" Sorrento Special Issue of Journal of Molecular Evolution "

LOCAL MOLECULAR CLOCKS IN THREE NUCLEAR GENES: DIVERGENCE TIMES FOR RODENTS AND OTHER MAMMALS, AND INCOMPATIBILITY AMONG FOSSIL CALIBRATIONS.

Emmanuel J. P. DOUZERY [1, *], Frédéric DELSUC [1], Michael J. STANHOPE [2, 3] & Dorothée HUCHON [1, 4]

[1] Laboratoire de Paléontologie, Paléobiologie et Phylogénie - CC064
Institut des Sciences de l'Evolution UMR 5554 / CNRS
Université Montpellier II; Place E. Bataillon
34 095 Montpellier Cedex 05 — France

[2] Biology and Biochemistry, Queen's University of Belfast, 97 Lisburn Rd., Belfast BT9 7BL, UK.

[3] Present address: Bioinformatics, GlaxoSmithKline, UP1345, 1250 South Collegeville Road, Collegeville PA 19426, USA.

(4) Florida State University, Biological Science Department, Tallahassee FL 32306-1100, USA.

[*] Corresponding author: Tel = 33 4 67 14 48 63 / Fax = 33 4 67 14 36 10 / e-mail = douzery@isem.univ-montp2.fr

ABSTRACT

Reconstructing the chronology of mammalian evolution is a debated issue between molecule- and fossil-based inferences. A methodological limitation of molecules is the evolutionary rate variation among lineages, precluding the application of the global molecular clock. We considered 2422 first and second codon positions of the combined ADRA2B, IRBP, and vWF nuclear genes for a well-documented set of placentals including an extensive sampling of rodents. Using seven independent calibration points and a maximum-likelihood framework, we evaluated whether molecular and paleontological estimates of mammalian divergence dates may be reconciled by the local molecular clocks approach allowing local constancy of substitution rates with variations at larger phylogenetic scales. To handle the difficulty of choosing among all possible rate assignments for various lineages, local molecular clocks were based on the results of branch-length and two-cluster tests. Extensive lineage-specific variation of evolutionary rates was detected, even among rodents. Cross-calibrations indicated some incompatibilities between divergence dates based on different paleontological references. To decrease the impact of a single calibration point, estimates derived from independent calibrations displaying only slight reciprocal incompatibility were averaged. The divergence dates inferred for the split between mice and rats (~13-19 Myr) was younger than previously published molecular estimates. The most recent common ancestors of rodents, primates and rodents, boreoeutherians, and placentals were estimated to be respectively ~60, 70, 75, and 78 Myr old. Global clocks, local clocks and quartet dating analyses suggested a Late Cretaceous origin of the crown placental clades followed by a Tertiary radiation of some placental orders like rodents.

Keywords: Local molecular clock — maximum likelihood — divergence times — phylogeny — fossil record — evolutionary rates — nuclear genes — mammals — rodents.

INTRODUCTION

With the exponential growth of DNA sequence data for a variety of organisms, it becomes increasingly attractive to place an evolutionary perspective on the results obtained at the molecular level with those obtained from morphology and paleontology. However, reconstructing the chronology of mammalian diversification has been a matter of controversy. Paleontological and molecular approaches disagree on both the tempo and mode of the early placental radiation (Alroy 1999; Benton 1999; Easteal 1999). Fossil evidence indicates an explosive radiation of placental orders in the Paleocene some 65 million years (Myr) ago, just after the Cretaceous-Tertiary (K-T) limit. On the contrary, molecular clock approaches analyzing both mitochondrial and nuclear data support a Cretaceous origin and diversification of placentals.

Discrepancies between clocks and rocks might have their origin in the incompleteness of the fossil record or possibly in methodological shortcomings associated with the molecular approach. Several sources of error might affect the accuracy of divergence dates deduced from DNA and protein data: the number of sites analyzed (Bromham et al. 2000), the number of species considered and the topology of the reference phylogeny (Sanderson and Doyle 2001), and the molecular dating method (Rodriguez-Trelles et al. 2002). However, the most confounding and widespread cause is likely to be the variation of evolutionary rates, precluding the use of the global molecular clock concept (Zuckerkandl and Pauling 1965). In eukaryotic genes, lineage specific substitution rate variations have been documented for protists (Philippe and Germot 2000), plants (Bousquet et al. 1992), fungi (Moncalvo et al. 2000), and various animals such as insects (Huelsenbeck 1998), cyclostomes (Mallatt and Sullivan 1998), or mammals (Huchon et al. 2000). Rate variations have also been detected for mitochondrial genomes among mammals (Gissi et al. 2000). Two mammalian orders have long been the subject of intensive studies regarding their evolutionary rates: rodents and

primates. For some molecular markers, rodents have been shown to evolve faster than other placentals (Wu and Li 1985; Adkins et al. 2001), and to display within-order rate heterogeneities (Huchon et al. 2000). Primates are also characterized by different rates of molecular evolution between lineages, namely faster rates in anthropoids (Adkins and Honeycutt 1994; Andrews et al. 1998; Andrews and Easteal 2000; Liu et al. 2001).

Because changes in evolutionary rate are widespread, a "global"-clock tree cannot always be easily obtained by the removal of fast- or slow- evolving taxa without sacrificing taxon sampling (Takezaki et al. 1995; Huchon et al. 2000; Murphy et al. 2001b). This is the reason why a so-called local molecular clock approach has been recently proposed by Yoder and Yang (2000). These authors hypothesize local homogeneity of evolutionary rates with variations at larger phylogenetic scales. In a maximum likelihood (ML) framework, Yoder and Yang (2000) attribute independent substitution rates to some fast-evolving (or slow-evolving) lineages while assuming rate constancy in the others. Molecular dates are then obtained by converting branch lengths on the ML tree to time since divergence using one calibration point. This method has already been applied to derive a molecular time scale for the evolution of primates (Yoder and Yang 2000), rodents (Huchon and Douzery 2001) and xenarthrans (Delsuc et al. 2001).

Among placentals, rodents are good candidates to test the properties of the local clock approach. Indeed, for some molecular markers, rodents have been shown to display extensive lineage-specific substitution rate heterogeneities in nuclear genes (Huchon et al. 2000; Adkins et al. 2001), even between closely related taxa (e.g., Fieldhouse et al. 1997; Michaux and Catzeflis 2000; Michaux et al. 2001; Rowe and Honeycutt 2002). Additionally, several molecular studies, based on both mitochondrial and nuclear genes, have provided a reasonably good picture of phylogenetic relationships of rodents and other placentals (Nedbal et al. 1994, 1996; Catzeflis et al. 1995; Huchon et al. 1999, 2000, 2002; Adkins et al. 2001; Huchon and Douzery 2001; DeBry and Sagel 2001; Murphy et al. 2001a). These studies have

demonstrated that rodents share an unique common ancestor, that they are composed of three major clades of mouse-like, squirrel-like, and Guinea-pig-like animals, and that their closest extant relatives are lagomorphs, followed by primates, flying-lemurs and tree-shrews. This relatively well resolved phylogeny allows its use as a molecular scaffold for dating purposes. Furthermore, a fairly good fossil record is available for rodents, at the order (e.g., Hartenberger 1998) and family (e.g., Vucetich et al. 1999) levels, providing thus a number of potentially reliable paleontological calibration points. Finally, the timing of two particular events of rodent evolution has been the matter of much controversy and represent important examples of the disagreements between the fossil and molecular dating approaches. On the one hand, the timing of rodent origins is controversial. The divergence between sciurognaths (squirrel-like rodents) and hystricognaths (Guinea-pig-like rodents) has been estimated to be 75-125 million years old by molecular dating (Janke et al. 1997; Kumar and Hedges 1998; Cao et al. 2000; Adkins et al. 2001) whereas fossils support a rodent radiation 55 Myr ago (Hartenberger 1998). On the other hand, the age of the split between mouse and rat has been highly debated. Whereas the fossil record suggests a Mus / Rattus split around 14 Myr (Jacobs and Downs 1994), global molecular clocks applied to DNA or protein sequence data provide mean divergence dates of at least 33 Myr (Nei et al. 2001), 35 Myr (Janke et al. 1994), 41 Myr (Kumar and Hedges 1998), or 42 Myr (Huchon et al. 2000). A noticeably younger molecular estimate is the 23 Myr figure of Adkins et al. (2001) based on the analysis of the growth hormone receptor gene.

The purpose of this paper is to re-evaluate the discrepancies between molecular and paleontological estimates of mammalian divergence dates, using the local molecular clock approach of Yoder and Yang (2000) to accommodate lineage-specific variations in evolutionary rates. In order to reduce potential caveats of molecular methods, we analyzed three nuclear genes for a well-documented interordinal set of placental mammals including an extensive taxon sampling of rodent families. Nuclear genes were favored because it has been

demonstrated that they contain more phylogenetic signal than mitochondrial genes (e.g., Springer et al. 2001a). Additionally, several independent calibration points—within and outside rodents—were considered in an attempt to evaluate whether the local molecular clock approach will be able to reconcile molecular data with the paleontological record.

MATERIAL & METHODS

Sequence data.

