

# **NEW LIGHT ON THE DATES OF PRIMATE ORIGINS AND DIVERGENCE**

Christophe Soligo <sup>1</sup>, Oliver Will <sup>2</sup>, Simon Tavaré <sup>3</sup>,

Charles R. Marshall <sup>4</sup> and Robert D. Martin <sup>5</sup>

- 1 Human Origins Programme, Department of Palaeontology, The Natural History Museum, Cromwell Road, London SW7 5BD, UK. Tel: ++20 7460 3616; Fax: ++20 7942 5546.  
e-mail: [ayeaye@aim.unizh.ch](mailto:ayeaye@aim.unizh.ch) (corresponding author)
- 2 Statistics Department, University of Washington, Box 354322, Seattle, WA 98195-4322, USA. e-mail: [oliveran@stat.washington.edu](mailto:oliveran@stat.washington.edu)
- 3 Program in Molecular and Computational Biology, University of Southern California, 835 West 37th Street, SHS172, Los Angeles, CA 90089-1340, USA. e-mail: [stavare@usc.edu](mailto:stavare@usc.edu)
- 4 Department of Earth & Planetary Sciences, Harvard University, 20 Oxford Street, Cambridge, MA 02138, USA. e-mail: [marshall@eps.harvard.edu](mailto:marshall@eps.harvard.edu)
- 5 The Field Museum, 1400 S. Lake Shore Drive, Chicago, IL 60605-2496, USA. e-mail: [rdmartin@fieldmuseum.org](mailto:rdmartin@fieldmuseum.org)

## 1. Introduction

The known fossil record for undoubted primates of modern aspect (i.e. confined to Euprimates and excluding Plesiadapiformes) dates back to the beginning of the Eocene epoch, about 55 million years ago (mya), and it is widely accepted among primate paleontologists that primates originated during the preceding Paleocene epoch, some 60-65 mya. A parallel conclusion has been reached for most orders of placental mammals, and it is generally assumed that the origin and radiation of most if not all placental orders with extant representatives took place after the extinction of dinosaurs at the end of the Cretaceous. In common parlance, the Age of Mammals followed on from the Age of Dinosaurs. A comparable explanation has been given for the adaptive radiation of modern birds. All such interpretations depend on the common procedure of dating the origin of a group by the earliest known fossil representative, perhaps adding a safety margin of a few million years in tacit but conservative recognition of the fact that the earliest known fossil is unlikely to coincide exactly with the time of origin. Such direct dating from the fossil record faces two problems: (1) If the fossil record represents a very poor sample, the first known fossil representative of a given group is likely to be considerably more recent than the actual origin of that group. (2) Various kinds of bias in the fossil record may introduce further error. In this light, it has been suggested that a relatively low sampling level of the fossil record for primates has led to substantial underestimation of their time of origin (Martin 1986, 1990, 1993; Tavaré *et al.*, in press).

Correct timing of the initial emergence of a group such as the primates is of great importance if the mechanisms that led to its evolution are to be understood, as both biotic and abiotic environmental conditions can be taken into account only if the origin of the group and the prevailing environmental conditions can be accurately correlated chronologically.

In the present chapter, we review available paleontological and molecular evidence pertinent to the timing of the origin of the Primates. We also present new analyses using a recently developed statistical method that estimates times of origin of clades based on their modern diversity, their known fossil record, diversification models, and estimates of relative sampling intensities.

## **2. The Fossil Record**

Before proceeding any further, it is necessary to draw a crucial distinction between the time of initial divergence of a given group, such as the primates, and the age of the last common ancestor of all known, diagnosable members of that group (Fig. 1). In a phylogenetic tree, the initial time of origin of any given taxon is indicated by the point of divergence between that taxon and its most closely related sister taxon (node 1 in Figure 1). Initially, the taxon of interest might diverge from its closest relatives as a lineage lacking the characteristic morphological features of its later descendants and then exist for some time before developing recognizable diagnostic characters. A considerable temporal gap may therefore occur between the initial divergence of a taxon and the emergence of diagnostic morphological characteristics as recognized by paleontologists (i.e. between nodes 1 and 2 in Figure 1). With respect to the evolution of placental mammals, this point has been succinctly expressed by Madsen *et al.* (2001, p.613): "Easteal (1999) suggested that primitive placentals from the Cretaceous may have diversified phylogenetically before they diverged morphologically and acquired the diagnostic features of ordinal level crown-group clades." The upper limit for the temporal gap between the initial divergence of a taxon and the emergence of diagnostic morphological characteristics is set by the estimated age of the last common ancestor of modern lineages within the taxon (node 3 in Figure 1), or by the age of the oldest known clearly recognizable fossil representative of the taxon, whichever is older.

It should be noted that inferred phylogenetic relationships, in conjunction with the fossil record, may be used to extend minimum estimates of divergence times in some cases (Norell, 1992; Smith, 1994). Under the assumption that sister groups had the same time of origin, the later-appearing sister group is assumed to have existed at least by the time of first appearance of the earlier-appearing sister group. The range extension for the later-appearing sister group is referred to as a ghost lineage (Norell, 1992). In the case of the primates, the uncertainties that prevail regarding both the composition of and the relationships within Archonta – the supraordinal grouping to which primates are often allocated – make it difficult to apply the concept of the ghost lineage. It can be noted, however, that none of the modern orders of Archonta extends back much beyond the time of the earliest known primate fossils. The oldest known fossils belonging to Scandentia are from Eocene deposits (Tong, 1988) while the oldest fossils tentatively attributed to Volitantia (Dermoptera + Chiroptera) are Late Paleocene (Stucky and McKenna, 1993), which would extend the expected range of the primates back by no more than a few million years. Among extinct groups of archontans, the Plesiadapiformes and the Mixodectidae (as possible members of Dermoptera) are potentially relevant (Hooker, 2001). If confirmed to represent the sister group of primates, either of these would extend the expected range of primates back to the early Paleocene.

Undoubted primates (equated here with Euprimates) first appear in the fossil record at the beginning of the Eocene period in western Europe and North America. A reported primate from the Late Paleocene of Morocco (Sigé *et al.*, 1990), *Altiatlasius*, has recently been reassigned to the Plesiadapiformes (Hooker *et al.*, 1999) and is therefore not considered here. The absence from the known fossil record of any pre-Eocene primates of modern aspect is usually interpreted as evidence that the order originated not long before that period, around 60 mya and no earlier than 65 mya.

