

# Genomic Evidence for Large, Long-Lived Ancestors to Placental Mammals

J. Romiguier,<sup>1</sup> V. Ranwez,<sup>1,2</sup> E.J.P. Douzery,<sup>1</sup> and N. Galtier<sup>\*,1</sup>

<sup>1</sup>CNRS, Université Montpellier 2, UMR 5554, ISEM, Montpellier, France

<sup>2</sup>Montpellier SupAgro, UMR 1334, AGAP, Montpellier, France

\*Corresponding author: E-mail: nicolas.galtier@univ-montp2.fr.

Associate editor: Naruya Saitou

## Abstract

It is widely assumed that our mammalian ancestors, which lived in the Cretaceous era, were tiny animals that survived massive asteroid impacts in shelters and evolved into modern forms after dinosaurs went extinct, 65 Ma. The small size of most Mesozoic mammalian fossils essentially supports this view. Paleontology, however, is not conclusive regarding the ancestry of extant mammals, because Cretaceous and Paleocene fossils are not easily linked to modern lineages. Here, we use full-genome data to estimate the longevity and body mass of early placental mammals. Analyzing 36 fully sequenced mammalian genomes, we reconstruct two aspects of the ancestral genome dynamics, namely GC-content evolution and nonsynonymous over synonymous rate ratio. Linking these molecular evolutionary processes to life-history traits in modern species, we estimate that early placental mammals had a life span above 25 years and a body mass above 1 kg. This is similar to current primates, cetartiodactyls, or carnivores, but markedly different from mice or shrews, challenging the dominant view about mammalian origin and evolution. Our results imply that long-lived mammals existed in the Cretaceous era and were the most successful in evolution, opening new perspectives about the conditions for survival to the Cretaceous–Tertiary crisis.

**Key words:** phylogeny, GC-content, dN/dS ratio, GC-biased gene conversion, placentalia, fossils.

## Introduction

It is commonly assumed that early mammals were small creatures that only evolved into a variety of forms and sizes after the massive extinction of large reptiles, at the Cretaceous/Tertiary (KT) boundary, 65 Ma (Dawkins 2004; Feldhamer et al. 2007). This scenario is consistent with theoretical considerations: Cope's rule (Alroy 1998) states that current living lineages generally descend from small ancestors, because large forms have a short-term advantage but tend to be more prone to extinction in the long run. The hypothesis of a small ancestral size is also largely supported by the fossil record: most of the Cretaceous mammals are smaller than a few inches, whereas post-KT deposits include numerous large mammals (Luo 2007; Smith et al. 2010).

Paleontology, however, is not conclusive regarding the ancestry of extant mammals due to the difficulty of linking Cretaceous and Paleocene fossils to modern lineages (Archibald et al. 2001; Asher et al. 2005). Genomic data provide an attractive opportunity to characterize the ancestral features of extant species: genome dynamics is imprinted by species life-history traits (Nikolaev et al. 2007; Nabholz et al. 2008; Romiguier et al. 2010), and ancestral genome characters can be reconstructed by phylogenetic methods (Galtier et al. 1999; Blanchette et al. 2004; Boussau et al. 2009; Lartillot and Poujol 2011). If molecular evolution can be traced back to the last common ancestor of extant placentals, then we could potentially learn about its macroscopic characteristics, even though this ancestor is not physically observable because missing from the fossil record.

In mammals, two genomic variables are known to correlate with species life-history traits. First, species longevity and body mass influence the ratio of nonsynonymous (= amino acid changing, dN) to synonymous (dS) nucleotide substitution rates. It has been shown that large and long-lived species display a higher dN/dS ratio, on average, than small and short-lived ones, presumably because of the smaller average population sizes, and hence the less effective purifying selection, in long-lived animals (Nikolaev et al. 2007). Second, large species tend to show a lower GC3 (percentage of G and C at the third position of codons) than small species. This effect is supposed to be caused by GC-biased gene conversion (Duret and Galtier 2009), a mechanism by which a biased DNA-repair process favors G and C alleles during meiotic recombination. Because short-lived species experience a higher rate of meiosis per time unit, their genome shows a faster divergence in gene GC3 and an increase in average GC3 (Romiguier et al. 2010).

Here, we quantify the influence of species longevity on gene coding sequence evolutionary dynamics (dN/dS ratio and GC3 divergence) in modern placental mammals. Then we reconstruct ancestral gene sequences using nonhomogeneous phylogenetic models and estimate the dN/dS and GC3 dynamics in the deepest branches of the mammalian tree. On the basis of the existing correlation between genomic processes and traits, we estimate the maximal life span of early placental mammals. Our analysis suggests that the ancestors of living placentals had a longevity and body mass similar to current primates, cetartiodactyls, or carnivores but differed

markedly from mice or shrews, contradicting the prevailing view about mammalian origins and evolution.

## Materials and Methods

### Phylogeny and Divergence Dates

The phylogeny used in this study (fig. 1) was adapted from Meredith et al. (2011). Divergence dates were obtained from the TimeTree of Life database (<http://www.timetree.org>), which summarizes current paleontological and molecular knowledge. In Rodentia, the divergence dates proposed by this database are inconsistent with the tree topology assumed here. We used the following dates, consistent with recent literature on the subject (Springer et al. 2003; Huchon et al. 2007): *Mus/Rattus*: 16.4 My; *Mus/Dipodomys*: 70 My; *Mus/Cavia*: 73 My; and *Mus/Spermophilus*: 74 My.