The nucleotide sequence data set used here was previously obtained by Huchon et al. (2002). It includes 21 rodents, 19 other placental mammals and two marsupials for three nuclear markers: the Alpha 2B Adrenergic Receptor gene (ADRA2B ; 1170 sites), partial exon 1 of the Interphotoreceptor Retinoid Binding Protein gene (IRBP ; 1227 sites), and exon 28 of the von Willebrand Factor gene (vWF ; 1236 sites). Stationary base composition was evaluated at the 1% level of chi-square tests by comparing the nucleotide composition of each sequence to the frequency distribution assumed in the ML model. Non-stationary base composition was detected for ADRA2B, IRBP, and vWF, taken individually or in combination. The two marsupial outgroups (Macropus sp., Didelphis sp.) systematically deviated from base composition stationary assumptions whatever the data set considered. This was not the case for placental taxa, where the deviating taxa were different from one data set to an other. Dryomys nitedula, Mus musculus, and Bradypus tridactylus were heterogeneous for ADRA2B, as were Marmota monax and Orycteropus afer for IRBP, Dipodomys merriami, Thomomys talpoides, Lama sp., Physeter catodon, and Procavia capensis for vWF, and Mus musculus, Rattus norvegicus, Castor canadensis, Dryomys nitedula, Petromus typicus, Lepus crawshayi, Oryctolagus cuniculus, Physeter catodon, Orycteropus afer, Procavia capensis, and Bradypus tridactylus for the concatenated data set.

A closer examination revealed that, as expected, the base composition at third codon positions was responsible for this heterogeneity. Therefore, we decided to exclude all third codon positions, and combined the three markers in order to average lineage- and gene-specific rate heterogeneities. This data set of concatenated first and second codon positions represents a total of 2422 aligned sites for which stationary base composition was respected for all placental taxa.

Maximum likelihood analyses.

The General Time Reversible (GTR or REV) model of sequence evolution was chosen for nucleotides (Lanave et al. 1984; Yang 1994). Base composition was empirically estimated from the data set. Rate heterogeneity among DNA sites was described by a discrete Gamma distribution with 8 categories (Γ_8) (Yang 1996). All analyses of local molecular clocks and corresponding datings were conducted under PAML (Yang 1997), version 3.1. The highest-likelihood topology used as a reference for the dating was the one identified by Huchon et al. (2002) using the concatenation of ADRA2B, IRBP, and vWF. Optimal ML parameters were A↔C = 2.26, A↔G = 6.34, A↔T = 0.92, C↔G = 1.68, C↔T = 4.48, and G↔T = 1.00 for the substitution rate matrix, and $\alpha = 0.47$ for the Gamma distribution, yielding a log-likelihood of ln$\underline{L}$ = -26054.39. This topology was compatible with the most recent studies on placental ordinal phylogenetic relationships in depicting the monophyly of Rodentia, Glires, Euarchontoglires, Laurasiatheria, Boreoeutheria, and Afrotheria (Madsen et al. 2001; Murphy et al. 2001a). Minor departures from the topology robustly identified by Murphy et al. (2001b) on 16.4 kb of nuclear and mitochondrial DNA involved only weakly supported nodes for the position of Eulipotyphla (here represented by Erinaceus), Pholidota (Manis), Dermoptera (Cynocephalus), and the location of the root of placentals (see Delsuc et al. 2002, for a discussion of this point).

Relative rate tests.

Rate heterogeneity among taxa was detected using the method developed by Takezaki et al. (1995) as implemented in the two-cluster and branch-length tests of the LINTRE package (www.bio.psu.edu/People/Faculty/Nei/Lab). The two-cluster test was a relative rate test that examined whether there is a change in substitution rate between the two descendant lineages (i.e., cluster) on each node of the bifurcating tree. The branch length test evaluated the deviation of the root-to-tip distance of each terminal taxa relative to the average root-to-tip length. Tests were conducted under the Tamura-Nei distance with $\alpha = 0.47$ (the GTR distance was not available under LINTRE). Fastest- and slowest-rate branches and species were identified by the two-cluster and the branch-length tests at the stringent $P < 0.01$ significance level. They then served as a basis for the definition of the local molecular clocks.

Definition of local clocks.

The use of the standard likelihood ratio test for comparison of models with and without clock (Felsenstein 1988) was not appropriate here because it did not allow a comparison of two models with the same number of parameters (i.e., in our case, models including the same number of local molecular clocks). Therefore, following the pioneering approach of Kishino and Hasegawa (1990), the performances of the different clock models were evaluated using the Akaike Information Criterion (AIC; Akaike 1974), the lowest AIC values being indicative of the model that best fit the data. The AIC was calculated as $-2 \times \ln L + 2 \times$ (number of free parameters of the model), where $\ln L$ was the log-likelihood of the tree.

Calibration points.

We chose seven calibration points spanning the different placental clades with a special focus on rodents, following Huchon et al. (2002): (i) Radiation of Caviomorpha at 31

Myr (Walton 1997; Wyss et al. 1993) ; (ii) Mus vs. Rattus at 14 Myr (Jacobs and Downs 1994) ; (iii) Glis vs. Dryomys at 28.5 Myr (identification of Glirinae since the Late Oligocene: Hartenberger 1994) ; (iv) Aplodontia vs. Marmota at 37 Myr (identification of Sciuridae since Late Eocene: McKenna and Bell 1997) ; (v) Ochotona vs. Leporidae at 37 Myr (identification of ochotonids since Late Eocene: McKenna and Bell 1997) ; (vi) Lama vs. other cetartiodactyls (e.g., Physeter) at 63 Myr (Gingerich and Uhen 1998) ; and (vii) Dugong vs. Procavia at 60 Myr (identification of Paenungulata since Paleocene: Gheerbrant et al. 1996). Within Glires, ages corresponded to the minimum estimates for the crown group considered. These seven paleontological calibration points were independently used on the best local clock phylogram to obtain seven sets of divergence date estimates for the main placental clades.

RESULTS

Extensive lineage-specific variation of evolutionary rates.

The results of the branch length and two-cluster tests are recapitulated in Figure 1. Four taxa are evolving significantly faster than other placentals—the sloth (Bradypus), the hedgehog (Erinaceus), and two geomyoid rodents (Dipodomys, Thomomys)—, whereas eight are evolving slower—a carnivore (Felis), a perissodactyl (Equus), a primate (Homo), a flying-lemur (Cynocephalus), and four rodents (Aplodontia, Glis, Anomalurus, and Castor). For all these taxa, the ratio of the mean root-to-tip distance to its standard-error was higher than 3, leading to a highly significant branch length test ($P < 0.01$). The two-cluster test recorded significant differences ($P < 0.01$) in evolutionary rates for terminal branches—e.g., Dugong versus

Procavia—as well as internal branches—e.g., the ones subtending sciuroids and glirids (Figure 1).

Setting the local molecular clocks.

The results of the branch length test (cf. Figure 1) are used to define the local molecular clocks. Nine models combining different local clocks are evaluated (Figure 2). All taxa but the faster- and slower-evolving are attributed the local default rate of $\mathcal{R}_D = 1.00$. For the four faster-evolving species, there are two possibilities. Either a single local clock is defined for Bradypus, Erinaceus, Dipodomys and Thomomys (the so-called model A [lnL = -26161.14]: Figure 2), or different local clocks are defined, one for the sloth, one for the hedgehog, and one for the two geomyoid rodents (model D [lnL = -26159.02]). Regarding the latter point, the definition of an unique evolutionary rate for both Dipodomys and Thomomys is obviously a more parsimonious solution relative to the definition of one independent local clock for each rodent. Moreover, the two-cluster test detected a rate variation along the branch leading to the two geomyoids relative to Castor: the same local clock is thus attributed to the ancestral branch of geomyoids and their descendants Dipodomys and Thomomys. For the eight slower-evolving taxa, the same possibilities exist, namely defining either a single local clock for all (model B [lnL = -26131.45]) or independent clocks for each taxon (model G [lnL = -26122.96]). To better describe the data, the previous models are combined in order to simultaneously consider slow and fast local clocks on the same tree: model C [lnL = -26102.76] combines A and B, and model H [lnL = -26094.70] combines D and G.

The next step was to try to increase the log-likelihood of the tree by attributing new local rates to species that were initially considered as evolving according to the default rate $r_D = 1.00$. These taxa were chosen based on the results of the two-cluster test. In order to reduce the number of parameters, we defined only one local clock—either slow or fast—per pairs of terminal branches, and chose the rate yielding the higher log-likelihood. Cynopterus, Manis, and Marmota

were thus identified as slower-evolving species, and Procavia, Ochotona, Massoutiera, and Cavia as faster-evolving taxa. Local clocks for these taxa were incorporated in models with an increasing number of independent rates: model E ($\ln L$ = -26085.97), F ($\ln L$ = -26089.24), and I ($\ln L$ = -26081.32) (cf. Figure 2).

The difficulty of this multi-local clocks approach is to find a reasonable compromise between the number of parameters added—here the local clocks—and the gain of log-likelihood. As a measure of this compromise, we followed Kishino and Hasegawa (1990) and used the AIC. For example on model E, the log-likelihood of the model with five local clocks is $\ln L$ = -26085.97, with the following number of ML free parameters: five for the GTR model, one for the rate heterogeneity, and 45 for the branch lengths (42 taxa + 5 local clocks – 2 parameters as described in Yoder and Yang (2000)). The AIC of this model is therefore equal to 2 x 26085.97 + 2 x 51 = 52273.94. This value is the lowest found here, even relative to the AIC of the hypothesis without any clock ($AIC_{\text{no clock}}$ = 52282.78). Figure 2 recapitulates the relationship between the AIC values and the number of local clocks defined.