However, the ages of the first known fossil representatives of certain other mammalian groups are in themselves incompatible with the interpretation that the placental lineage leading to primates diverged only 60-65 mya. The best illustration of this is provided by studies of artiodactyl relationships. It has long been accepted that cetaceans and artiodactyls are sister-groups, but recent molecular evidence has uniformly indicated that cetaceans are actually nested within the artiodactyls and that their closest relatives are hippopotamuses. This conclusion, initially suggested by immunological data (Sarich, 1993), is now supported by nuclear gene sequences (Graur and Higgins, 1994; Gatesy *et al.*, 1996, 1999; Gatesy, 1997; Madsen *et al.*, 2001, Murphy *et al.*, 2001a, 2001b), by insertions of interspersed elements (retrotransposons) in the nuclear genome (Nikaido *et al.*, 1999), and by complete mitochondrial genomes (Ursing and Arnason, 1998). In fact, recent evidence from two early terrestrial relatives of cetaceans, *Ichthyolestes* and *Pakicetus* (Thewissen *et al.*, 2001), has confirmed that they share the unique tarsal morphology of artiodactyls and are therefore more closely related to them than to mesonychians, which were long thought to be the direct sister group of cetaceans. Although a cladistic analysis of the morphological data did not confirm a specific link between cetaceans and hippopotamuses, there is undoubtedly a closer link between cetaceans and artiodactyls than hitherto believed by paleontologists. The molecular evidence now uniformly indicates that the following sequence of divergences occurred during the evolution of the hoofed mammals (ungulates): (1) between odd-toed perissodactyls and even-toed artiodactyls; (2) within artiodactyls between camels+pigs and ruminants+hippos+cetaceans; (3) between ruminants and hippos+cetaceans; (4) between hippos and cetaceans. Given that the first known fossil representative of the cetaceans is dated to 54 mya (Bajpai and Gingerich 1998), it follows that the initial divergence in this well-supported sequence of 4 splits in ungulate evolution must have occurred at a relatively early date and that the separation between ungulates and the lineage leading to primates must have

taken place even earlier. A date of only 60-65 mya for the divergence of the primate lineage from other lineages of placental mammals hence seems inherently improbable. It seems likely, instead, that the early evolution of primates has simply remained undocumented in the known fossil record.

Early placental mammals seem to be generally poorly documented in the known fossil record. This is strikingly illustrated by the case of bats (order Chiroptera). Modern bats constitute a widespread and diverse group containing almost a thousand species, including about 165 megachiropterans (Old World fruit-bats) and some 815 microchiropterans (Corbet and Hill, 1991). As with primates of modern aspect, the earliest known clearly identifiable bat fossils date back to the beginning of the Eocene (about 55 mya) in North America, Europe, Africa, and Australia, although one report extends this back into the latest Paleocene, to 56 mya. The first relatively complete bat skeletons are known from Early Eocene deposits in North America (*Icaronycteris*) and from Early/Middle Eocene deposits in Europe (*Archaeonycteris*, *Hassianycteris*, and *Palaeochiropteryx*). By this time, all of the major defining morphological features of bats can be identified, notably including the development of a wing membrane (patagium) between digits II and V of the hand and extreme backward rotation of hindlimbs for suspension, involving extensive remodelling of the pelvis and ankle joint. Furthermore, all 4 Eocene bat genera that are documented by relatively complete skeletons show weak to moderate enlargement of the cochlea, indicating the development of some degree of echolocation capacity. For this and other reasons, a recent review of morphological evidence (Simmons and Geisler, 1998) concludes that these 4 genera are more closely related to microchiropterans than to megachiropterans and branched off successively from the lineage leading to the common ancestor of microchiropterans, such that they are an integral part of the adaptive radiation that led to modern bats. Yet there are no known fossils documenting the transition from a generalized early placental ancestor to the highly

specialized, immediately recognizable condition of the earliest known bat skeletons. Furthermore, there is an obvious and extreme bias in the geographical occurrence of well-preserved bat fossils. Whereas at least 4 skeletons of *Icaronycteris* have been reported from a single site in North America (Green River, Wyoming, approx. 53 mya), all the others (some 100 skeletons of *Archaeonycteris*, *Hassianycteris*, and *Palaeochiropteryx*) have been discovered at the European site of Messel, southern Germany (approx. 49 mya). With some of the exquisitely preserved bat skeletons from Messel, remains of the stomach contents are also present. Analysis of these has revealed moth wing scales indicating dietary habits comparable to those of modern microchiropteran bats.

The fossil record for Old World fruit-bats (megachiropterans) is even less informative. The earliest known remnant is a single tooth identified as that of a megachiropteran found in upper Eocene deposits of Thailand (Ducrocq *et al.*, 1993). Given that microchiropterans are reliably documented from the earliest Eocene, this indicates (at the very least) a ghost lineage of some 15 my prior to the earliest known megachiropteran.

Furthermore, a recent cladistic analysis of archontan relationships using both cranial and postcranial characters has provided evidence for a Cretaceous origin of bats (Hooker, 2001). In the cladogram issued from that study, bats branch off at a lower node than both the extinct genus *Deccanolestes* – a possible primitive Archontan – and the extinct family Nyctitheriidae. Therefore, the early Paleocene age of the oldest known nyctithere and the latest Cretaceous age of *Deccanolestes* imply that the divergence of bats from other known mammals occurred at least as long ago as the latest Cretaceous (Hooker, 2001).

Overall, it is obvious that there are very large gaps in the fossil record for bats. In particular, the transition to the shared morphology of all known bats is not documented at all.