### Orthologous Genes and Alignments

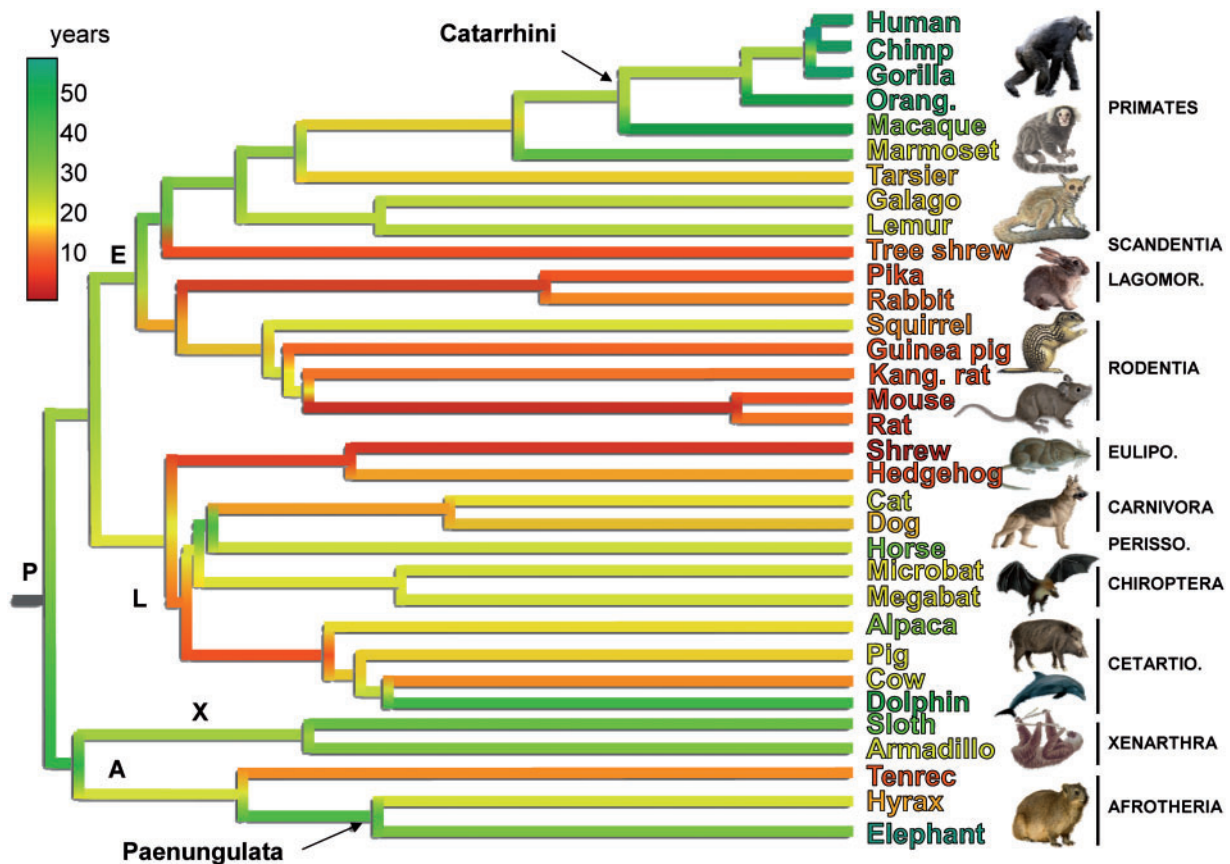
Aligned orthologous gene coding sequences were obtained from the ORTHOMAM database (Ranwez et al. 2007) version 6, which is based on ENSEMBL orthology annotations. Alignments were cleaned using the Gblocks program with default parameters (Castresana 2000).

### Life-History Traits

The maximum life span and body mass of placental mammals were retrieved from the AnAge database (de Magalhaes and Costa 2009), build 11. For each of the 33 placental species in the genomic data set, longevity was estimated by taking the mean longevity across all species of its family, for example, dog longevity was calculated as the mean of the longevitys of all documented Canidae species. The family average was considered here as an estimate of the long-term average, thus avoiding potential problems due to recent changes in longevity (e.g., in domesticated species). Human was excluded when calculating the Homnidae average because the maximal record in human (120 years) is irrelevant from an evolutionary viewpoint.

### Ancestral GC3 Estimation

For each of the 787 genes shared by all 33 species and the outgroups *Ornithorhynchus*, *Macropus*, and *Monodelphis*, we estimated ancestral GC3 at each node of the placental tree using Galtier and Gouy's nonhomogeneous model of sequence evolution (Galtier and Gouy 1998), as implemented in the Bio++ library (Dutheil et al. 2006; Dutheil and Boussau 2008), with branch lengths being separately estimated for each gene. This model is such that the GC content can



**Fig. 1.** Predicted evolution of longevity across the placental phylogeny. The maximal longevity of current species (family average) is indicated by the color of species names (red, short lived; green, long lived, see the top-left scale bar). Time-independent estimates of ancestral longevity are similarly indicated by the color of branches. P, most recent common ancestor (MRCA) of Placentalia; E, Euarchontoglires MRCA; L, Laurasiatheria MRCA; X, Xenarthra MRCA; A, Afrotheria MRCA.

fluctuate in time and across lineages, with each branch of the tree being assigned its own process, and its own equilibrium GC content. Outgroups are necessary for reliable estimation of GC3, especially, at the most basal nodes of the tree.

### GC3 Plots

For each pair of species, plots of gene GC3 in species 1 versus species 2 were drawn using all orthologous genes available for the considered pair. This number varied across species pairs, from 3,682 (*Choloepus/Ornithorhynchus*) to 11,526 (Human/Chimpanzee). Kendall's coefficient was used to measure the level of correlation of GC3 across genes between any two species or ancestors. To ensure comparability across species pairs, this coefficient was only calculated using the subset of 787 genes available in all 33 species of the data set. Importantly, in these correlation analyses, the observed GC3 values at tip nodes (extant species) were replaced with values predicted by the model, that is, expected GC3 values assuming that the estimated branch lengths and parameters of the nucleotide substitution model are true. This was necessary because, for a given gene, the model tends to smooth estimated GC3 across nodes of the tree and slightly underestimate the heterogeneity of GC3 across nodes. The comparison between internal and terminal pairs of nodes is therefore only relevant when estimated values are used for all nodes. Using actual GC3 values at tip nodes would lead to (perhaps implausibly) higher estimates of ancestral longevity.

### GC3-Based Ancestral Longevity Estimation

The time-corrected index of GC3 conservation that was used for a given pair of species was  $\gamma = -t/\log(\tau)$ , where  $t$  is the divergence time, and  $\tau$  is Kendall's correlation coefficient. This is based on the assumption of an exponential decay in the coefficient correlation over time, at a rate that is inversely proportional to the species longevity, here considered as constant over time and between the two considered lineages. Under this assumption, we have  $\tau = \exp(-\alpha t/l)$ , where  $l$  is the species longevity, and  $\alpha$  is a scaling factor. The  $\gamma$  index was calculated for the (Catarrhini ancestor and Paenungulata ancestor) pair and for 13 independent pairs of modern species of comparable divergence times. Pairs of modern species were chosen so as to maximize the number of within-order comparisons and avoid within-family comparisons. Among Afrotheria, the hyrax/tenrec pair was preferred to the hyrax/elephant pair to better match the assumption of equal longevity between species within a pair. The divergence time of the ancestral Catarrhini ancestor/Paenungulata ancestor pair was defined as  $([PI - Ca] + [PI - Pa])/2$ , where PI is the age of the placental ancestor (105 My), Ca is the age of the Catarrhini ancestor (30 My), and Pa is the age of the Paenungulata ancestor (61 My). Regression analyses were performed using R, similar to all the statistical analyses in this study.