Another difficulty of the multi-local clocks approach is to choose among all the possible rate assignments for the various lineages (Sanderson 1997). Because of the tremendous number of possibilities, we here evaluated only nine models (A to I) which span quite a large range in number of local clocks (2 to 18). The best model (E) found is based on 5 local clocks (Figure 2). In this model, Massoutiera is fast-evolving, and Bradypus, Procavia, Erinaceus, Ochotona, Dipodomys and Thomomys, and Cavia are the fastest-evolving taxa, and conversely Cynopterus, Manis, Marmota and Aplodontia are slow-evolving, and Felis, Equus, Homo, Cynocephalus, Glis, Anomalurus, and Castor are the slowest-evolving taxa (Figure 3).

The ML estimate of the local clocks also provides a measure of the evolutionary rate contrast among these lineages. For the first and second codon positions of the three concatenated nuclear genes ADRA2B, IRBP, and vWF, the fastest-evolving species exhibit a 2.6-fold excess of evolutionary rates relative to the slowest ones. When 18 independent local clocks are defined

(model I: Figure2), a greater excess of 3.8-fold is detected between Cavia (local rate of 1.67) and Anomalurus (0.44).

Crossed-calibrations based on seven points.

Seven calibration points are mapped on the 5-local clocks phylogram (Figure 3) in order to estimate: i) the divergence dates between the calibrating taxa, i.e., those taxa for which a paleontological age of divergence is suggested ; and ii) the age of the most recent common ancestors (MRCA) of 6 major mammalian clades, i.e., the murines (Mus versus Rattus), rodents, glires, euarchontoglires, boreoeutherians, and placentals.

Table 1 presents the results of the crossed-calibrations. To evaluate the accuracy of each calibration point—i.e., its ability to correctly estimate the true date of divergence—we searched whether a given calibration point yielded a 95% confidence interval (mean date ± 1.96 standard-error) containing the supposed splitting date of the six remaining calibration points. Cross-calibrations indicate that our calibration points are not fully congruent between them, i.e., none of them could give 95% confidence intervals that include the paleontological date of the six other calibration points (Table 1). Three calibration points (Caviomorpha at 31 Myr, Mus / Rattus at 14 Myr, and Sciuroidea at 37 Myr) could recover two paleontological dates. Two calibration points (Gliridae at 28.5 Myr and Cetartiodactyla at 63 Myr) could recover one paleontological date. Paenungulata at 60 Myr was unable to be congruent with any of the six other calibration points. Interestingly, some calibration points were reciprocally and accurately compatible. For example, the glirids at 28.5 Myr suggest Mus and Rattus be separated 12.8 Myr ago—the 95% confidence interval is 10.2-15.4 Myr, and contains the 14 Myr fossil estimate—, and conversely, Mus / Rattus at 14 Myr suggest Glis and Dryomys separated 31.3 Myr—the confidence interval is 26.5-36.1 Myr, and contains the 28.5 Myr fossil estimate. Three other pairs of calibration points are reciprocally compatible: caviomorphs and sciuroids, caviomorphs and lagomorphs, and Mus / Rattus and sciuroids (Table 1).

Divergence dates of mammals.

Table 2 presents the results of the local molecular-clock dating for the age of the MRCA of 6 mammalian clades. As expected from previous works indicating that divergence times depend upon the choice of calibration points (Huchon et al. 2000; Soltis et al. 2001), the age of the MRCA of murines, rodents, glires, euarchontoglires, boreoeutherians, and placentals is highly variable depending on the calibrating taxa. For example, the Gliridae calibration point suggests the MRCA of rodents be ~47 Myr old, while the Paenungulata point suggests it to be ~121 Myr old. To summarize the results based on the seven independent calibration points, we referred to the median age of the corresponding MRCA. When the seven calibration points are considered, the median ages for the MRCA of murines, rodents, glires, euarchontoglires, boreoeutherians, and placentals are ~18, 66, 73, 78, 83, and 87 Myr respectively (Table 2). However, more recent estimates are obtained after the removal of the incompatible cetartiodactyl and paenungulate calibrations (see Table 2). In this case, the median ages of the five calibrations are respectively ~16, 60, 66, 70, 75, and 78 Myr.

DISCUSSION

Incompatibility of divergence times calculated from independent calibrations points.

A great variance of molecular estimates of divergence times has been described for mammals (Bromham et al. 1999; Huchon et al. 2000). As an illustration of this variance, molecular datings reported for the split between hystricognath and sciurognath rodents range from values as disparate as 75 to 125 Myr (Janke et al. 1997; Kumar and Hedges 1998; Cao et al. 2000; Adkins et al. 2001). Problematically, this 50 Myr difference between two extreme

estimates is of the same order of magnitude as the paleontological age of 55 Myr for modern placentals (Hartenberger 1998). In this paper, we attempted to understand the origins of the conflict between molecules and fossils by using an approach accounting for evolutionary rate variations—the ML local molecular clocks (Yoder and Yang 2000)—, and a set of seven independent calibrations spread over the placental tree.

Mean local molecular clock estimates for the age of the MRCA of murines, rodents, glires, euarchontoglires, boreoeutherians, and placentals over these seven calibrations are respectively ~19 ± 7, 71 ± 25, 79 ± 27, 84 ± 30, 89 ± 31, and 94 ± 33 Myr. The large standard-errors reflect the strong dependence of date estimates upon the choice of calibration points. In order to better understand the compatibility properties of these seven calibrations, we evaluated the level of inadequacy between paleontological (A_{PAL}) and molecular (A_{MOL}) divergence times (cf. Table 1). This inadequacy ($I_{X / Y}$) was quantified for each calibration point (e.g., node X), and for each of the 6 MRCAs under focus (e.g., Y), by the absolute difference between the date estimated by the local clocks and the age expected under paleontological hypotheses, standardized by the latter age, and expressed in percentage, i.e., by $I_{X / Y} = 100 \times |A_{PAL}$ (node X) $- A_{MOL}$ (node Y) $| / A_{PAL}$ (node X). If we rank the seven calibration points by decreasing order of agreement between fossils and molecules as evaluated by an increasing mean percent of inadequacy over these seven values, we have sciuroids (mean $I_{X / Y}$ ± standard error = 24% ± 15), caviomorphs (25% ± 16), lagomorphs and Mus / Rattus (28% ± 17), glirids (33% ± 18), cetartiodactyls (41% ± 22), and then paenungulates (100% ± 42). Clearly, datings obtained using the paenungulate—and to a lesser extent the cetartiodactyl—calibration points are at odds with estimates derived from other paleontological calibrations. This incompatibility of the cetartiodactyl and paenungulate calibrations relative to the five other points chosen within Glires may be explained by limitations of either the local clock method and / or the fossil record. (i) The taxon sampling for paenungulates is sparse in our phylogram (Figure 1), which might have affected the

accuracy of branch length estimation within this clade. (ii) The cetartiodactyl calibration point is based on controversial fossil inferences. The 63 Myr value is inferred from the Mesonychia-Arctocyonia divergence (Gingerich and Uhen 1998), which assumes that Mesonychia are the sister clade of cetaceans. This hypothesis has been recently challenged by new paleontological discoveries (Thewissen et al. 2001).

Choosing among independent calibration points

We here consider mainly calibrations within rodents, with at least one point in each of the major rodent clades (Adkins et al. 2001; Huchon et al. 2002): sciuroids and glirids in the Sciuroidea + Gliridae clade, murines in the Anomaluromorpha + Castoridae + Geomyoidea + Myodonta clade, and caviomorphs in the Ctenohystrica. Together with the lagomorph point, it is clear that the rodent calibrations were characterized by greater reciprocal compatibility. Conversely, the greatest inadequacy between molecular and paleontological ages was given by the two points outside Glires: the cetartiodactyls and the paenungulates. This observation contradicts the point of view of Adkins et al. (2001) who suggested that "the inclusion of a rodent calibration point to estimate rodent divergence dates introduces a confounding influence [on the molecular dates]".

Actually, the phylogenetic distance $D_{X/Y}$ between each pair (X, Y) of calibration points—expressed in substitution number per 100 sites—can be measured by the node-to-node sum of the branch lengths on the highest-likelihood phylogram (Figure 1). A positive correlation is found between this distance and the mean percent of inadequacy $I_{X/Y}$ between fossil and molecular dates ($I_{X/Y} = 9.7 \times D_{X/Y} - 29.0$; $r^2 = 0.42$; data not shown). In other words, an additional 9.7% of inadequacy between the expected paleontological age and the local clock estimate accumulates for each additional substitution per 100 sites separating the MRCAs under focus on the ML tree. This observation is not unexpected given that the most confounding cause for errors in the molecular dating is likely to be the variation of

evolutionary rates among lineages. There are more possibilities for increasing (or decreasing) substitution rate variations to accumulate for more distant calibrating nodes—e.g., the paenungulate point is farther than the caviomorph one relative to the Mus / Rattus split—, and therefore more possibilities of obtaining reciprocally incompatible local clock estimates of divergence times. If there is a problem of over- or under-estimation with a given calibration point, its uniformly inflating or deflating impact will be the lower on the nodes located in its closest vicinity.