### 3. The Molecular Evidence

Since 1994, evidence concerning the time of divergence between primates and other orders of placental mammals that conflicts with a direct reading of the known fossil record has been steadily accumulating from several independent studies of DNA sequence data. In a comparative analysis of the marsupial *Didelphis virginiana* and several placentals, taking sequence data for 8 mitochondrial genes with rates of evolution not significantly differing from a molecular clock model, a calibration date of 130 mya for the marsupial/placental divergence yielded a date of  $93 \pm 12$  mya for the divergence between human (representing primates) and a group representing carnivores, artiodactyls, and cetaceans (ferungulates) (Janke *et al.*, 1994). Subsequently, using sequence information for a large sample of nuclear genes showing relatively constant rates of change in mammals and birds, and taking a calibration date of 310 mya for the separation between diapsid and synapsid reptiles, divergence times between primates and artiodactyls and between primates and rodents were both estimated to be around 90 mya or older (Hedges *et al.*, 1996). In a follow-up study based on a larger sample of species and nuclear gene sequences, it was found that inferred molecular dates calibrated in this way agree with most early (Paleozoic) and late (Cenozoic) paleontological dates, but that major gaps are apparent in the Mesozoic fossil record. It was inferred that at least 5 lineages of placental mammals arose more than 100 mya and that most modern orders diverged before the end of the Cretaceous (Kumar and Hedges, 1998). On a separate tack, combined analysis of DNA sequences from 3 mitochondrial genes and 2 nuclear genes indicated that adaptive radiation from a specific common ancestor gave rise to a group of African mammals containing golden moles, hyraxes, manatees, elephants, elephant shrews, and aardvarks ("Afrotheria"). Using 9 different calibration points within the mammalian tree (including a date of 130 my for the marsupial/eutherian split and a date of 60 my for the ruminant/cetacean split), the mean divergence time between Afrotheria and other



orders of mammals (including primates) was estimated at about 90 mya (Springer *et al.*, 1997). In yet another approach, sequence data for the complete cytochrome b gene were used to generate a tree showing divergences between various mammal species, including 10 primates, and the tree was calibrated by taking a date of 60 mya for the split between artiodactyls and cetaceans. This calibration indicated that primates diverged from other orders of mammals at about 90 mya and that the split between haplorhine and strepsirrhine primates took place about 80 mya (Arnason *et al.*, 1996). The data set was subsequently expanded to include new sequence data for the baboon, and a double calibration based on the fossil record for ungulates was applied: 60 my for the divergence between artiodactyls and cetaceans and 50 my for the divergence between equids and rhinocerotids among perissodactyls. The time of divergence between ungulates and primates was estimated at 95 mya, while the split between strepsirrhines and haplorhines was confirmed to be in the region of 80 mya (Arnason *et al.*, 1998). These studies consistently indicate that primates diverged from other placental mammals about 90 mya.

A date of 90 mya for the divergence between primates and other placentals has received further consistent support from several very recent studies. A new statistical technique for handling the variation of the molecular clock between lineages was applied to complete mitochondrial genome sequences for 23 mammalian species. Using a calibration of 56.5 mya for the split between hippos and cetaceans, the method found a divergence time of 97.6 mya for primates from a sister clade containing Artiodactyla, Perissodactyla, and Carnivora (Huelsenbeck *et al.*, 2000). Another group of investigators constructed a phylogenetic tree for 26 placental taxa using up to 8665 bp of nuclear DNA. In supplementary information for their paper, they report only the time of the basal split of placental mammals at 111-118 mya using two calibration points: elephants and hyraxes splitting at 60 mya and hippos and cetaceans splitting at 55 mya. However, we can interpolate their figure and

conclude that their tree supports a primate divergence of approximately 90 mya (Madsen *et al.*, 2001). Subsequently, this data set was combined with that used in a parallel study (Murphy *et al.*, 2001a) to yield an overall sequence set of 16,397 bp and to generate a consensus phylogeny for placental mammals (Murphy *et al.*, 2001b). This combined study provided further confirmation for the existence of 4 superordinal groupings (Afrotheria, Xenarthra, Laurasiatheria, and Euarchontoglires, the latter including Primates). Afrotheria was the first of these groupings to diverge, at an estimated date of 103 mya, while the divergence between Laurasiatheria and Euarchontoglires was estimated at 79-88 mya.

It should be noted that all of the molecular trees cited were calibrated using the ages of various known fossil representatives of lineages external to the order Primates. Given that first recorded fossil representatives must in all cases indicate *minimum* dates for times of divergence, it is striking that a relatively consistent result emerges with respect to inference of the time of divergence of primates. (This is perhaps because comparatively well-documented parts of the mammalian tree were selected as sources of calibration dates). It should also be emphasized that the primary concern in calibration of molecular trees to date has been the time of divergence of primates from other orders of placental mammals (node 1 in Figure 1). There has been relatively little interest in dating the last common ancestor of living primates (node 3 in Figure 1), although genetic distances uniformly indicate that the temporal gap between the initial divergence of primates and their common ancestor must have been relatively small.

#### **4. Quantifying the Incompleteness of the Fossil Record**

As already noted above, the earliest known unequivocal fossil primates are of basal Eocene age (about 55 mya), and the standard view is that primates originated no earlier than

about 65 mya, close to and probably above the K/T-boundary, with their initial radiation following the extinction of the dinosaurs at the end of the Cretaceous.

Although the molecular evidence, when calibrated with various fossil dates outside the primate tree, consistently indicates that the lineage leading to living primates diverged from other placental mammal lineages about 90 my ago (node 1 in Figure 1), it is conceivable that the diagnostic features of known living and fossil primates did not emerge until some time after this divergence (node 2 in Figure 1), and that the last common ancestor of living primates (node 3 in Figure 1) may be even more recent. It might therefore be imagined that a species-poor lineage with barely differentiated morphological features did indeed diverge from other placental mammals some 90 mya, but did not lead to morphologically recognizable primates until 60-65 mya. This could potentially explain the disparity between the known fossil record and molecular-based estimates of the time of divergence between primates and other mammals. However, available molecular evidence concerning the first divergence among living primates, between strepsirrhines and haplorhines, indicates that it took place relatively soon after the primates diverged from other placental mammals. In what appear to be the only published calibrations of the first divergence among living primates, a date of about 80 mya is indicated (Arnason *et al.*, 1996, 1998). Hence, if the diagnostic morphological features shared by all living primates and their known fossil relatives can be attributed to common ancestry rather than to convergent evolution (as is generally assumed), these features must have been present at an early stage. If primates diverged from other placental mammals about 90 mya, the diagnostic features of the group must accordingly have been developed by about 80 mya, well before the end of the Cretaceous, and a major gap must therefore exist preceding the known fossil record. The extent of that gap may in part be due to the K/T mass extinction. A loss of taxa at the K/T boundary and the possibility that some taxa were slow to recover from that event might to some extent explain the difficulty encountered

in finding primates of modern aspect in the Paleocene. It should be noted that there is evidence indicating that biological recovery from major extinctions may take as long as 10 million years (Kirchner and Weil, 2000). However, in order to adequately interpret apparent discrepancies between molecular and fossil data it is necessary to develop methods that can quantitatively estimate degrees of incompleteness within the fossil record.