### dN/dS Analysis

The branch-specific nonsynonymous/synonymous substitution rate ratio was calculated through substitution mapping,

in the spirit of Jobson et al. (2010). For each codon of each gene of the data set, synonymous and nonsynonymous changes were mapped onto the branches of the tree by probabilistic mapping (Dutheil et al. 2005), assuming Yang and Nielsen's (1998) model of coding sequence evolution. Then for each branch, the numbers of synonymous and nonsynonymous changes were summed across genes and their ratio calculated. This approach gives equal weight to codons, not to genes, and is computationally faster than maximum-likelihood approaches (Romiguier et al. 2012). Summing counts across genes avoids taking the ratio between small numbers, thus escaping the upward bias reported by Wolf et al. (2009) when estimating dN/dS from very short amounts of sequence divergence. Linear regression of species longevity on log-transformed branch-specific dN/dS ratio was performed using terminal branches only. Then, for each internal branch of the tree, ancestral longevity was predicted based on estimated dN/dS ratio.

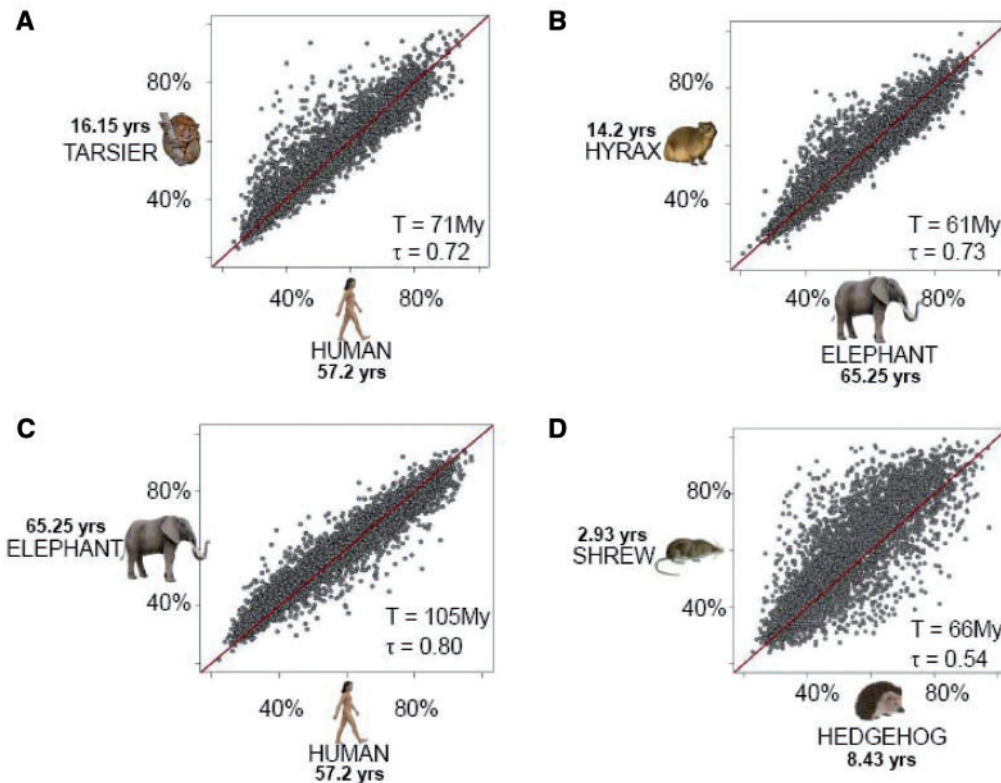
## Results

The phylogeny of the 33 placental mammalian species used in this study is shown in figure 1. These species, whose complete genomes have been sequenced, widely differ in their life-history traits. Here, we focused on maximum life span, represented by the color of species names in figure 1. Our aim is to estimate the maximum life span of the ancestors to these living species, here represented by the root and other basal nodes of the tree. We first introduce the main arguments of this study based on illustrative examples before presenting the whole statistical analyses.

### GC3 Plots

We examined the rate of genomic divergence using GC3 plots, in which the GC content at codon third positions of orthologous genes is compared between two species. Figure 2 shows GC3 plots for four species pairs, including human–tarsier (two primates, diverged ca. 71 Ma, fig. 2A), elephant–hyrax (two afrotherians, diverged ca. 61 Ma, fig. 2B), and human–elephant (fig. 2C), which diverged approximately 105 Ma, and whose common ancestor is that of all extant placentals. The first two plots illustrate the faster GC3 dynamics of short-lived species—many genes have undergone a substantial increase in GC3 in tarsier and hyrax lineages, when compared with human and elephant lineages, resulting in asymmetric plots. This is in agreement with the documented trend for GC3 enrichment in small mammals (Romiguier et al. 2010).

Notably, the human–elephant plot reveals a high level of GC3 conservation between these two species, such that the human/elephant correlation coefficient, here measured by Kendall's  $\tau = 0.80$ , is higher than those computed for the human/tarsier ( $\tau = 0.72$ ) and elephant/hyrax ( $\tau = 0.73$ ) pairs, despite the > 1.5-fold longer divergence period. The 105 My of human/elephant divergence was more conservative regarding GC3 than the 64 My of cow/pig divergence ( $\tau = 0.77$ ), and the 40 My of rabbit/pika divergence ( $\tau = 0.78$ ), not to speak of



**FIG. 2.** Typical GC3 plots between pairs of long-lived and/or short-lived species. The family average in maximal longevity is given for each species.

that of mouse/kangaroo rat (70 My,  $\tau = 0.57$ ) or shrew/hedgehog (66 My,  $\tau = 0.54$ , [fig. 2D](#)).