Setting the local molecular clocks

The divergence times obtained by our approach directly depends upon the setting of the different local rate models (A to I). The incorporation of different local rates in the tree is contingent on the interpretation of results from the branch length and two-cluster tests. With the two-cluster test, we chose to attribute local clocks to only one of the two branches displaying contrasted rates of evolution. By focusing on terminal branches, we aim to avoid over-parameterization of the models. Alternative possibilities do exist such as defining local clocks on internal branches of the phylogram—e.g., those leading to several calibrating nodes (cetartiodactyls, sciuroids, glirids, and murines: cf. Figure 1). Due to the high number of potential combinations of local clocks, we only explored nine models incorporating from two to 18 local clocks, but we cannot rule out the possibility that a better fit of the data might be found by using more complex models that would give lower AIC relative to the 5-clocks model (E). Future developments of the local clock approach might involve the construction of algorithms to explore the attribution of rates to taxa and identify the best local clock tree.

Detecting rate heterogeneities among lineages by the local clocks.

We observed a better fit of the data with less branch rates (AIC of 52273.94 with five local clocks against 52282.78 without clock). Under maximum likelihood, an independent rate is

usually attributed to each branch of the phylogram, yielding 2N – 3 rates for an unrooted bifurcating tree with N taxa (42 species for our data). In the particular case of the first and second codon positions of ADRA2B + IRBP + vWF, we here show that defining as few as five rates instead of 2 x 42 – 3 = 81 reasonably fits the DNA sequence data. This might suggest that the 2N – 3 branch parameters actually correspond to an over-parameterization of the ML model, at least in the present case.

With our 5-rates model, we identified extensive rate variation between placental lineages, but also within rodents, with detection of a three-fold contrast between the slowest and fastest evolving taxa. We incorporated an extensive taxon sampling within rodents spanning all sciurognath families and all hystricognath superfamilies. In light of our results, it is no longer accurate to suggest that rodents as a group are evolving faster than man (e.g., Wu and Li 1985). Indeed, for the three nuclear exons under focus, rodents contain both among the slowest- and fastest-evolving placental species, with respectively the scaly-tailed flying squirrel Anomalurus (local clock $R_{S8} = 0.44$: cf. Figure 2, model I) versus the pocket mouse Thomomys and the kangaroo rat Dipodomys (local clock $R_{F2} = 1.51$), and the Guinea pig (local clock $R_{F6} = 1.67$). One should note that these local clock values should not be considered as absolute values. Indeed, the estimate of each local rate is influenced by the choice of the others, as this can also be noticed from the results of Kishino and Hasegawa (1990). For example, in model A (Figure 2), a single rate of 1.82 is estimated for all fast-evolving taxa. In model B, a single rate of 0.57 is estimated for all slow-evolving taxa. When the two models are combined (model C), the ML estimate of the fast rate decreases to 1.51, and the slow rate increases to 0.62.

Potential caveats of using relative rate tests to detect evolutionary rate variation have been pointed out by Bromham et al. (2000). These authors suggested that relative rate tests such as triplet relative rates or likelihood ratio tests are limited in power to detect rate variations when using short sequences. Here we used ~2400 first and second codon positions, giving us more than 90% chance of detecting a two-fold (or more) lineage-specific rate variation (Bromham et al.

2000). It is thus likely that most of the rate variation has been accounted for in our definition of local clocks. However, if moderate rate variations remain unincorporated in our local clock approach, they should have only slightly affected our date estimates.

The chronology of the diversification of placentals.

Despite the undeniable existence of gaps in the fossil record (Easteal 1999), some molecular dates are clearly incompatible with paleontology. For example, regarding the age of divergence of crown placentals, two rodent calibration points yield dates that are recent (~62-68 Myr for the murine and glirid calibrations: Table 2) relative to the fossil record, which suggests that placentals were already diversified 85-90 Myr ago (e.g., Archibald et al. 2001). Conversely, the date of 159 Myr for the divergence of Placentalia (paenungulate calibration: Table 2) seems ancient relative to the fossil record because stem eutherians are identified in the Early Cretaceous around 125 Myr ago (Ji et al. 2002).

We are thus confirming previous observations of the strong dependence of molecular date estimates of cladogenesis events upon the choice of the fossil references used for calibrations (Lee 1999; Huchon et al. 2000; Soltis et al. 2002). In order to decrease the impact of using a single calibration point, we averaged estimates derived from 5 independent calibrations displaying only slight reciprocal incompatibility—i.e., the caviomorph, murine, glirid, sciuroid, and lagomorph points. The MRCAs of rodents, primates and rodents, boreoeutherians, and placentals were therefore estimated to be respectively ~ 60, 70, 75, and 78 Myr old.

More generally, molecular dates of divergence based on the use of a single calibration point are very crude, especially when times are computed from very recent calibrations (Ayala et al. 1998) or when time estimates for a group are derived from extrapolation on the times calculated for another (van Tuinen and Hedges 2001). It is thus attractive to simultaneously incorporate information from two independent calibration points in molecular

dating, as proposed by Rambaut and Bromham (1998) in their quartet dating method. This method allows rate variation in each of the two lineages under focus. Application of quartet dating provided the following ranges of mean divergence time estimates: 55.8 Myr for the MRCA of rodents, 68.0-69.4 for the one of primates and rodents, 79.0 for the one of boreoeutherians, and 76.1-103.0 for placentals (Eizirik et al. 2001; Murphy et al. 2001b; Huchon et al. 2002). When several quartets of taxa are available to evaluate the age of a given node, local clock estimates are quite close to the most recent quartet dating estimates. Thus, molecular analyses based on global clocks, local clocks, and quartet dating seem to converge toward a Late Cretaceous origin of the crown placental clades (Easteal 1999; Huchon et al. 2000; Eizirik et al. 2001; Murphy et al. 2001b) followed by a Tertiary radiation of some orders like rodents (Huchon et al. 2002) and chiropterans (Springer et al. 2001b).

The age of the split between mice and rats.

Among our seven phylogenetically independent calibration points, we incorporated the controversial Mus / Rattus split. Because of the consistently deeper molecular estimates of the MRCA of mouse and rat (23-42 Myr: Janke et al. 1994; Kumar and Hedges 1998; Huchon et al. 2000; Adkins et al. 2001; Nei et al. 2001) relative to the fossil record (14 Myr: Jacobs and Downs 1994), the inclusion of this calibration in our analyses may have biased the estimates of the other dates towards more recent ages. However, the five less-incompatible calibration points yielded younger estimates (~13-19 Myr) relative to those previously published (Table 2). Even the more incompatible cetartiodactyl calibration provides a divergence time (22 Myr) virtually identical to the one obtained by Adkins et al. (2001) based on the DNA comparison of the growth hormone receptor gene. By contrast, the paenungulate calibration—shown to be highly incompatible with the other calibrations (see above)—yielded 33 Myr (Table 2), an estimate identical to the one obtained by Nei et al. (2001) on the concatenation of 104 protein sequences.

A proposed explanation for the discrepancy between molecules and fossils for the age of the split between mice and rats is the high substitution rates of these two rodents. For example, murid rodents—the family to which mouse and rat belong—display particularly high substitution rates in their mitochondrial genomes (Gissi et al. 2000). This peculiarity possibly related to a change in mutational process might have obscured their phylogenetic placement for a long time (Lin et al. 2002). Relative to previous molecular estimates, the younger dates obtained by Adkins et al. (2001), and in this paper, might be explained by a denser taxon sampling in rodents that breaks the long isolated branch leading to murids, allowing a more accurate estimation of its length.

Conclusions and perspectives.

Additional causes may be invoked to explain why different fossil calibrations can yield sometimes radically different date estimates using the local clock method. i) The choice of the genes. We combined three genetically independent nuclear markers in order to average the individual behavior of each exon ; ii) The choice of taxa. Substitution rate variations are widespread in mammalian genomes, and it is possible that rate variation would be better detected with a larger number of taxa. By decreasing the average length of terminal branches, an increased taxon sampling in a given clade will increase the accuracy of the reconstructed phylogeny (Rannala et al. 1998). For this reason, calibration points subtending a few isolated branches (e.g., paenungulates, sciuroids, and glirids) might be less relevant than those subtending a denser sampling (e.g., cetartiodactyls and caviomorphs). It is thus likely that the dating estimates obtained from, for example, the paenungulate calibration point, might be improved if more afrotherians were taken into account. iii) The choice of the topology. The tree we used has been confidently identified and evaluated by Huchon et al. (2002), and is congruent with those independently inferred by Adkins et al. (2001), Madsen et al. (2001), and Murphy et al. (2001a, b). We should keep in mind that the confidence intervals on

molecular dates are roughly of the same order of magnitude whatever the choice of sites, taxa, and topologies (Sanderson and Doyle 2001). It is thus likely that the mean and standard errors on divergence dates obtained with the local clock approach are not strongly depending upon the choice of ADRA2B, IRBP, and vWF as molecular markers, a denser species representation within Glires, and the highest-likelihood phylogram as the reference topology for molecular dating purposes. However, the present study may benefit from an expanded sampling of markers among nuclear genes, and of taxa among the three other recently identified major placental clades (Laurasiatheria, Xenarthra, and Afrotheria: Madsen et al. 2001; Murphy et al. 2001a,b).