A simple calculation by Martin (1993) indicated that only 3% of extinct primate species have so far been documented. Rough correction for underestimation of the time of origin led to the inference that ancestral primates existed about 80 mya. This preliminary inference has now been confirmed by our newly developed statistical approach (Tavaré *et al.*, in press), which is based on an estimate of species preservation derived from a model of the diversification pattern of the analyzed group. The method takes into account the number of extant species, the mean species lifetime, the ages of the bases of the relevant stratigraphic intervals, the numbers of fossil species found in those intervals, and the relative sizes of the sampling intensities in each interval. It can be used to estimate either (1) the age of the last common ancestor of living primates, or (2) the age of the first morphologically recognizable primate. A logistic diversification model was chosen in which logistic growth is parameterized by the time at which diversity reached 90% of its present value. Various diversification models can be explored with our method, but logistic growth is the most biologically realistic model (Raup *et al.*, 1973), as it matches the general expectation of an equilibrium diversity level. The great diversity of Holarctic primates during the Eocene indicates that at least 90% of modern diversity would already have been reached by the Middle Eocene. Consequently logistic growth was parameterized at 49 mya. We used a mean species lifetime of 2.5 my, but our results were relatively insensitive to changes in this value.

Our approach is based on modeling the speciation process as a non-homogeneous Markov branching process with a specified diversification curve. This is a process in which

species live for a random amount of time, go extinct and are replaced by a random number of species. The lifetime of the species and the number of descendant species are not affected by any of the other species alive at that time. This is a commonly accepted model for the diversification of a clade (MacArthur and Wilson, 1963; Nee *et al.*, 1994; Kubo and Iwasa, 1995). The branching process allows us to compute the expected number of species alive in a given stratigraphic interval. Assuming that any species alive in such an interval can be fossilized and found with the same probability, we may calculate the expected number of species found as fossils in each stratigraphic interval. Our statistical method is based on matching the observed and expected number of fossil finds in each interval as closely as possible, and a parametric bootstrap approach is used to assess bias in the estimates and to find approximate confidence intervals.

Using this approach, we first determined an estimate of the age of the last common ancestor of living primates, (i.e. the time of divergence between strepsirrhines and haplorhines) as 81.5 mya, with an approximate 95% confidence interval of (72.0, 89.6) mya (Tavaré *et al.*, in press). This closely agrees with the only available molecular estimates of the strepsirrhine-haplorhine divergence (Arnason *et al.*, 1996, 1998).

The age of the last common ancestor of living primates thus determined corresponds to node 3 in Figure 1. It gives the minimum age for the presence of morphological characteristics considered to be shared derived features (autapomorphies) of primates of modern aspect, assuming that all known fossil primates of modern aspect belong within the phylogenetic tree for extant primates. The present consensus view is that the earliest known primates of modern aspect (Early Eocene adapiforms and omomyiforms) are sister groups of modern strepsirrhines and haplorhines, respectively. However, it is conceivable that the adapiforms and/or the omomyiforms diverged prior to the last common ancestor of modern primates. To allow for this possibility, we here extend our previously published analyses

(Tavaré *et al.*, in press) to estimate the time of the initial diversification of the primate clade (node 2 in Figure 1), which can be taken as the age of the first morphologically diagnosable primates of modern aspect. In our initial estimations of the time of divergence between living strepsirrhines and haplorhines, we considered only simulated trees in which an initial bifurcation led to living representatives on both sides. This implicitly incorporated the assumption that all known fossil primates of modern aspect are nested within the tree including all living primates. To allow for the possibilities that defining features of living primates might have emerged prior to their last common ancestor and that some fossil primates might have diverged prior to that ancestor, the analysis was repeated without the constraint of an initial bifurcation with surviving representatives on both sides of the tree.

In order to estimate the age of node 3 in Figure 1, we start the speciation process from 2 initial species, both leading to living descendants. However, to estimate the age of node 2 in Figure 1, the speciation process starts from a single species. We are then assuming that this first species and all its descendants would be identifiable as primates of modern aspect by a paleontologist. It is important to recognize that the combination of features distinguishing primates from their mammalian relatives – and probably distinguishing the first primates from earlier ancestors in the lineage leading to them – are unlikely to have evolved simultaneously. As a result, designation of the first morphologically recognizable primates on a temporal scale can only be hypothetical, and the estimate of their age is an approximate indication of when the acquisition of primate characteristics took place.

Repeating the model specifications that were used to estimate the age of the strepsirrhine-haplorhine divergence, the age of the first morphologically recognizable primates (node 2 in Figure 1) is estimated at 85.9 mya, with a 95% confidence interval of (73.3, 95.7) mya. Note that the estimate for the strepsirrhine-haplorhine divergence is only 4.5 my younger than the best estimate for age of the first morphologically recognizable primate.

If the notion of a first morphologically recognizable primate provokes discomfort, it is reassuring to know that the relatively short time span between this construct and the last common ancestor of living primates allows one to use the age of the first morphologically recognizable primate as a proxy for the age of the last common ancestor of living primates.

Therefore, for the most realistic model settings, i.e. assuming a logistic growth model with 90% of modern diversity reached by the base of the Middle Eocene, our estimates of the emergence and subsequent diversification of primates of modern aspect are in broad agreement with molecular estimates of divergence times (Table 1). Other diversification models such as linear or exponential growth, as well as parameterization of the logistic growth with more recent dates, all result in age estimates for the presence of the first morphologically recognizable primates that are even older.

In stark contrast with our results, Gingerich and Uhen (1994) argued, on the basis of a formalization of Martin's (1993) heuristic approach, that there is only a 5 in a billion chance ( $5 \times 10^{-9}$ ) that primates originated 80 mya, and that, at a 95% confidence level, the origin of primates was located somewhere between 55 and 63 mya. Using our updated data on the number of fossil primate species, the probability that primates originated 80 mya calculated in this way in fact declines even further to a mere  $2 \times 10^{-18}$ . However, although modern species diversity is initially entered into the model by Gingerich and Uhen (1994), it eventually falls out of the equation that is applied. The results of the calculation are the same regardless of the number of modern primate species or estimated preservation rates and are as such based solely on the existing fossil record. As a consequence, their model is set to return the highest probability for the scenario in which the time of origin of a group is equal to the age of the oldest known fossil of that group. It therefore simply states that the more a scenario differs from a direct reading of the existing fossil record, the less likely it is to be real, thus entering precisely the kind of circularity which we have aimed to eliminate (Tavaré *et al.*, in press).