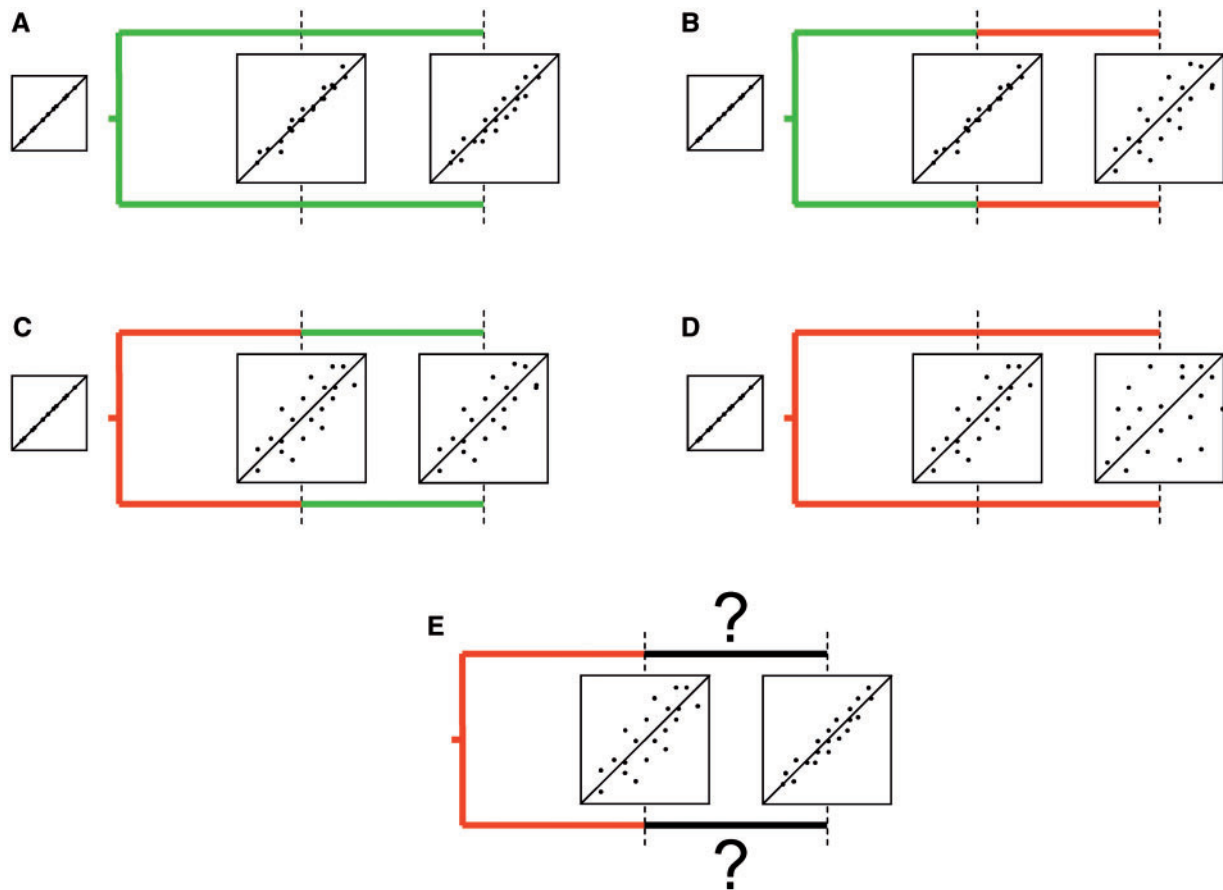
This result confirms that long-lived species evolve slowly regarding GC3 genes, while also strongly suggesting that all human and elephant ancestors were long-lived animals. If the first primates and afrotherians had had a short life span, then the GC3 of these ancestors should have diverged quickly during early placental evolution, and reached a state similar to [figure 2D](#), thus indelibly marking the human/elephant plot. We suggest that an ancestral GC3 pattern typical of short-lived taxa ([fig. 2D](#)) cannot lead to a modern GC3 pattern typical of long-lived species ([fig. 2C](#)) during the course of evolution: even if GC-biased gene conversion had stopped, the random accumulation of AT→GC and GC→AT mutations through genetic drift is not expected to drive individual gene GC3 back to similar values in two independently evolving lineages (see [fig. 3](#) for illustration). Note that high levels of GC3 conservation were also found between human and the long-lived sloth ( $\tau = 0.78$ ), and between elephant and sloth ( $\tau = 0.75$ ), extending the rationale to the xenarthran ancestral branch.

### Ancestral Longevity Estimation

To quantify the early rate of GC3 divergence, we reconstructed ancestral gene GC3 using the maximum-likelihood method and a nonhomogeneous model of sequence evolution. From these reconstructions, we built GC3 plots between internal nodes of the tree and measured levels of GC3 conservation between these ancestors. We focused on divergence

between the most recent common ancestor (MRCA) to extant Catarrhini (Old World monkeys and apes) and the MRCA to extant Paenungulata (elephants, sirenians, and hyraxes; [fig. 1](#)). These two MRCAs diverged over an average 59.5 My since their common placental ancestor (see Materials and Methods). Their level of GC3 correlation is  $\tau = 0.84$ . We compared this number to levels of GC3 correlation calculated between independent pairs of extant species with a comparable divergence time ( $59.5 \pm 30$  My). We found that among 13 such species pairs, only the long-lived chimpanzee/*Macaca* ( $\tau = 0.93$ , 44 years) and cow/dolphin ( $\tau = 0.85$ , 36 years) pairs showed levels of GC3 correlation similar to, or higher than, the Paenungulata/Catarrhini ancestral pair ([supplementary table S1, Supplementary Material](#) online). All 11 species pairs with an average longevity lower than 30 years diverged more rapidly than Paenungulata/Catarrhini regarding GC3.

With this data set, a linear regression of longevity on a time-corrected GC3 conservation level (see Materials and Methods) revealed a strong, positive correlation ( $r^2 = 0.93$ ) and predicted that the average longevity during early Catarrhini/Paenungulata divergence was  $33.3 \pm 7.6$  years ([fig. 4](#)). Among the 897 placental species with a documented maximal longevity, 174 (19%) are within this range, and 65 (7%) are more long lived. The list of modern species matching the  $33.3 \pm 7.6$  prediction interval includes 76 primates, 38 cetartiodactyls, and 30 carnivores but only five rodents (mostly porcupines), whereas rodents represent  $\sim 40\%$  of placental species overall. In this list of 144 species, 95% of the body mass distribution is within (0.65–18 kg) in arboreal