The lineage-specific variation of evolutionary rates documented in numerous mammalian genomes requires the development of powerful molecular dating methods. Global clocks, local clocks and quartet dating are limited in their application by the assumption of rate constancy at either large or reduced phylogenetic scale. Non- or semi-parametric (Sanderson 1997; Sanderson 2002) and Bayesian (Thorne et al. 1998; Huelsenbeck et al. 2000; Kishino et al. 2001; Aris-Brosou and Yang 2002; Thorne and Kishino 2002) methods that model rate variation through time are another way to accommodate lineage-specific rate heterogeneity. The performance of these methods relies on the distribution of the parameters used to model the evolution of rates, and their use within placental mammals appears promising (Cao et al 2000). Improvements in dating methods in concert with the increase of gene and taxon sampling in molecular studies is likely to yield more accurate divergence dates. Higher reliability of molecular dating estimates combined to new paleontological discoveries might help to reconcile rocks and clocks.

ACKNOWLEDGMENTS

This work has been supported by the TMR Network "Mammalian phylogeny" (contract FMRX – CT98 – 022) of the European Community to MJS, the "Genopole Montpellier Languedoc-Roussillon", and the "Action Bioinformatique inter-EPST" of the CNRS. We wish to thank the "Centre de Ressources" INFOBIOGEN (Evry, France), and Jacques Demaille and Denis Pugnère (Institut de Génétique Humaine de Montpellier, CNRS UPR 1142) for computing facilities. We thank an anonymous referee for comments on the manuscript. This is the contribution N° 2002-070 of the Institut des Sciences de l'Evolution de Montpellier (UMR 5554 - CNRS).

REFERENCES

Adkins RM, Honeycutt RL (1994) Evolution of the primate cytochrome c oxidase subunit II gene. J. Mol. Evol. 38:215-231

Adkins RM, Gelke EL, Rowe D, Honeycutt RL (2001) Molecular phylogeny and divergence time estimates for major rodent groups: evidence from multiple genes. Mol. Biol. Evol. 18:777-791

Akaike H (1974) A new look at the statistical model identification. IEEE Trans. Automat. Contr. AC-19:716-723

Alroy J (1999) The fossil record of North American mammals: evidence for a Paleocene evolutionary radiation. Syst. Biol. 48:107-118

Andrews TD, Jermiin LS, Easteal S (1998) Accelerated evolution of cytochrome b in simian primates: adaptive evolution in concert with other mitochondrial proteins? J. Mol. Evol. 47:249-257.

Andrews TD, Easteal S (2000) Evolutionary rate acceleration of cytochrome c oxidase subunit I in simian primates. J. Mol. Evol. 50:562-568.

Archibald JD, Averianov AO, Ekdale EG (2001) Late Cretaceous relatives of rabbits, rodents, and other extant eutherian mammals. Nature 414:62-65

Aris-Brosou S, Yang Z (2002) Effects of models of rate evolution on estimation of divergence dates with special reference to the metazoan 18S ribosomal RNA phylogeny. Syst. Biol. 51:703-714

Ayala FJ, Rzhetsky A, Ayala FJ (1998) Origin of the metazoan phyla: molecular clocks confirm paleontological estimates. Proc. Natl. Acad. Sci. USA 95:606-611

Benton M (1999) Early origins of modern birds and mammals: molecules vs. morphology. Bioessays 21:1043-1051

Bousquet J, Strauss SH, Doerksen AH, Price RA (1992) Extensive variation in evolutionary rate of rbcL gene sequences among seed plants. Proc. Natl. Acad. Sci. USA 89:7844-7848

Bromham L, Phillips MJ, Penny D (1999) Growing up with dinosaurs: molecular dates and the mammalian radiation. Trends Ecol. Evol. 14:113-118

Bromham L, Penny D, Rambaut A, Hendy MD (2000) The power of relative rates tests depends on the data. J. Mol. Evol. 50:296-301

Cao Y, Fujiwara M, Nikaido M, Okada N, Hasegawa M (2000) Interordinal relationships and timescale of eutherian evolution as inferred from mitochondrial genome data. Gene 259:149-158

Catzeflis FM, Hänni C, Sourrouille P, Douzery E (1995) Molecular systematics of hystricognath rodents: the contribution of sciurognath mitochondrial 12S rRNA sequences. Mol. Phylogenet. Evol. 4:357-360

DeBry RW, Sagel RM (2001) Phylogeny of Rodentia (Mammalia) inferred from the nuclear-encoded gene IRBP. Mol. Phylogenet. Evol. 19:290-301

Delsuc F, Catzeflis FM, Stanhope M, Douzery EJP (2001) The evolution of armadillos, anteaters and sloths depicted by nuclear and mitochondrial phylogenies: implications for the status of the enigmatic fossil Eurotamandua. Proc. R. Soc. Lond. B 268:1605-1615

Delsuc F, Scally M, Madsen O, Stanhope MJ, de Jong WW, Catzeflis F, Springer MS, Douzery EJP (2002) Molecular phylogeny of living xenarthrans and the impact of character and taxon sampling on the placental tree rooting. Mol. Biol. Evol. 19:1656-1671

Easteal S (1999) Molecular evidence for the early divergence of placental mammals. BioEssays 21:1052-1058

Eizirik E, Murphy WJ, O'Brien SJ (2001) Molecular dating and biogeography of the early placental mammal radiation. J. Hered. 92:212-219

Felsenstein J (1988) Phylogenies from molecular sequences: inference and reliability. Ann. Rev. Genet. 22:521-565

Fieldhouse D, Yazdani F, Golding GB (1997) Substitution rate variation in closely related rodent species. Heredity 78:21-31

Gheerbrant E, Sudre J, Cappetta H (1996) A Palaeocene proboscidean from Morocco. Nature 383:68-70

Gingerich PD, Uhen MD (1998) Likelihood estimation of the time of origin of Cetacea and the time of divergence of Cetacea and Artiodactyla. Paleontologia Electronica 2:1-47

Gissi C, Reyes A, Pesole G, Saccone C (2000) Lineage-specific evolutionary rate in mammalian mtDNA. Mol. Biol. Evol. 17:1022-1031

Hartenberger J-L (1994) The evolution of the Gliroidea. In: Tomida Y, Li C-K, Setoguchi T (eds) Rodent and lagomorph families of Asian origins and diversification. National Science Museum Monographs, Tokyo, pp 19-33

Hartenberger J-L (1998) Description de la radiation des Rodentia (Mammalia) du Paléocène supérieur au Miocène ; incidences phylogénétiques. C. R. Acad. Sci. Paris, Sciences de la terre et des planètes 326:439-444

Huchon D, Catzeflis FM, Douzery EJP (1999) Molecular evolution of the nuclear von Willebrand Factor gene in mammals and the phylogeny of rodents. Mol. Biol. Evol. 16:577-589

Huchon D, Catzeflis F, Douzery EJP (2000) Variance of molecular datings, evolution of rodents, and the phylogenetic affinities between Ctenodactylidae and Hystricognathi. Proc. R. Soc. Lond. B 267:393-402

Huchon D, Douzery EJP (2001) From the Old World to the New World: a molecular chronicle of the phylogeny and biogeography of hystricognath rodents. Mol. Phylogenet. Evol. 20:238-251

Huchon D, Madsen O, Sibbald MJJB, Ament K, Stanhope M, Catzeflis F, de Jong WW, Douzery EJP (2002) Rodent phylogeny and a timescale for the evolution of Glires: evidence from an extensive taxon sampling using three nuclear genes. Mol. Biol. Evol. 19:1053-1065

Huelsenbeck JP (1998) Systematic bias in phylogenetic analysis: is the Strepsiptera problem solved? Syst. Biol. 47:519-537

Huelsenbeck JP, Larget B, Swofford D (2000) A compound poisson process for relaxing the molecular clock. Genetics 154:1879-1892

Jacobs LL, Downs WR (1994) The evolution of murine rodents in Asia. In: Tomida Y, Li C-K, Setoguchi T (eds) Rodent and lagomorph families of Asian origins and diversification. National Science Museum Monographs, Tokyo, pp 149-156

Janke A, Feldmaier-Fuchs G, Thomas K, Von Haeseler A, Pääbo S (1994) The marsupial mitochondrial genome and the evolution of placental mammals. Genetics 137:243-256

Janke A, Xu X, Arnason U (1997) The complete mitochondrial genome of the wallaroo (Macropus robustus) and the phylogenetic relationship among Monotremata, Marsupialia, and Eutheria. Proc. Natl. Acad. Sci. USA 94:1276-1281

Ji Q, Luo Z-X, Yuan C-X, Wible JR, Zhang J-P, Georgi J-A (2002) The earliest known eutherian mammal. Nature 416:816-822

Kishino H, Hasegawa M (1990) Converting distance to time: application to human evolution. Methods Enzymol. 183:550-570

Kishino H, Thorne JL, Bruno WJ (2001) Performance of a divergence time estimation method under a probabilistic model of rate evolution. Mol. Biol. Evol. 18:352-361

Kumar S, Hedges SB (1998) A molecular timescale for vertebrate evolution. Nature 392:917-920

Lanave C, Preparata G, Saccone C, Serio G (1984) A new method for calculating evolutionary substitution rates. J. Mol. Evol. 20:86-93

Lee MS (1999) Molecular clock calibrations and metazoan divergence dates. J. Mol. Evol. 1999 49:385-391

Lin Y-H, Waddell PJ, Penny D (2002) Pika and vole mitochondrial genomes increase support for both rodent monophyly and Glires. Gene 294:119-129

Liu J-C, Makova KD, Adkins RM, Gibson S, Li W-H (2001) Episodic evolution of growth hormone in primates and emergence of the species specificity of human growth hormone receptor. Mol. Biol. Evol. 18:945-953

