. Sunagar, E. A. Undheim, S. A. Ali, A. C. Alagon, T. Ruder, T. N. Jackson, S. Pineda Gonzalez, G. F. King, A. Jones, A. Antunes, B. G. Fry, Dracula's children: Molecular evolution of vampire bat venom. *J. Proteomics* **89**, 95–111 (2013).
48. I. M. B. Francischetti, T. C. Assumpção, D. Ma, Y. Li, E. C. Vicente, W. Uieda, J. M. Ribeiro, 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 (2013).
49. C. D. Phillips, R. J. Baker, Secretory gene recruitments in vampire bat salivary adaptation and potential convergences with sanguivorous leeches. *Front. Ecol. Evol.* **3**, 1–11 (2015).
50. J.-E. Ma, L. M. Li, H. Y. Jiang, X. J. Zhang, J. Li, G. Y. Li, L. H. Yuan, J. Wu, J. P. Chen, Transcriptomic analysis identifies genes and pathways related to myrmecophagy in the Malayan pangolin (*Manis javanica*). *PeerJ* **5**, e4140 (2017).
51. R. M. Nowak, *Walker's Mammals of the World* (Johns Hopkins Univ. Press, 1999).
52. M. Hofreiter, H. N. Poinar, W. G. Spaulding, K. Bauer, P. S. Martin, G. Possnert, S. Pääbo, A molecular analysis of ground sloth diet through the last glaciation. *Mol. Ecol.* **9**, 1975–1984 (2000).
53. M. S. Bargo, The ground sloth *Megatherium americanum*: Skull shape, bite forces, and diet. *Acta Palaeontol. Pol.* **46**, 173–192 (2001).
54. R. K. McAfee, Feeding mechanics and dietary implications in the fossil sloth *Neocnus* (Mammalia: Xenarthra: Megalonychidae) from Haiti. *J. Morphol.* **272**, 1204–1216 (2011).
55. M. C. M. McKenna, A. A. R. Wyss, J. J. Flynn, Paleogene pseudoglyptodont xenarthrans from central Chile and Argentine Patagonia. *Am. Mus. Novit.* **3536**, 1–18 (2006).
56. S. F. Vizcaíno, G. H. Cassini, J. C. Fernicola, M. Susana Bargo, Evaluating habitats and feeding habits through ecomorphological features in glyptodonts (Mammalia, xenarthra). *Ameghiniana* **48**, 305–319 (2011).
57. E. Gheerbrant, Paleocene emergence of elephant relatives and the rapid radiation of African ungulates. *Proc. Natl. Acad. Sci. U.S.A.* **106**, 10717–10721 (2009).
58. E. Gheerbrant, M. Amaghzaz, B. Bouya, F. Goussard, C. Letenneur, *Ocepeia* (middle Paleocene of Morocco): The oldest skull of an afrotherian mammal. *PLOS ONE* **9**, e89739 (2014).
59. E. Gheerbrant, A. Filippo, A. Schmitt, Convergence of afrotherian and laurasiatherian ungulate-like mammals: First morphological evidence from the Paleocene of Morocco. *PLOS ONE* **11**, e0157556 (2016).
60. N. B. Simmons, J. H. Geisler, Phylogenetic relationships of *Icaronycteris*, *Archaeonycteris*, *Hassianycteris*, and *Palaeochiropteryx* to extant bat lineages, with comments on the evolution of echolocation and foraging strategies in Microchiroptera. *Bull. Am. Museum Nat. Hist.* **235**, 1–182 (1998).
61. N. B. Simmons, K. L. Seymour, J. Habersetzer, G. F. Gunnell, Primitive Early Eocene bat from Wyoming and the evolution of flight and echolocation. *Nature* **451**, 818–821 (2008).
62. E. R. Price, A. Brun, E. Caviedes-Vidal, W. H. Karasov, Digestive adaptations of aerial lifestyles. *Phys. Chem. Chem. Phys.* **30**, 69–78 (2015).
63. L. D. Martin, in *Carnivore Behavior, Ecology, and Evolution*, J. L. Gittleman, Ed. (Springer, 1989), pp. 536–568.
64. M. Spaulding, J. J. Flynn, Phylogeny of the Carnivoramorpha: The impact of postcranial characters. *J. Syst. Palaeontol.* **10**, 653–677 (2012).
65. G. Storch, *Eomanis waldi*, ein Schuppentier aus dem Mittel-Eozän der 'Grube Messel' bei Darmstadt (Mammalia: Pholidota). *Senckenbergiana Lethaea* **59**, 503–529 (1978).
66. T. J. Gaudin, R. J. Emry, J. R. Wible, The phylogeny of living and extinct pangolins (Mammalia, Pholidota) and associated taxa: A morphology based analysis. *J. Mamm. Evol.* **16**, 235–305 (2009).
67. F. Solé, E. Gheerbrant, M. Amaghzaz, B. Bouya, Further evidence of the African antiquity of hyaenodontid ('Creodonta', Mammalia) evolution. *Zool. J. Linn. Soc.* **156**, 827–846 (2009).
68. K. D. Rose, L. T. Holbrook, R. S. Rana, K. Kumar, K. E. Jones, H. E. Ahrens, P. Missiaen, A. Sahni, T. Smith, Early Eocene fossils suggest that the mammalian order Perissodactyla originated in India. *Nat. Commun.* **5**, 5570 (2014).
69. L. N. Cooper, E. R. Seiffert, M. Clementz, S. I. Madar, S. Bajpai, S. T. Hussain, J. G. Thewissen, Anthracobunids from the middle eocene of India and Pakistan are stem perissodactyls. *PLOS ONE* **9**, e109232 (2014).
70. F. Welker, M. J. Collins, J. A. Thomas, M. Wadsley, S. Brace, E. Cappellini, S. T. Turvey, M. Reguero, J. N. Gelfo, A. Kramarz, J. Burger, J. Thomas-Oates, D. A. Ashford, P. D. Ashton, K. Rowsell, D. M. Porter, B. Kessler, R. Fischer, C. Baessmann, S. Kaspar, J. V. Olsen, P. Kiley, J. A. Elliott, C. D. Kelstrup, V. Mullin, M. Hofreiter, E. Willerslev, J.-J. Hublin, L. Orlando, I. Barnes, R. D. E. MacPhee, Ancient proteins resolve the evolutionary history of Darwin's South American ungulates. *Nature* **522**, 81–84 (2015).
71. M. Buckley, Ancient collagen reveals evolutionary history of the endemic South American "ungulates." *Proc. Biol. Sci.* **282**, 20142671 (2015).
72. M. Westbury, S. Baleka, A. Barlow, S. Hartmann, J. L. A. Paijmans, A. Kramarz, A. M. Forasiepi, M. Bond, J. N. Gelfo, M. A. Reguero, P. López-Mendoza, M. Taglioretti, F. Scaglia, A. Rinderknecht, W. Jones, F. Mena, G. Billet, C. de Muizon, J. L. Aguilar, R. D.E. MacPhee, M. Hofreiter, A mitogenomic timetree for Darwin's enigmatic South American mammal Macrauchenia patachonica. *Nat. Commun.* **8**, 15951 (2017).
73. J. G. M. Thewissen, S. T. Hussain, Postcranial Osteology of the most Primitive Artiodactyl: *Diacodexis pakistanensis* (Dichobunidae). *Anat. Histol. Embryol.* **19**, 37–48 (1990).
74. J. M. Theodor, in *The Evolution of Artiodactyls*, D. R. Prothero, S. E. Foss, Eds. (JHU Press, 2007), pp. 32–58.
75. Q. Li, X. Ni, An early Oligocene fossil demonstrates treeshrews are slowly evolving "living fossils." *Sci. Rep.* **6**, 18627 (2016).
76. L. Marivaux, L. Bocat, Y. Chaimanee, J.-J. Jaeger, B. Marandat, P. Srisuk, P. Tafforeau, C. Yamee, J.-L. Welcomme, Cynocephalid dermopterans from the Palaeogene of South Asia (Thailand, Myanmar and Pakistan): Systematic, evolutionary and palaeobiogeographic implications. *Zool. Scr.* **35**, 395–420 (2006).
77. M. Cartmill, Rethinking primate origins. *Science* **184**, 436–443 (1974).
78. B. A. Williams, R. F. Kay, E. C. Kirk, New perspectives on anthropoid origins. *Proc. Natl. Acad. Sci. U.S.A.* **107**, 4797–4804 (2010).
79. X. Ni, D. L. Gebo, M. Dagosto, J. Meng, P. Tafforeau, J. J. Flynn, K. C. Beard, The oldest known primate skeleton and early haplorhine evolution. *Nature* **498**, 60–64 (2013).