**Fig. 3.** Predicted behavior of GC3 plot under four scenarios of life span evolution. Two periods of time, early and late, are considered. Green branches are for long-lived lineages and red branches for short-lived ones. GC3 plots between the two diverging lineages are represented at three time points: initial stage (no divergence, correlation coefficient = 1), end of the early period, and end of the late period. In scenario A, an ancestral long life span is kept throughout the two periods. In this case, gene GC3 evolution is slow, and a strong level of correlation is to be expected from GC3 plots. Scenario B represents convergent evolution from a long ancestral life span to a short derived life span. Here, the GC3 plot is only little perturbed during the early period, like in (A), but a fast decay of correlation coefficient is expected during the late period. In scenarios C and D, the ancestral life span was short and was either kept throughout (D) or convergently evolved to a derived long life span (C). In both cases, a low correlation coefficient is expected for GC3 plots, if only because the intermediate GC3 plot is weakly correlated. Under scenario C, convergent evolution toward a long life span is not expected to result in an increased correlation coefficient but rather in a freeze of GC3 plot. Reverse evolution toward a higher correlation coefficient, represented in panel E, can only occur under an implausible scenario in which a large number of initially diverged genes would convergently reach a common equilibrium GC3 in the two lineages.

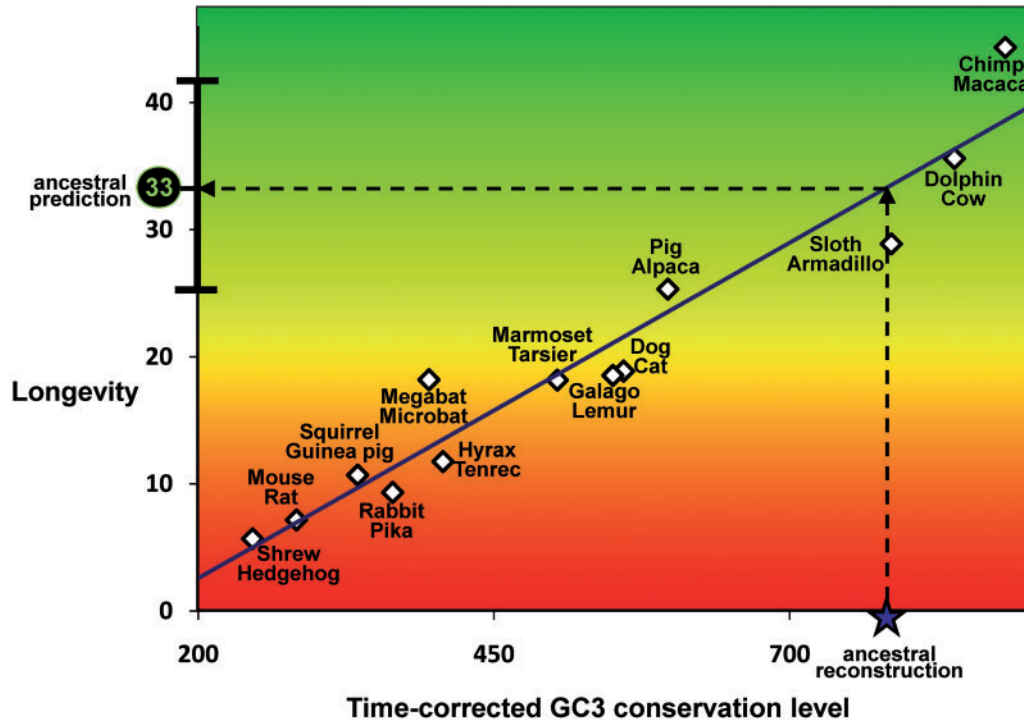
species and within (3.75–800 kg) in terrestrial species—arboreal mammals are known to be longer lived than terrestrial mammals, for a given body mass (Shattuck and Williams 2010). Notably, all members of the former Lipotyphla group (“insectivores”: Eulipotyphla, Afrosoricida, Macroscelidae, Scandentia, and Dermoptera), often considered as having retained placental plesiomorphies (Madsen et al. 2001; Asher 2005), have a maximal longevity of less than 19 years (mean = 7 years), that is, well outside the prediction interval.

### dN/dS Analysis

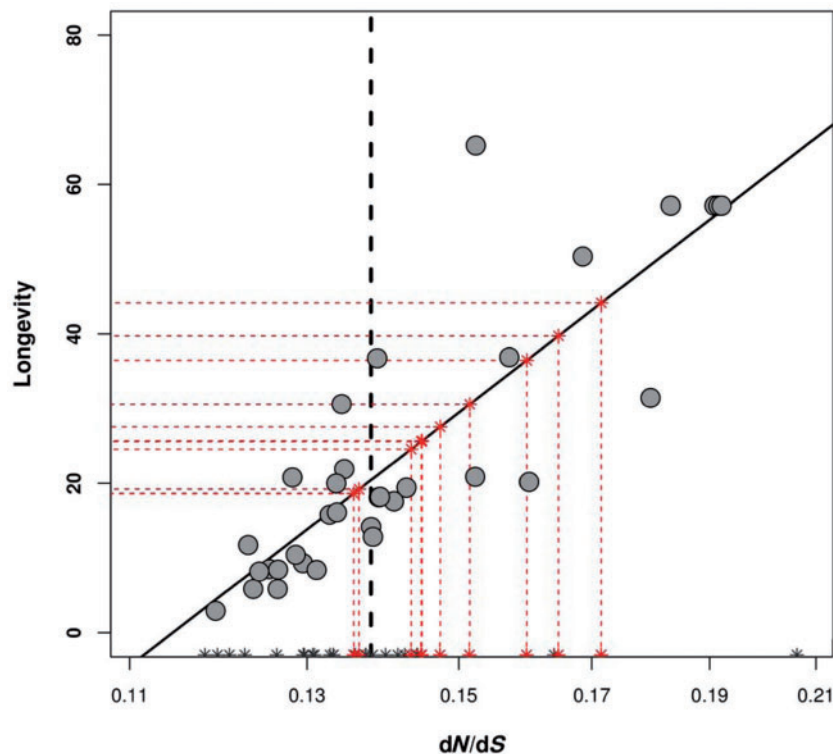
The above calculations rely on prior knowledge of the age of internal nodes in the placental tree, here taken as errorless data. Although somewhat consensual in recent molecular phylogenetic literature (Meredith et al. 2011), divergence dates between mammals are obviously uncertain (Murphy and Ezirik 2009). Even though our GC3-based analysis is

only dependent on relative, not absolute, divergence dates, confirming these results in a time-independent manner would appear desirable. To achieve this aim, we focused on terminal branches of the tree and measured the branch-specific dN/dS ratio. This time-independent statistics is known to be correlated to species life-history traits: dN/dS is higher, on average, in long-lived species than in short-lived ones (see earlier). We here confirmed this relationship: a significant ( $r^2 = 0.86$ ,  $P$  value < 0.0001), positive correlation was obtained between log-transformed terminal branch dN/dS and species longevity (fig. 5). Then, we measured the dN/dS ratio in internal branches of the placental tree. On the basis of the linear regression of figure 4, we predicted the average longevity in ancestral placental lineages, as shown by colors in figure 1.