Madsen O, Scally M, Douady CJ, Kao DJ, DeBry RW, Adkins R, Amrine H, Stanhope MJ, de Jong WW, Springer MS (2001) Parallel adaptative radiations in two major clades of placental mammals. Nature 409:610-614

Mallatt J, Sullivan J (1998) 28S and 18S rDNA sequences support the monophyly of lampreys and hagfishes. Mol. Biol. Evol. 15:1706-1718

McKenna MC, Bell SK (1997) Classification of mammals above the species level. Columbia University Press, New York

Michaux J, Catzeflis F (2000) The bushlike radiation of muroid rodents is exemplified by the molecular phylogeny of the LCAT nuclear gene. Mol. Phylogenet. Evol. 17:280-293

Michaux J, Reyes A, Catzeflis F (2001) Evolutionary history of the most speciose mammals: molecular phylogeny of muroid rodents. Mol. Biol. Evol. 18:2017-2031

Moncalvo J-M, Lutzoni FM, Rehner SA, Johnson J, Vilgalys R (2000) Phylogenetic relationships of agaric fungi based on nuclear large subunit ribosomal DNA sequences. Syst. Biol. 49:278-305

Murphy WJ, Eizirik E, Johnson WE, Zhang YP, Ryder OA, O'Brien S (2001a) Molecular phylogenetics and the origins of placental mammals. Nature 409:614-618

Murphy WJ, Eizirik E, O'Brien SJ, Madsen O, Scally M, Douady C, Teeling E, Ryder OA, Stanhope MJ, de Jong WW, Springer MS (2001b) Resolution of the early placental mammal radiation using Bayesian phylogenetics. Science 294:2348-2351

Nedbal MA, Allard MW, Honeycutt RL (1994) Molecular systematics of hystricognath rodents: evidence from the mitochondrial 12S rRNA gene. Mol. Phylogenet. Evol. 3:206-220

Nedbal MA, Honeycutt RL, Schlitter DA (1996) Higher-level systematics of rodents (Mammalia, Rodentia): evidence from the mitochondrial 12S rRNA gene. J. Mammal. Evol. 3:201-237

Nei M, Xu P, Glazko G (2001) Estimation of divergence times from multiprotein sequences for a few mammalian species and several distantly related organisms. Proc. Natl. Acad. Sci. USA 98:2497-2502

Philippe H, Germot A (2000) Phylogeny of eukaryotes based on ribosomal RNA: long-branch attraction and models of sequence evolution. Mol. Biol. Evol. 17:830-834

Rambaut A, Bromham L (1998) Estimating divergence dates from molecular sequences. Mol. Biol. Evol. 15:442-448

Rannala B, Huelsenbeck JP, Yang Z, Nielsen R (1998) Taxon sampling and the accuracy of large phylogenies. Syst. Biol. 47:702-710

Rodriguez-Trelles F, Tarrio R, Ayala FJ (2002) A methodological bias toward overestimation of molecular evolutionary time scales. Proc. Natl. Acad. Sci. USA 99:8112-8115

Rowe DL, Honeycutt RL (2002) Phylogenetic relationships, ecological correlates, and molecular evolution within the cavioidea (Mammalia, Rodentia). Mol. Biol. Evol. 19:263-277

Sanderson MJ (1997) A nonparametric approach to estimating divergence times in the absence of rate constancy. Mol. Biol. Evol. 14:1218-1231

Sanderson MJ, Doyle JA (2001) Sources of error and confidence intervals in estimating the age of angiosperms from rbcL and 18S rDNA data. Am. J. Bot. 88:1499-1516

Sanderson MJ (2002) Estimating absolute rates of molecular evolution and divergence times: a penalized likelihood approach. Mol. Biol. Evol. 19:101-109

Soltis PS, Soltis DE, Savolainen V, Crane PR, Barraclough TG (2002) Rate heterogeneity among lineages of tracheophytes: integration of molecular and fossil data and evidence for molecular living fossils. Proc. Natl. Acad. Sci. USA 99:4430-4435

Springer MS, DeBry RW, Douady C, Amrine HM, Madsen O, de Jong WW, Stanhope MJ (2001a) Mitochondrial versus nuclear gene sequences in deep-level mammalian phylogeny reconstruction. Mol. Biol. Evol. 18:132-143

Springer MS, Teeling E, Madsen O, Stanhope MJ, deJong WW (2001b) Integrated fossil and molecular data reconstruct bat echolocation. Proc. Natl. Acad. Sci. USA 98:6241-6246

Takezaki N, Rzhetsky A, Nei M (1995) Phylogenetic test of the molecular clock and linearized trees. Mol. Biol. Evol. 12:823-833

Thewissen JG, Williams EM, Roe LJ, Hussain ST (2001) Skeletons of terrestrial cetaceans and the relationship of whales to artiodactyls. Nature 413:277-81

Thorne JL., Kishino H (2002) Divergence time and evolutionary rate estimation with multilocus data. Syst. Biol. 51:689-702

Thorne JL, Kishino H, Painter IS (1998) Estimating the rate of evolution of the rate of molecular evolution. Mol. Biol. Evol. 15:1647-1657

van Tuinen M, Hedges SB (2001) Calibration of avian molecular clocks. Mol. Biol. Evol. 18:206-213

Vucetich MG, Verzi DH, Hartenberger J-L (1999) Review and analysis of the radiation of the South American Hystricognathi (Mammalia, Rodentia). C. R. Acad. Sci. Paris, Earth & Planetary Sciences 329:763-769

Walton AH (1997) Rodents. In: Kay RF, Madden RH, Ciffelli RL, Flynn JJ (eds) Vertebrate paleontology in the neotropics. The Miocene fauna of La Venta, Columbia. Smithsonian Institution Press, Washington and London, pp 392-409

Wu C-I, Li W-H (1985) Evidence for higher rates of nucleotide substitution in rodents than in man. Proc. Natl. Acad. Sci. USA 82:1741-1745

Wyss AR, Flynn JJ, Norell MA, Swisher III CC, Charrier R, Novacek MJ, McKenna MC (1993) South America's earliest rodent and recognition of a new interval of mammalian evolution. Nature 365:434-437

Yang Z (1994) Estimating the pattern of nucleotide substitution. J. Mol. Evol. 39:105-111

Yang Z (1996) Among-site rate variation and its impact on phylogenetic analyses. Trends Ecol. Evol. 11:367-372

Yang Z (1997) PAML: a program package for phylogenetic analysis by maximum likelihood. CABIOS 13:555-556

Yoder AD, Yang Z (2000) Estimation of primate speciation dates using local molecular clocks. Mol. Biol. Evol. 17:1081-1190

Zuckerkandl E, Pauling L (1965) Evolutionary divergence and convergence in proteins. In: Bryson V, Vogel HJ (eds) Evolving genes and proteins. Academic Press, New York, pp 97-166