The problem with such an approach can be illustrated by applying the method of Gingerich and Uhen (1994) to the complete gap that exists in the primate fossil record during the middle Oligocene. That gap, between the Fayum primates of the early Oligocene and the earliest occurrence of platyrrhines in the fossil record of South America in the late Oligocene, is likely to cover around 6 my. Application of the method of Gingerich and Uhen (1994) yields a vanishingly small probability of  $2 \times 10^{-19}$  that primates existed during that gap.

In an analysis of evolutionary and preservational constraints on the times of divergence of eutherian mammals, Foote *et al.* (1999) concluded that molecular estimates of the times of origin of the living eutherian orders could be correct only if the preservation potential per lineage per million years was at least an order of magnitude smaller than it appeared to be. They consequently argued that it was unlikely for these ordinal divergences to have occurred as deep in the Cretaceous as the molecular clock data suggest. This conclusion, however, is not matched by our analyses of the fossil record of primates (Tavaré *et al.*, in press). The reason for this discrepancy seems to lie in the estimated preservation potential of mammalian lineages. Foote *et al.* estimate the preservation potential for Cenozoic mammals to be between 0.25 and 0.37/lineage/my (Foote, 1997, Foote *et al.*, 1999), and that of Cretaceous mammals to be 0.03. Significantly, the average values for the preservation potential based on our approach are 0.023/lineage/my for the known fossil record of primates, and 0.003/lineage/my for the time prior to the first known fossils. These values are, in fact, an order of magnitude smaller than those determined by Foote *et al.* (1999). It thus seems that our two very different methods of analyses of the fossil record are not in conflict; where we differ is in the estimated preservability of taxa.



## 5. Preservation Bias in the Fossil Record

There are several reasons why the preservation rates calculated by Foote *et al.* (1994) are likely to be overestimates. These all relate to the problem of circularity when interpreting the completeness of the fossil record through analysis of the fossil record alone. First, methods for assessing the completeness of the fossil record based exclusively on the fossil record can only account for gaps that occur within known lineages. They are insensitive to the existence of larger gaps, both chronological and geographical, and will overestimate completeness where such gaps occur. Foote (1997) demonstrated that the method used by Foote *et al.* (1999) will overestimate preservation potential where chronological gaps occur, with larger gaps within a given chronological range resulting in a larger overestimation. Even simple temporal variation in preservation probability will in most cases cause a slight to moderate overestimation of completeness (Foote, 1997). The primate fossil record as a whole has two large gaps. One, already noted above, extends over a period of about 6 my during the middle Oligocene. The other is the gap between the origin of the order and its first fossil appearance, a gap which most would agree to be at least 5 to 10 my and which we estimate to be over 25 my. More gaps become apparent when individual lineages are considered. In the most dramatic primate example, documentation of Malagasy lemurs was, until very recently, strictly limited to subfossils just a few thousand years old. Yet it was known that lemurs must have existed much earlier, as the sister-group (lorisiforms) is documented by fossils that are at least 20 my old (Szalay and Delson, 1979), and possibly over 30 my old (Simons, 1995), thus documenting a ghost lineage (Norell, 1992) for lemurs extending at least that far back in time. Very recently, a strepsirrhine primate (*Bugtilemur*) interpreted as a possible relative of the lemur family Cheirogaleidae has been recovered from Early Oligocene deposits of Pakistan (Marivaux *et al.*, 2001). Rather than closing a gap, however, this new find illustrates just how little may be known about key aspects of primate evolution. The lemurs are a diverse group of

modern primates known, until now, exclusively from the island of Madagascar. To explain the presence of a lemur in the Oligocene of Pakistan combined with the, as yet, total absence of fossil lemurs from anywhere else in the world, requires the contemplation of some fairly elaborate biogeographical scenarios.

Substantial geographical gaps are, in fact, likely to be the rule during the earlier phases of primate evolution. Living primates are essentially confined to tropical and subtropical climates (Martin 1990; Figure 2a). Support for the inference that this was also true in the past comes from the fact that primates only ever populated substantial parts of the northern continents when these areas supported subtropical climates at times of markedly increased global temperatures, during the Eocene and the Miocene. Yet 47% of all known fossil primate species come from restricted areas of North America and Europe and, for the first half of paleontologically documented primate evolution, sites yielding fossil primates are largely restricted to these two regions (Figure 2c). A direct reading of the known fossil record would suggest that primates originated some time during the Paleocene in the northern continents and subsequently migrated southwards. An alternative interpretation is that primates originated earlier in the relatively poorly documented southern continents and expanded northwards when climatic conditions permitted during the Eocene and, to a lesser degree, during the Miocene. The preservation rates proposed by Foote *et al.* (1999) for modern eutherian mammals as a whole are based either entirely (for the Cenozoic rates) or to more than three quarters (for the Cretaceous rates) on North American faunas. North America is the best sampled region in the world, and estimates based on that region will necessarily overestimate the preservation rates of groups with an almost worldwide distribution.

Our method also implies that approximately 5% and no more than 7% of all primate species that have ever existed are known from the fossil record. This low value does not seem unrealistic, as only 6-7% of all living primate species are known from the fossil record, a

record that is expected to be better than the average, given that it is dominated by easily collected and relatively common Pleistocene sediments. In addition, the belief underlying any direct reading of the fossil record – namely that most of primate evolution has by now been unearthed and described – is easily refuted by the ongoing rate of publication of new species of fossil primates (Figure 3).

To produce a precise estimate for the time of origin of primates (or any clade) using our method requires knowledge of the true diversity curve of the clade, of the relative sampling intensities of each stratigraphic interval, and of the mean species longevity (although the first is the most important in influencing the resulting estimates). As the values of these parameters are not known in detail, our estimates of the time of origin of primates must remain provisional. It is significant, however, that a number of our models produce dates concordant with various molecular estimates using calibration with fossil dates outside the primate tree (Janke *et al.*, 1994; Hedges *et al.*, 1996; Kumar and Hedges, 1998; Springer *et al.*, 1997; Arnason *et al.*, 1996, 1998; Huelsenbeck, 2000; Eizirik *et al.*, 2001; Murphy *et al.*, 2001b).