Although the prediction interval for any specific branch was wide, this analysis essentially corroborated the GC3-based estimates. Among the 10 internal branches separating the



**FIG. 4.** Linear regression of longevity on time-corrected GC3 conservation. Each dot is for a pair of species. x axis is the time-corrected index of GC3 conservation  $-t/\log(\tau)$ , where  $t$  is the divergence time between the two considered species, and  $\tau$  is Kendall's correlation coefficient calculated from GC3 plots. y axis is the average maximal longevity of the considered pair. The star marks the position of the Catarrhini MRCA/Paenungulata MRCA pair in x axis. The estimated longevity and prediction interval for this pair of ancestors is shown on y axis (Supplementary Material online).



**FIG. 5.** Linear regression of longevity on dN/dS ratio. Each dot is for a terminal branch of the placental tree. x axis: dN/dS ratio (log scale). y axis: species maximal longevity. Vertical dotted line: median dN/dS ratio across terminal branches. Stars in x axis represent the estimated dN/dS ratio in internal branches of the tree. Red stars correspond to the 10 internal branches from the path connecting the Catarrhini MRCA to the Paenungulata MRCA. Estimated ancestral longevity for these 10 branches are shown by projection onto the y axis. Raw data available from supplementary table S2, Supplementary Material online.

Catarrhini MRCA from the Paenungulata one (red stars in fig. 5), the predicted longevity ranged from 19 to 45 years and averaged 29.0 years, confirming the results of figure 4. This analysis suggests that the MRCAs of Afrotheria, Xenarthra, Laurasiatheria, and Euarchontoglires—the four major placental superorders (indicated by their initial in fig. 1)—were long lived, whereas the MRCAs of Eulipotyphla, Lagomorpha, and, surprisingly, Cetartiodactyla, were short-lived animals. In this analysis, the average dN/dS ratio of internal branches (0.141) was almost identical to the average dN/dS ratio of terminal branches (0.140), and no correlation was observed between the estimated dN/dS ratio and the phylogenetic depth of an internal branch, here defined as the age of its bottom node ( $r^2 = -0.09$ ,  $P$  value = 0.6). These results suggest that our inferences are not affected by a systematic upward bias of dN/dS estimates in ancient branches, as could be expected in case of substitutional saturation or alignment errors in deeply branching lineages.

### Controls

Additional control analyses were achieved to check the robustness of ancestral longevity estimates. The impact of missing data was assessed by removing from the data set alignments that included at least one sequence containing a proportion of gaps above some threshold. The results were only slightly affected. When the threshold was set at 50%, only 205 genes were retained, and time-dependent predicted ancestral longevity for the Catarrhini/Paenungulata MRCA was  $32.8 \pm 7.8$ . The impact of CpG dinucleotides, which are hypermutable toward TpG or CpA in mammals, was assessed by removing all CpG-affected codon sites from the alignments. A codon site was considered CpA affected as soon as any position in the codon was involved in a CpG, a TpG, or a CpA doublet in >50% of the sequences. The results were robust with respect to this removal (time-dependent predicted ancestral longevity for the Catarrhini/Paenungulata MRCA:  $29.7 \pm 9.8$ ). This control is especially needed knowing that a shift in CpG substitution rate at the time of early placental divergence has been reported (Arndt et al. 2003). Finally, qualitatively similar results were obtained when we used the age of female sexual maturity, rather than maximal life span, as a predictor of GC3 dynamics.

### Discussion

Here, we used a phylogenetic approach to reconstruct ancestral genome dynamics and estimate ancestral life-history traits, based on the evidence in modern taxa that species traits influence molecular evolution. This analysis, which benefits from the strength of nonhomogeneous models of sequence evolution, illustrates the power of genomic data to unravel species evolutionary history. The link between species characteristics and GC-content dynamics has been previously observed (Romiguier et al. 2010; Nabholz et al. 2011), but its implications regarding ancestral trait reconstruction have never been considered so far.

The main signal we extract from the data is related to the much higher evolutionary instability of gene GC3 in

short-lived than in long-lived species. We suggest that the high level of conservation of gene GC3 during early placental divergence is only compatible with a long ancestral life span, because putative short-lived ancestors would have left indelible mark in modern GC3 plots (fig. 1). This was formalized in our method by the assumption of an exponential decay of correlation coefficient in time, which neglects possible instances of convergent evolution in gene GC3. The very strong correlation we observed between time-corrected GC3 divergence level and species longevity (fig. 4) suggests that this assumption is largely met by the data and that GC3 plots correctly capture the evolutionary dynamics of longevity.

This result is reinforced by the analysis of dN/dS variations between lineages, an independent source of information that corroborated the GC3-based inferences. Note that the agreement between GC3 plots and dN/dS is not limited to the deepest branches of the tree. The report of a low dN/dS ratio in the ancestral Cetartiodactyl branch, for instance, is consistent with the relatively low level of GC3 correlation between long-lived members of this group and other long-lived mammals. The GC3 correlation coefficient between alpaca and horse (0.72), for instance, is lower than between the more distantly related horse and *Macaca* (0.8) or horse and sloth (0.75)—to talk only about species of maximum life span  $\sim 30$  years. So the short-lived predicted ancestral Cetartiodactyl has apparently left an indelible mark in the GC3 plots of this group, in agreement with the rationale of figure 3.

Our study of past genome dynamics predicts that the maximum life span of the ancestors to modern placentals was more than 20 years and probably  $\sim 30$  years. According to our results, these early placental mammals were either large-sized terrestrial or medium-sized arboreal animals. Ancestral mouse-like or shrew-like life-history traits are excluded, questioning one of the most frequently told stories in evolutionary biology. Our analysis suggests that very small size in placental mammals is a derived state, which evolved several times independently (e.g., in Rodentia, Chiroptera, Eulipotyphla, and Macroscelidea), in contradiction with Cope's rule (Monroe and Bokma 2010). We note that the hypothesis of a long ancestral life span for placentals appears consistent with the ecological theory. Evolution of the placenta (and more generally viviparity) means that reproduction is delayed to the benefit of juvenile survival, that is, increased investment in parental care, which could appear unexpected in a short-lived, r-strategy species (Stearns 1992). This comment is only loosely relevant to our results, though, because the evolution of the placenta may have preceded to a non-negligible extent of the diversification of the MRCA of modern placental superorders.