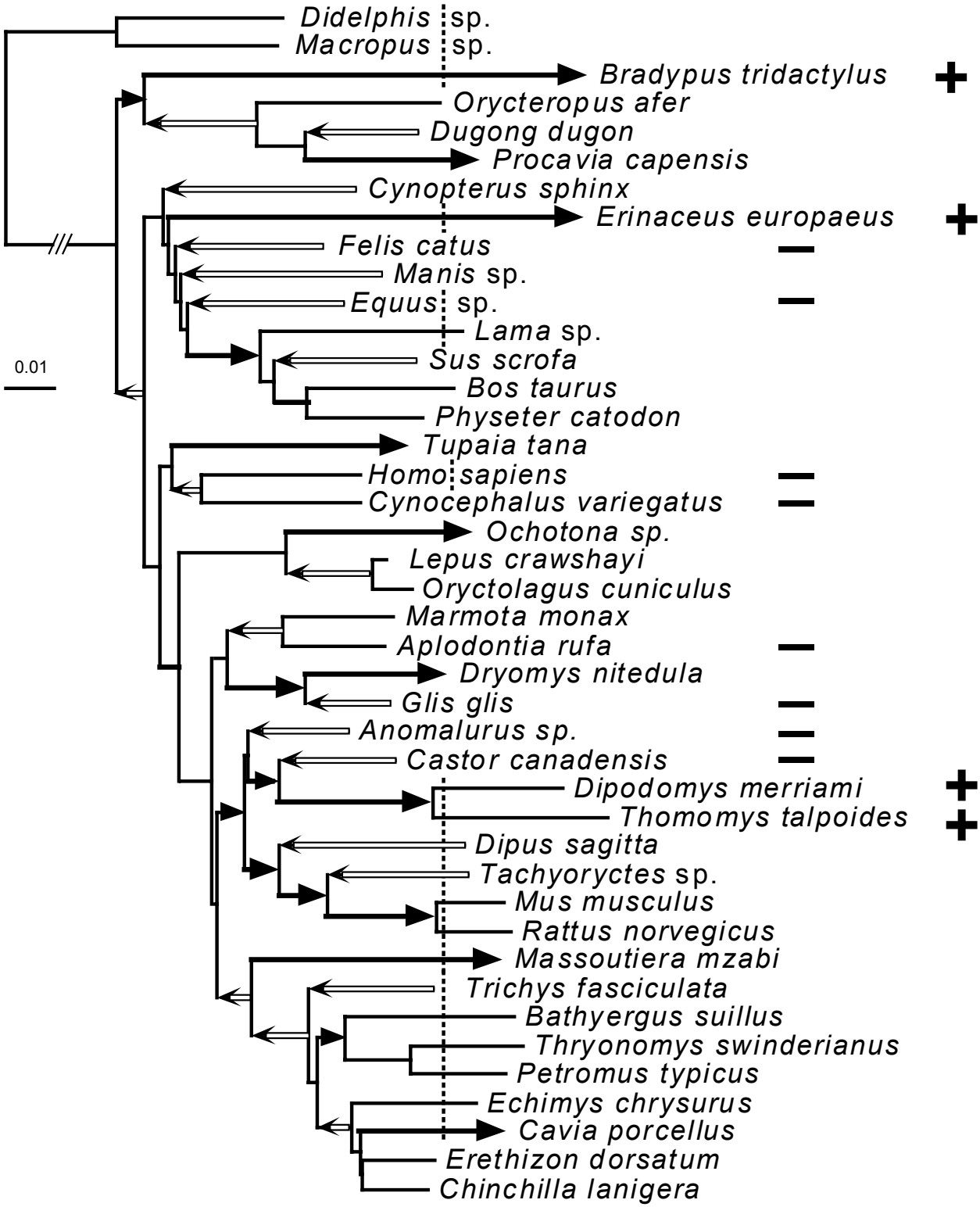
FIGURES

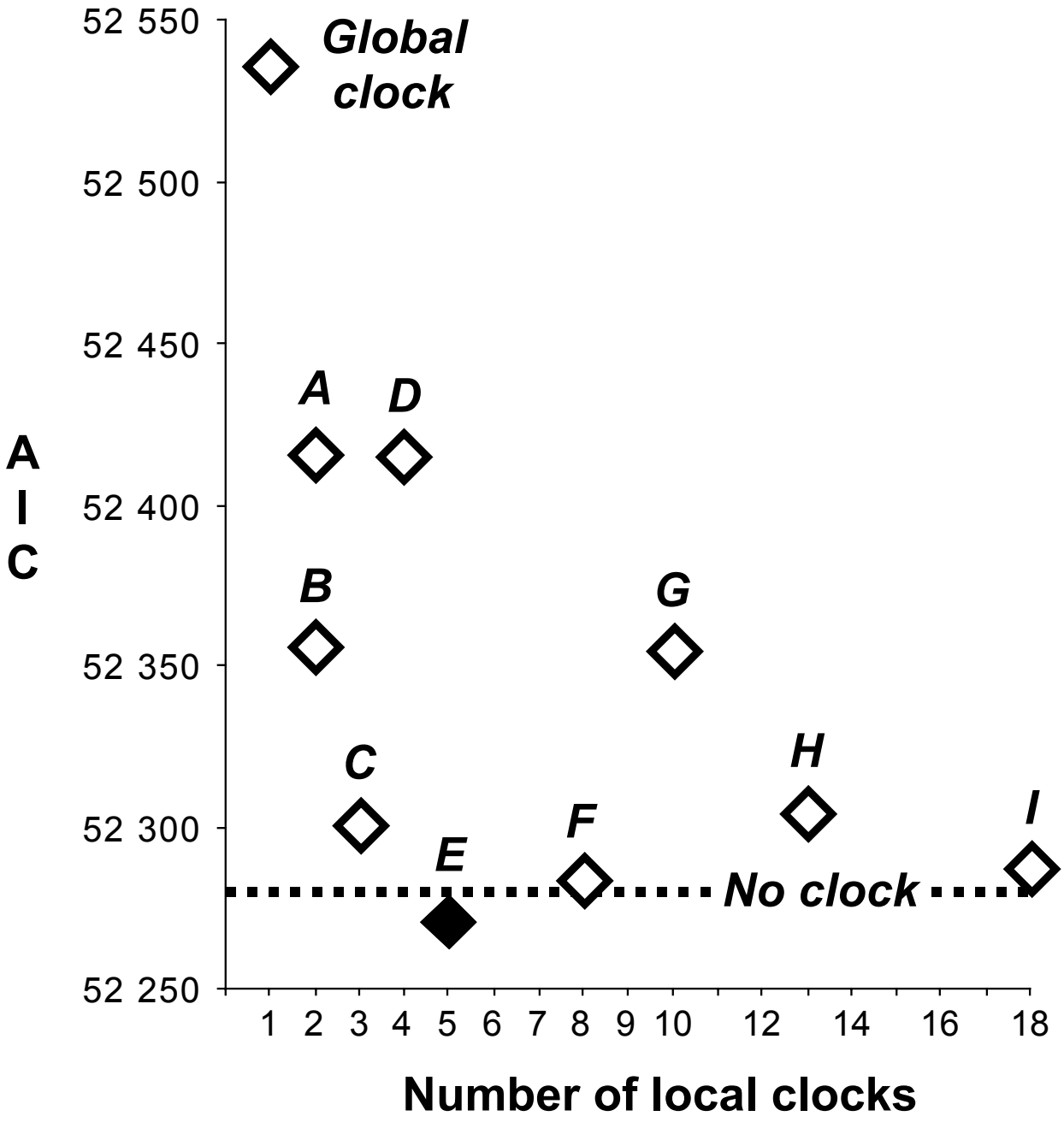
Figure 1. Extensive nucleotide substitution rate variations in the first two codon positions of the ADRA2B + IRBP + vWF nuclear genes between placental mammals. The vertical dashed line indicates the mean value of the root-to-tip distance of the 40 placental taxa. Significantly faster- or slower-evolving species are respectively indicated by a "+" or a "—" as evidenced by the branch length test. Significantly faster- and slower- evolving branches as evidenced by the two-cluster test are respectively indicated by black arrows towards the right and white arrows towards the left. The scale unit corresponds to expected number of nucleotide substitutions per site. The log-likelihood of this tree is lnL = –26054.36, and its AIC is 52282.78. In the clocklike constrained model—with a single global clock—a significant loss of log-likelihood is observed (lnL = -26222.37 ; AIC = 52538.74).

Figure 2. The relationship between the Akaike Information Criterion (AIC) and the number of local clocks. Nine different models of local clocks (**A** to **I**) have been evaluated according to their AIC. Open diamonds correspond to models for which the AIC values exceed the one of the model without any clock. The black diamond corresponds to the situation where the AIC is minimized (see also Figure 3). In addition to the default rate $r_D = 1.00$, the following combinations of local clocks (r_F for fast rates and r_S for slow rates) have been defined under maximum likelihood: **A** — $r_F = 1.82$ (a single very fast rate) ; **B** — $r_S = 0.57$ (a single very slow rate) ; **C** (combination of models **A** and **B**) — $r_F = 1.51$ and $r_S = 0.62$; **D** — $r_{F1\ (Erinaceus)} = 1.83$, $r_{F2\ (Dipodomys + Thomomys)} = 2.14$, $r_{F3\ (Bradypus)} = 1.61$ (different very fast rates) ; **E** (cf. Figure 3) — $r_{F1} = r_{F2} = r_{F3} = r_{F4\ (Procavia)} = r_{F5\ (Ochotona)} = r_{F6\ (Cavia)} = 1.54$ (a single very fast rate), $r_{F7\ (Massoutiera)} = 1.24$ (fast), $r_{S2\ (Felis)} = r_{S4\ (Equus)} = r_{S5\ (Homo + Cynocephalus)} = r_{S7\ (Glis)} = r_{S8\ (Anomalurus)} = r_{S9\ (Castor)} = 0.59$ (very slow), and $r_{S1\ (Cynopterus)} = r_{S3\ (Manis)} = r_{S6\ (Marmota + Aplodontia)} = 0.73$ (slow) ; **F** — $r_F = 1.55$ (a very fast rate), $r_S = 0.63$ (a very slow rate), and different intermediate rates: $r_{F4} = 1.43$, $r_{F5} =$

1.45, R_{F6} = 1.66, R_{F7} = 1.25, and $R_{F8\ (Tupaia)}$ = 0.91 ; **G** — different slow rates: R_{S1} = 0.61, R_{S2} = 0.48, R_{S3} = 0.69, R_{S4} = 0.55, R_{S5} = 0.60, R_{S6} = 0.64, R_{S7} = 0.51, R_{S8} = 0.41, and R_{S9} = 0.51 ; **H** (combination of models D and G) — different fast rates: R_{F1} = 1.49, R_{F2} = 1.51, and R_{F3} = 1.48, and different slow rates: R_{S1} = 0.69, R_{S2} = 0.52, R_{S3} = 0.74, R_{S4} = 0.62, R_{S5} = 0.63, R_{S6} = 0.69, R_{S7} = 0.54, R_{S8} = 0.42, R_{S9} = 0.57 ; **I** — additional different fast rates: R_{F1} = 1.52, R_{F2} = 1.57, R_{F3} = 1.52, R_{F4} = 1.43, R_{F5} = 1.45, R_{F7} = 1.25, and R_{F6} = 1.67, and additional different slow rates: R_{S1} = 0.70, R_{S2} = 0.53, R_{S3} = 0.76, R_{S4} = 0.63, R_{S5} = 0.65, R_{S6} = 0.72, R_{S7} = 0.56, R_{S8} = 0.44, R_{S9} = 0.59, and R_4 = 0.92.