The poor sampling that we have inferred for the primate fossil record is unlikely to be restricted to that group. Cretaceous divergence times for primates and other modern orders of mammals should now be considered the more likely scenario, in which case the influence of continental drift has probably been considerable (Martin, 1990; Eastal *et al.*, 1995; Hedges *et al.*, 1996; Murphy *et al.*, 2001b). Clearly, fossil evidence from appropriate regions is needed to test this proposition. In the case of primates, it can be predicted that early members of the order showing characteristic morphological features lived somewhere in the southern continents (i.e. on parts of the former Gondwanaland) approximately 85 mya.

## Acknowledgments

We thank Marian Dagosto and Matt Ravosa for inviting us to contribute to this volume and for organizing the conference on primate origins. ST and OW were supported in part by NSF grant DBI 95-04393. CRM was supported in part by NSF grant, EAR-9258045, by NASA grant NCC2-1053 and by American Chemical Society Petroleum Research Fund Grant, 31316-AC8. CS was supported by Janggen-Poehn-Stiftung, by the G. & A. Claraz Schenkung and by a grant to RDM from the A.H. Schultz-Stiftung.

## References

- Arnason, U., Gullberg, A., Janke, A., and Xu, X.-f. (1996). Pattern and timing of evolutionary divergences between hominoids based on analyses of complete mtDNAs. *J. mol. Evol.* 43:650-661.
- Arnason, U., Gullberg, A., and Janke, A. (1998). Molecular timing of primate divergences as estimated by two nonprimate calibration points. *J. mol. Evol.* 47:718-727.
- Bajpai, S., and Gingerich, P. D. (1998). A new Eocene archaeocete (Mammalia, Cetacea) from India and the time of origin of whales. *Proc. Natl. Acad. Sci. USA* 95: 15464–15468.
- Corbet, G. E., and Hill, J. E. (1991). *A World List of Mammalian Species*, 3<sup>rd</sup>. edition, Oxford University Press, Oxford.

- Ducrocq, S., Jaeger, J.-J., and Sigé, B. (1993). Un mégachiroptère dans l'Eocène supérieur de Thaïlande: Incidence dans la discussion phylogénique du groupe. *Neues Jahrb. Geol. Paläontol. Monatschr.* 9:561-575.
- Easteal, S., Collett, C., and Betty, D. (1995). *The Mammalian Molecular Clock*, R.G. Landes, Austin, Texas.
- Eizirik, E., Murphy, W. J., and O'Brien, S. J. (2001). Molecular dating and biogeography of the early placental mammal radiation. *J. Hered.* 92:212-219.
- Foote, M. (1997). Estimating taxonomic durations and preservation probability. *Paleobiology* 23:278-300.
- Foote, M., Hunter, J. P., Janis, C. M., and Sepkoski, J. J. Jr. (1999). Evolutionary and preservational constraints on origins of biological groups: divergence times of eutherian mammals. *Science* 283:1310-1314.
- Gatesy, Y. (1997). More DNA support for a Cetacea/Hippopotamidae clade: the blood-clotting protein gene g-fibrinogen. *Mol. Biol. Evol.* 14:537-543.
- Gatesy, Y., Hayashi, C., Cronin, M. A., and Arctander, P. (1996). Evidence from milk casein genes that cetaceans are close relatives of hippopotamid artiodactyls. *Mol. Biol. Evol.* 13:954-963.

- Gatesy, J., Milinkovitch, M., Waddell, V., and Stanhope, M. (1999). Stability of cladistic relationships between Cetacea and higher-level artiodactyl taxa. *Syst. Biol.* 48:6-20.
- Gingerich, P. D., and Uhen, M. D. (1994). Time of origin of primates. *J. Hum. Evol.* 27:443-445.
- Graur, D., and Higgins, D. G. (1994). Molecular evidence for the inclusion of cetaceans within the order Artiodactyla. *Mol. Biol. Evol.* 11:357-364.
- Hedges, S. B., Parker, P. H., Sibley, C. G., and Kumar, S. (1996). Continental breakup and the ordinal diversification of birds and mammals. *Nature, Lond.* 381:226-229.
- Hooker, J.J. (2001). Tarsals of the extinct insectivoran family Nyctitheriidae (Mammalia): evidence for archontan relationships. *Zool. J. Linn. Soc.* 132:501-529.
- Hooker, J. J., Russell, D. E., and Phélizon, A. (1999). A new family of Plesiadapiformes (Mammalia) from the Old World Lower Paleogene. *Palaeontology* 42:377-407.
- Huelsenbeck, J. P., Larget, B., and Swofford, D. (2000). A compound Poisson process for relaxing the molecular clock. *Genetics* 154:1879-1892.
- Janke, A., Feldmaier-Fuchs, G., Thomas, W. K., von Haeseler, A., and Pääbo, S. (1994). The marsupial mitochondrial genome and the evolution of placental mammals. *Genetics* 137:243-256.

- Kirchner, J. W., and Weil, A. (2000). Delayed biological recovery from extinctions throughout the fossil record. *Nature, Lond.* 404:177-180.
- Kubo, T., and Iwasa, Y. (1995) Inferring the rates of branching and extinction from molecular phylogenies. *Evolution* 49:695-704.
- Kumar, S., and Hedges, S. B. (1998). A molecular timescale for vertebrate evolution. *Nature, Lond.* 392:917-920.
- MacArthur, R. H., and Wilson, E. O. (1963) An equilibrium theory of insular zoogeography. *Evolution* 17:373-387.
- Madsen, O., Scally, M., Douady, C. J., Kao, D., DeBry, R.W., Adkins, R., Armbrave, H. M., Stanhope, M., de Jong, W.W., and Springer, M. S. (2001) Parallel adaptive radiations in two major clades of placental mammals. *Nature, Lond.* 409:610-614.
- Marivaux, L., Welcomme, J.-L., Antoine, P.-O., Métais, G., Baloch, I. M., Benammi, M., Chaimanee, Y., Ducrocq, S., and Jaeger, J.-J. (2001). A fossil lemur from the Oligocene of Pakistan. *Science* 294:587-591.
- Martin, R. D. (1986). Primates: a definition. In: Wood, B. A., Martin, L. B., and Andrews, P. (eds.), *Major Topics in Primate and Human Evolution*, Cambridge University Press, Cambridge; pp 1-31.