From a paleontological standpoint, our results call for re-examination of large or arboreal mammalian taxa as potential stem groups to extant placental superorders. The Cenozoic fossil record includes relatively large mammals, such as the late Cretaceous, dinosaur-eating *Repenomamus giganteus* (15 kg) (Hu et al. 2005), which is indicative of the existence of potentially long-lived mammals before the KT crisis.

This species, however, is only distantly related to placentals. The 81 Cretaceous eutherian fossils documented in the Paleobiology database are much smaller, their body mass being below 500 g. If the MRCA of current placentals lived in the Cretaceous, as suggested by molecular dating (Meredith et al. 2011), then our results would imply that a diversified fauna of large Cretaceous placentals has been missed by paleontologists, enlarging further the gap between molecules and morphology regarding early mammalian evolution. Our results, if trusted, also have implications regarding divergence date estimates. If early placentals were long lived, and therefore experienced a lower substitution rate than predicted by clock-relaxed models, then the divergence date between the main placental lineages would be even older than typical molecular estimates, reinforcing the discrepancy that exists between molecular and fossil time divergences. This is illustrated by the difference between our study and that of Steiper and Seifert (2012), who predicted a much smaller ancestral primate than we do when they tried to reconcile molecular divergence rates with the fossil record.

However, let us recall that our study is only weakly dependent on absolute divergence dates. The lack of documented stem Placentalia Cretaceous fossils is problematic vis-à-vis molecular dates irrespective of our results. We also note that an ancestral body mass of 1 kg, or even less, is compatible with our analysis if the placental MRCA was arboreal, and some of the best preserved Cretaceous eutherian fossils were arboreal (Ji et al. 2002; Goswami et al. 2011). Finally, we emphasize that many of our ancestral inferences in this study are consistent with the fossil record. Ancestral proboscidiids, for instance, are predicted to be much smaller/shorter lived than extant elephants by our analysis, in agreement with the fossil record for this group (Gheerbrant et al. 1996; Gheerbrant 2009). The same remark applies to the horse lineage, in which our relatively low ancestral longevity estimate is compatible with the relatively small size of early equid fossils (Froehlich 2002). The relatively long ancestral longevity predicted in the squirrel lineage is consistent with the large size and arboreal life style of their most primitive known fossils, *Paramys* and *Ailuravus* (Rose 2005). Finally, the cetartiodactyl MRCA is predicted to be very short lived in our analysis, which is consistent with the rabbit size of the oldest-known cetartiodactyl, *Diacodexis* (Rose 1982). In this group, our method predicts an ancestral longevity that falls outside the longevity range of extant taxa and still makes sense from a paleontological point of view.

If the molecular time scale is trusted, then our study would suggest that long-lived Mesozoic mammals (medium-sized arboreal or large terrestrial) not only have existed but also were successful during evolution, so that modern placental mammals, including the smallest ones, descend from larger ancestors. We note that the generally long-lived birds, crocodiles, and turtles have also passed the KT crisis and suggest that long life span species might benefit from an increased ability to survive sudden, several year-long episodes of food shortage and reduced growth rate.

## Supplementary Material

Supplementary Material, tables S1 and S2 are available at *Molecular Biology and Evolution* online (<http://www.mbe.oxfordjournals.org/>).

## Acknowledgments

The authors thank Vincent Cahais for technical help and Nicolas Lartillot for thoughtful discussions on the manuscript. This work was supported by Agence Nationale de la Recherche grants ANR-08-GENM-036-01 “CoGeBi” and ANR-10-BINF-01-02 “Ancestrome” and by European Research Council Advanced Grant ERC 232971 “PopPhyl.” This is contribution number ISEM-2012-109.

## References

- Alroy J. 1998. Cope's rule and the dynamics of body mass evolution in North American fossil mammals. *Science* 280:731–734.
- Archibald JD, Averianov AO, Ekdale EG. 2001. Late Cretaceous relatives of rabbits, rodents, and other extant eutherian mammals. *Nature* 414:62–65.
- Arndt PF, Petrov DA, Hwa T. 2003. Distinct changes of genomic biases in nucleotide substitution at the time of Mammalian radiation. *Mol Biol Evol.* 20:1887–1896.
- Asher RJ. 2005. Insectivoran-grade placentals. In: Rose KD, Archibald JD, editors. *The rise of placental mammals*. Baltimore (MD): Johns Hopkins University Press. p. 50–70.
- Asher RJ, Meng J, Wible JR, McKenna MC, Rougier GW, Dashzeveg D, Novacek MJ. 2005. Stem Lagomorpha and the antiquity of Glires. *Science* 307:1091–1094.
- Blanchette M, Green ED, Miller W, Haussler D. 2004. Reconstructing large regions of an ancestral mammalian genome in silico. *Genome Res.* 14:2412–2423.
- Boussau B, Blanquart S, Necsulea A, Lartillot N, Gouy M. 2009. Parallel adaptations to high temperatures in the Archaeal eon. *Nature* 456: 942–945.
- Castresana J. 2000. Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. *Mol Biol Evol.* 17:540.
- Dawkins R. 2004. *The ancestor's tale: a pilgrimage to the dawn of evolution*. New York: Houghton Mifflin.
- de Magalhaes JP, Costa J. 2009. A database of vertebrate longevity records and their relation to other life-history traits. *J Evol Biol.* 22: 1770–1774.
- Duret L, Galtier N. 2009. Biased gene conversion and the evolution of mammalian genomic landscapes. *Annu Rev. Genomics Hum Genet.* 10:285–311.
- Dutheil J, Boussau B. 2008. Non-homogeneous models of sequence evolution in the Bio++ suite of libraries and programs. *BMC Evol Biol.* 8:255.
- Dutheil J, Gaillard S, Bazin E, Glémin S, Ranwez V, Galtier N, Belkhir K. 2006. Bio++, a set of C++ libraries for sequence analysis, phylogenetics, population genetics, and molecular evolution. *BMC Bioinformatics* 7:188.
- Dutheil J, Pupko T, Jean-Marie A, Galtier N. 2005. A model-based approach for detecting coevolving positions in a molecule. *Mol Biol Evol.* 22:1919–1928.
- Feldhamer GA, Drickamer LC, Vessey VC, Merritt JF, Krajewski C. 2007. *Mammalogy: adaptation, diversity, ecology*. 3rd ed. Baltimore (MD): John Hopkins University Press.