Figure 3. Maximum likelihood phylogram with five local molecular clocks reconstructed from the first two codon positions of the ADRA2B + IRBP + vWF concatenated exons, and details of the seven fossil calibration points use for dating purposes. The local clocks (or rates: r) are indicated, with their maximum likelihood estimates. r_{VS}, r_S, r_D, r_F, and r_{VF} refer respectively to very slow (tight-dashed horizontal branches), slow (large-dashed branches), default (thin branches), fast (bold branches), and very fast rates (bolder branches). The log-likelihood of this tree is ln$\underline{L}$ = -26 085.97. Stars refer to the seven calibration points that span the whole placental tree: caviomorphs, Mus / Rattus, glirids, sciuroids, lagomorphs, cetartiodactyls, and paenungulates (C_1 to C_7 from bottom to top). The most recent common ancestors of five major embedded placental clades are indicated by arrows: R for Rodentia, G for Glires (i.e., rodents + lagomorphs), E for Euarchontoglires (Glires + Primates + Dermoptera + Scandentia), B for Boreoeutheria (euarchontoglires + laurasiatherians), and P for Placentalia.





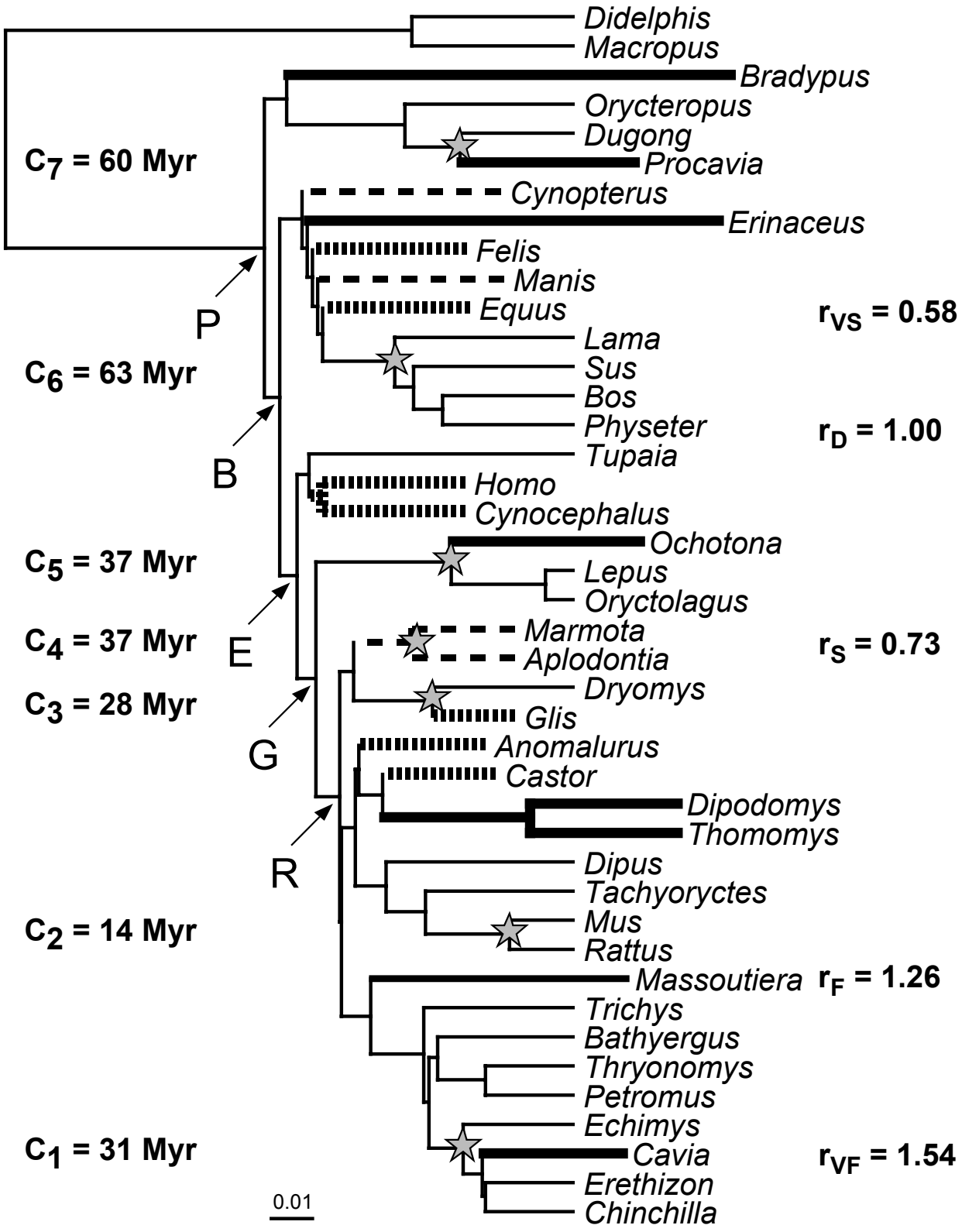


Table 1. Cross calibrations and reciprocal compatibility between seven paleontological references, as inferred by local molecular clock datings based on first and second codon positions of the three concatenated exons of ADRA2B + IRBP + vWF. Divergence dates in Myr are presented along with their standard-error in parentheses. Fossil calibration ages are given between square brackets. From top to bottom lines, divergences respectively correspond to *Echimys* vs. *Cavia*, *Mus* vs. *Rattus*, *Dryomys* vs. *Glis*, *Marmota vs. Aplodontia*, *Ochotona* vs. *Lepus*, *Lama* vs. *Physeter*, and *Dugong* vs. *Procavia*.

	CALIBRATION POINTS						
DIVERGENCES ESTIMATED	**Caviomorpha [31 Myr]**	***Mus* vs. *Rattus* [14 Myr]**	**Gliridae [28.5 Myr]**	**Sciuroidea [37 Myr]**	**Lagomorpha [37 Myr]**	**Cetartiodactyla [63 Myr]**	**Paenungulata [60 Myr]**
Caviomorpha	—	24.0 (1.6)	21.9 (1.5)	27.9 (1.9) *	32.7 (2.2) *	38.3 (2.6)	56.6 (3.9)
Murinae	18.1 (1.9)	—	12.8 (1.3) *	16.3 (1.7) *	19.1 (2.0)	22.3 (2.3)	33.0 (3.4)
Gliridae	40.4 (3.2)	31.3 (2.4) *	—	36.3 (2.8)	42.6 (3.3)	49.9 (3.9)	74.7 (5.8)
Sciuroidea	41.1 (3.5) *	31.8 (2.7) *	29.0 (2.4)	—	43.4 (3.6) *	50.8 (4.3)	75.1 (6.3)
Lagomorpha	35.1 (3.0) *	27.1 (2.3)	24.7 (2.1)	31.5 (2.7)	—	43.3 (3.7) *	64.0 (5.4)
Cetartiodactyla	51.0 (2.9)	39.5 (2.3)	36.0 (2.1)	45.8 (2.6)	53.8 (3.1)	—	93.0 (5.3)
Paenungulata	32.9 (2.9)	25.5 (2.2)	23.2 (2.0)	29.6 (2.6)	34.7 (3.0)	40.6 (3.6)	—

* Divergence dates which are accurately estimated by a given calibration point. For example, the 95%-confidence interval (age ± 1.96 standard-error) for the divergence time between the two sciuroids, Marmota and Aplodontia (41.1 ± 3.5 Myr), is 34.1 to 48.1 Myr as estimated under the Caviomorpha calibration at 31 Myr. This 34.1-48.1 Myr interval contains the fossil estimate for sciuroids (37 Myr).

Table 2. Local molecular clock datings of the age of the most recent common ancestor of six mammalian clades based on first and second codon positions of the three concatenated exons of ADRA2B + IRBP + vWF, and calibrated by seven paleontological references. Divergence dates are presented along with their standard-error in parentheses. For each of the six mammalian clades, italicized values correspond to the median of the seven divergence time estimates, and bold values correspond to the median of the five divergence age estimates after removal of the two non-compatible calibration points (cetartiodactyls and paenungulates).

	CALIBRATION POINTS						
DIVERGENCES ESTIMATED	**Caviomorpha [31 Myr]**	***Mus* vs. *Rattus* [14 Myr]**	**Gliridae [28.5 Myr]**	**Sciuroidea [37 Myr]**	**Lagomorpha [37 Myr]**	**Cetartiodactyla [63 Myr]**	**Paenungulata [60 Myr]**
Mus* vs. *Rattus	*18.1 (1.9)*	14.0 (1.4)	12.8 (1.3)	**16.3 (1.7)**	19.1 (2.0)	22.3 (2.3)	33.0 (3.4)
Rodentia	*66.3 (2.6)*	51.3 (2.0)	46.7 (1.8)	**59.6 (2.4)**	69.9 (2.8)	81.9 (3.2)	120.9 (4.8)
Glires	*73.4 (3.0)*	56.8 (2.3)	51.8 (2.1)	**66.0 (2.7)**	77.4 (3.2)	90.7 (3.7)	133.9 (5.5)
Euarchontoglires	*78.4 (3.1)*	60.7 (2.4)	55.3 (2.2)	**70.5 (2.8)**	82.7 (3.3)	96.9 (3.9)	143.1 (5.7)
Boreoeutheria	*82.9 (3.2)*	64.2 (2.5)	58.5 (2.3)	**74.5 (2.9)**	87.5 (3.4)	102.5 (4.0)	151.2 (5.9)
Placentalia	*87.3 (3.6)*	67.6 (2.8)	61.6 (2.5)	**78.5 (3.2)**	92.1 (3.8)	107.9 (4.4)	159.3 (6.5)