- Martin, R.D. (1990). *Primate Origins and Evolution: A Phylogenetic Reconstruction*, Princeton University Press, New Jersey).
- Martin, R. D. (1993). Primate origins: plugging the gaps. *Nature, Lond.* 363:223-234.
- Murphy, W. J., Eizirik, E., Johnson, W. E., Zhang, Y. P., Ryder, O. A., and O'Brien, S. J. (2001a). Molecular phylogenetics and the origins of placental mammals. *Nature, Lond.* 409:614-618.
- Murphy, W. J., Eizirik, E., O'Brien, S. J., Madsen, O., Scally, M., Douady, C. J., Teeling, E., Ryder, A. O., Stanhope, M. J., de Jong, W. W., and Springer, M. S. (2001b). Resolution of the early placental mammal radiation using Bayesian phylogenetics. *Science* 294:2348-2351.
- Nee, S., May, R. M., and Harvey, P. H. (1994) The reconstructed evolutionary process. *Phil. Trans. R. Soc. Lond. B* 344:305-311.
- Nikaido, M., Rooney, A. P., and Okada, N. (1999). Phylogenetic relationships among cetartiodactyls based on insertions of short and long interspersed elements: Hippopotamuses are the closest extant relatives of whales. *Proc. Natl. Acad. Sci. USA* 96:10261-10266.
- Norell, M. A. (1992). Taxic origin and temporal diversity: the effect of phylogeny. In: Novacek, M. J., and Wheller, Q. D. (eds.), *Extinction and Phylogeny*, Columbia University Press, New York, pp. 89-118.



- Raup, D. M., Gould, S. J., Schopf, T. M., and Simberloff, D. S. (1973) Stochastic models of phylogeny and the evolution of diversity. *J. Geol.* 81:525-542.
- Sarich, V. M. (1993). In: Szalay, F. S., Novacek, M. J., and McKenna, M. C. (eds.). *Mammal Phylogeny. Volume 1: Placentals*, Springer-Verlag, Berlin, pp. 103-114.
- Sigé, B., Jaeger, J.-J., Sudre, J., and Vianey-Liaud, M. (1990). *Altiaatlasius koulchii* n.gen. et sp., primate omomyidé du Paléocène supérieur du Maroc, et les origines des Euprimates. *Palaeontographica. Abt. A* 214:31-56.
- Simmons, N. B., and Geisler, J. H. (1998). Phylogenetic relationships of *Icaronycteris*, *Archaeonycteris*, *Hassianycteris*, and *Palaeochiropteryx* to extant bat lineages, with comments on the evolution of echolocation and foraging strategies in Microchiroptera. *Bull. Amer. Mus. nat. Hist.* 235:1-182.
- Simons, E. L. (1995). Egyptian Oligocene primates: a review. *Yb. Phys. Anthropol.* 38:199-238.
- Smith, A. B. (1994). *Systematics and the Fossil Record*, Blackwell Scientific, Oxford.
- Springer, M. S., Cleven, G. C., Madsen, O., de Jong, W. W., Waddell, V. G., Amrine, H. M., and Stanhope, M. J. (1997). Endemic African mammals shake the phylogenetic tree. *Nature, Lond.* 388:61-64.

Stucky, R. K., and McKenna, M. C. (1993). In: Benton, M. J. (ed.), *The Fossil Record II*. Chapman & Hall, London, pp. 739-771.

Szalay, F. S., and Delson, E. (1979). *Evolutionary History of the Primates*, Academic Press, New York & London.

Tavaré, S., Marshall, C. R., Will, O., Soligo, C., and Martin, R. D. (in press). Estimating the age of the last common ancestor of extant primates using the fossil record. *Nature, Lond.*

Thewissen, J. G. M., Williams, E. M., Roe, L. J., and Hussain, S. T. (2001). Skeletons of terrestrial cetaceans and the relationship of whales to artiodactyls. *Nature, Lond.* 413:277-281.

Tong, Y-s. (1988). Fossil tree shrews from the Eocene Hetaoyuan Formation of Xichuan, Henan. *Vert. PalAsiat.* 26:214-220.

Ursing, B. M., and Arnason, U. (1998). Analyses of mitochondrial genomes strongly support a hippopotamus-whale clade. *Proc. Roy. Soc. Lond. B* 265:2251-2255.

Table 1: Molecular and paleontological estimates of divergence and diversification times during early primate evolution. Paleontological estimates are derived from a statistical approach developed by Tavaré *et al.* (in press). Estimated nodes refers to the nodes in Figure 1. Molecular estimates are from Janke *et al.* (1994), Hedges *et al.* (1994), Kumar & Hedges (1998), Springer *et al.* (1997), Arnason *et al.* (1996, 1998).