- Froehlich DJ. 2002. Quo vadis eohippus? The systematics and taxonomy of the early Eocene equids (Perissodactyla). *Zool J Linnean Soc.* 134: 141–256.
- Galtier N, Gouy M. 1998. Inferring pattern and process: maximum likelihood implementation of a non-homogeneous model of DNA sequence evolution for phylogenetic analysis. *Mol Biol Evol.* 15: 871–879.
- Galtier N, Tourasse NJ, Gouy M. 1999. A nonhyperthermophilic common ancestor to extant life forms. *Science* 283:220–221.
- Gheerbrant E. 2009. Paleocene emergence of elephant relatives and the rapid radiation of African ungulates. *Proc Natl Acad Sci U S A.* 106: 10717–10721.
- Gheerbrant E, Sudre J, Cappetta H. 1996. A Palaeocene proboscidean from Morocco. *Nature* 383:68–70.
- Goswami A, Prasad GV, Upchurch P, Boyer DM, Seiffert ER, Verma O, Gheerbrant E, Flynn JJ. 2011. A radiation of arboreal basal eutherian mammals beginning in the late Cretaceous of India. *Proc Natl Acad Sci U S A.* 108:16333–16338.
- Hu Y, Meng J, Wang Y, Li C. 2005. Large Mesozoic mammals fed on young dinosaurs. *Nature* 433:149–152.
- Huchon D, Chevret P, Jordan U, Kilpatrick CW, Ranwez V, Jenkins PD, Brosius J, Schmitz J. 2007. Multiple molecular evidences for a living mammalian fossil. *Proc Natl Acad Sci U S A.* 104:7495–7499.
- Ji Q, Luo ZX, Yuan CX, Wible JR, Zhang JP, Georgi JA. 2002. The earliest known eutherian mammal. *Nature* 416:816–822.
- Jobson RW, Nabholz B, Galtier N. 2010. An evolutionary genome scan reveals the targets for longevity-related natural selection in mammals. *Mol Biol Evol.* 27:840–847.
- Lartillot N, Poujol R. 2011. A phylogenetic model for investigating correlated evolution of substitution rates and continuous phenotypic characters. *Mol Biol Evol.* 28:729–744.
- Luo ZX. 2007. Transformation and diversification in early mammal evolution. *Nature* 450:1011–1009.
- Madsen O, Scally M, Douady CJ, Kao DJ, DeBry RW, Adkins R, Amrine HM, Stanhope MJ, de Jong WW, Springer MS. 2001. Parallel adaptive radiations in two major clades of placental mammals. *Nature* 409: 610–614.
- Meredith RW, Janečka JE, Gatesy J, et al. (22 co-authors). 2011. Impacts of the Cretaceous terrestrial revolution and KPg extinction on mammal diversification. *Science* 334:521–524.
- Monroe MJ, Bokma F. 2010. Little evidence for Cope's rule from Bayesian phylogenetic analysis of extant mammals. *J Evol Biol.* 23:2017–2021.
- Murphy WJ, Eizirik E. 2009. Placental mammals (*Eutheria*). The timetree of life. In: Hedges SH, Kumar S, editors. New York: Oxford University Press. p. 471–474.
- Nabholz B, Glémin S, Galtier N. 2008. Strong variations of mitochondrial mutation rate across mammals—the longevity hypothesis. *Mol Biol Evol.* 25:120–130.
- Nabholz B, Künstner A, Wang R, Jarvis E, Ellegren H. 2011. Dynamic evolution of base composition: causes and consequences in avian phylogenomics. *Mol Biol Evol.* 28:2197–2210.
- Nikolaev SI, Montoya-Burgos JI, Popadin K, Parand L, Margulies EH. 2007. Life-history traits drive the evolutionary rates of mammalian coding and noncoding genomic elements. *Proc Natl Acad Sci U S A.* 104: 20443–20448.
- Ranwez V, Delsuc F, Ranwez S, Belkhir K, Tilak MK, Douzery EJ. 2007. OrthoMaM: a database of orthologous genomic markers for placental mammal phylogenetics. *BMC Evol Biol.* 7:241.
- Romiguier J, Figuet E, Galtier N, Douzery EJ, Boussau B, Dutheil JY, Ranwez V. 2012. Fast and robust characterization of time-heterogeneous sequence evolutionary processes using substitution mapping. *PLoS One* 7:e33852.
- Romiguier J, Ranwez V, Douzery E, Galtier N. 2010. Contrasting GC-content dynamics across 33 mammalian genomes: relationship with life history traits and chromosome size. *Genome Res.* 20: 1001–1009.
- Rose KD. 1982. Skeleton of diacodexis, oldest known artiodactyl. *Science* 216:621–623.
- Rose KD. 2005. The beginning of the age of mammals. Baltimore (MD): Johns Hopkins University Press.
- Shattuck MR, Williams SA. 2010. Arboreality has allowed for the evolution of increased longevity in mammals. *Proc Natl Acad Sci U S A.* 107:4635–4639.
- Smith FA, Boyer AG, Brown JH, et al. (20 co-authors). 2010. The evolution of maximum body size of terrestrial mammals. *Science* 330: 1216–1219.
- Springer MS, Murphy WJ, Eizirik E, O'Brien SJ. 2003. Placental mammal diversification and the Cretaceous-Tertiary boundary. *Proc Natl Acad Sci U S A.* 100:1056–1061.
- Stearns SC. 1992. The evolution of life histories. New York (NY): Oxford University Press.
- Steiper ME, Seiffert ER. 2012. Evidence for a convergent slowdown in primate molecular rates and its implications for the timing of early primate evolution. *Proc Natl Acad Sci U S A.* 109:6006–6011.
- Wolf JB, Künstner A, Nam K, Jakobsson M, Ellegren H. 2009. Nonlinear dynamics of nonsynonymous (dN) and synonymous (dS) substitution rates affects inference of selection. *Genome Biol Evol.* 1:308–319.
- Yang Z, Nielsen R. 1998. Synonymous and nonsynonymous rate variation in nuclear genes of mammals. *J Mol Evol.* 46:409–418.