Acknowledgments: We thank three anonymous reviewers for comments on a previous draft of the manuscript and J. Johnson and the Broad Institute for providing access to xenarthran genome assemblies. **Funding:** This research was supported by an NSF Postdoctoral Research Fellowship in Biology (award no. 1523943; C.A.E.), an NSF Postdoctoral Fellow Research Opportunities in Europe award (C.A.E.), the France-Berkeley Fund (F.D. and M.W.N.), a European Research Council consolidator grant (ConvergeAnt no. 683257; F.D.), and the People Programme (Marie Curie Actions) of the European Union's Seventh Framework Programme (FP7/2007-2013) under REA grant agreement no. PCOFUND-GA-2013-609102, through the PRESTIGE programme coordinated by Campus France (C.A.E.). This is contribution ISEM 2018-049 of the Institut des Sciences de l'Evolution. **Author contributions:** C.A.E. conceived the study, collected the data, analyzed the data, and wrote the manuscript with input from M.W.N. and F.D. **Competing interests:** The authors declare that they have no competing interests. **Data and materials availability:** All data needed to evaluate the conclusions in the paper are present in the paper and/or the Supplementary Materials. Additional data related to this paper may be requested from the authors.

Submitted 2 December 2017
Accepted 30 March 2018
Published 16 May 2018
10.1126/sciadv.aar6478

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

Chitinase genes (*CHIA*s) provide genomic footprints of a post-Cretaceous dietary radiation in placental mammals

Christopher A. Emerling, Frédéric Delsuc and Michael W. Nachman

Sci Adv **4** (5), eaar6478.
DOI: 10.1126/sciadv.aar6478