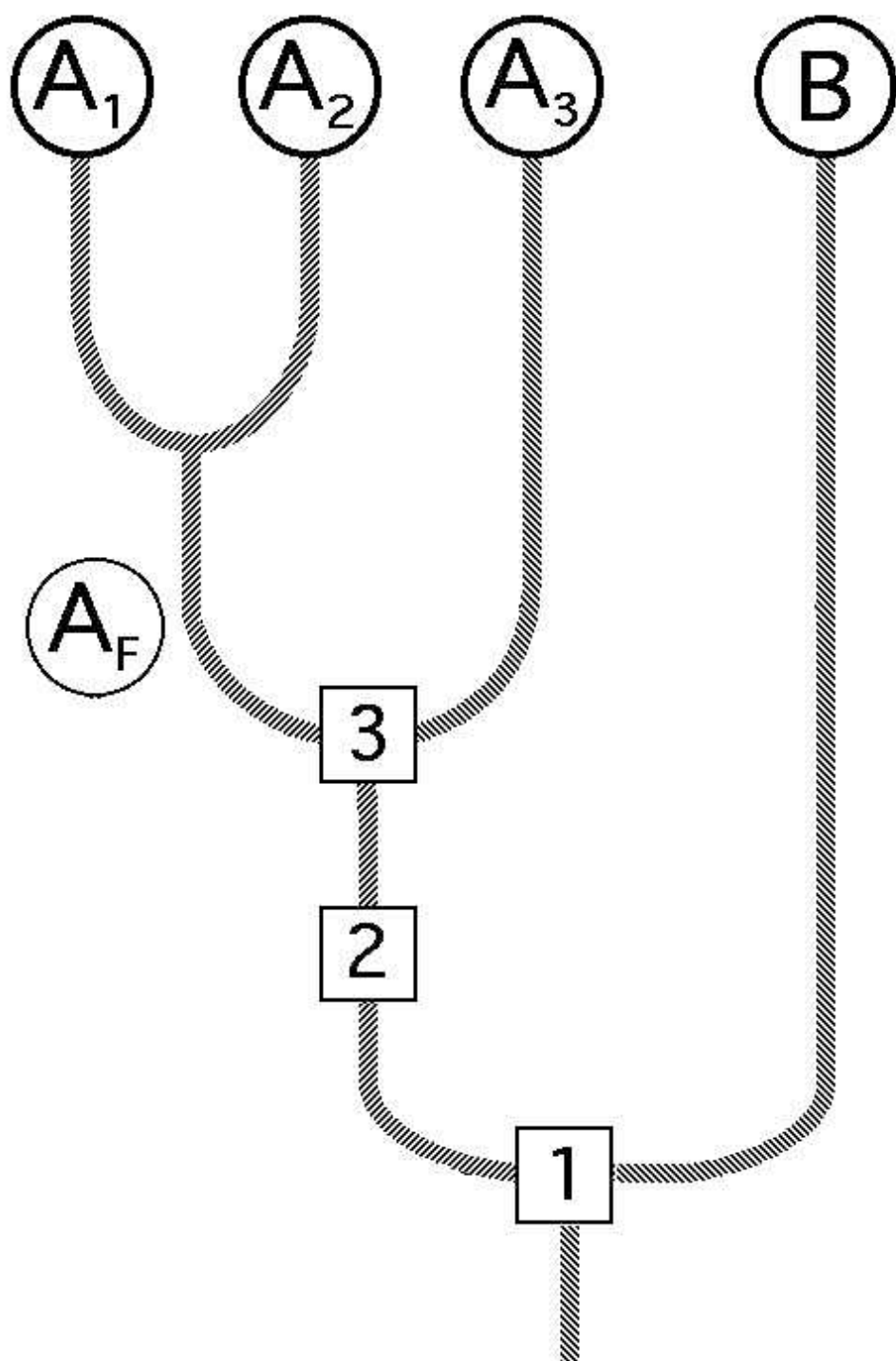
Estimated Node	Molecular estimates	Paleontological estimates
<b>Node 1:</b> divergence of the primate lineage from other modern mammals	~ 90 mya	NA
<b>Node 2:</b> initial diversification of primates/ first morphologically recognizable primates	NA	85.9 mya
<b>Node 3:</b> Divergence of strepsirrhines and haplorhines/ last common ancestor of living primates	~ 80 mya	81.5 mya

## Figure Captions

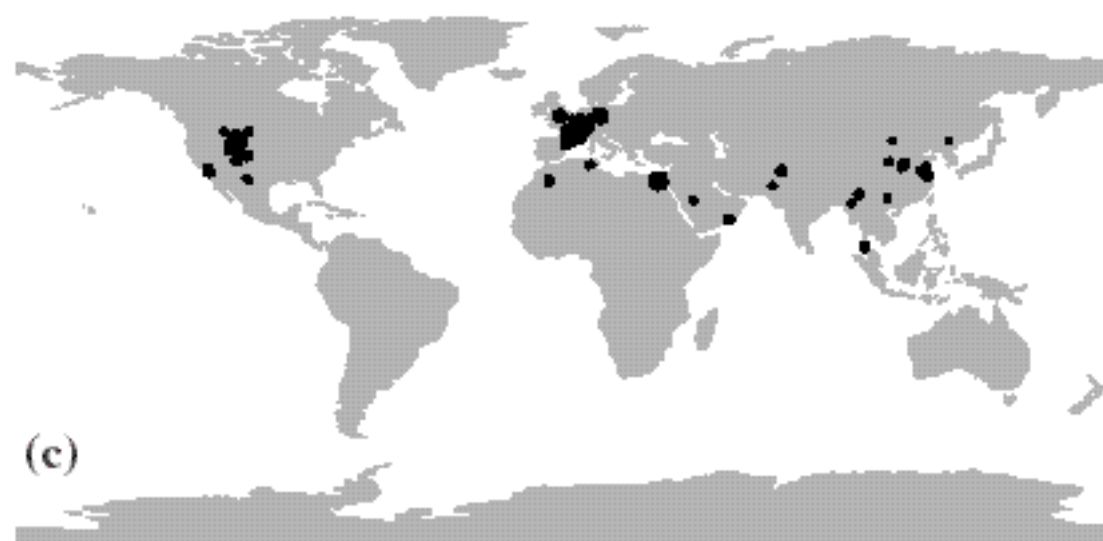
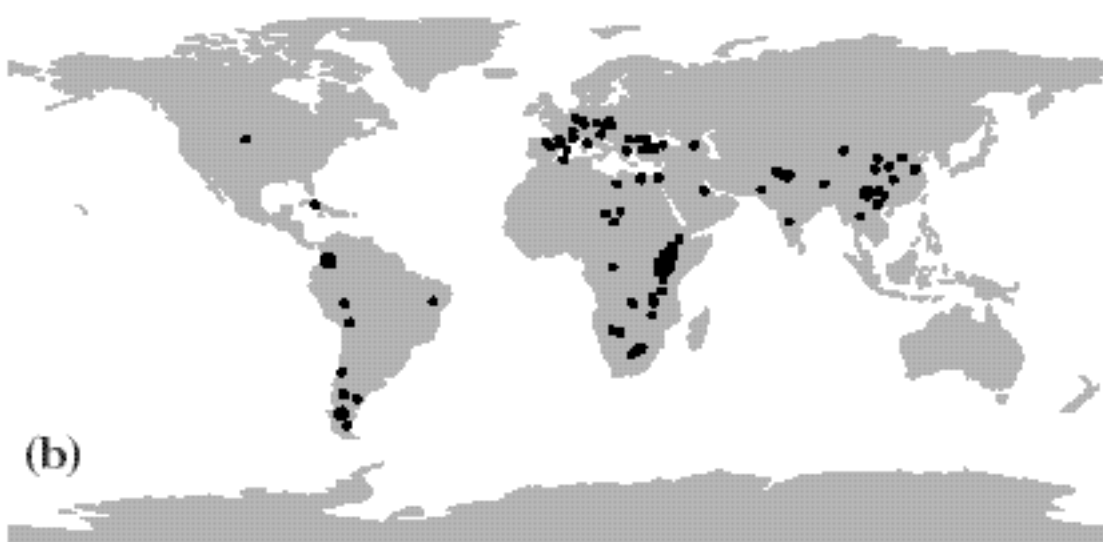
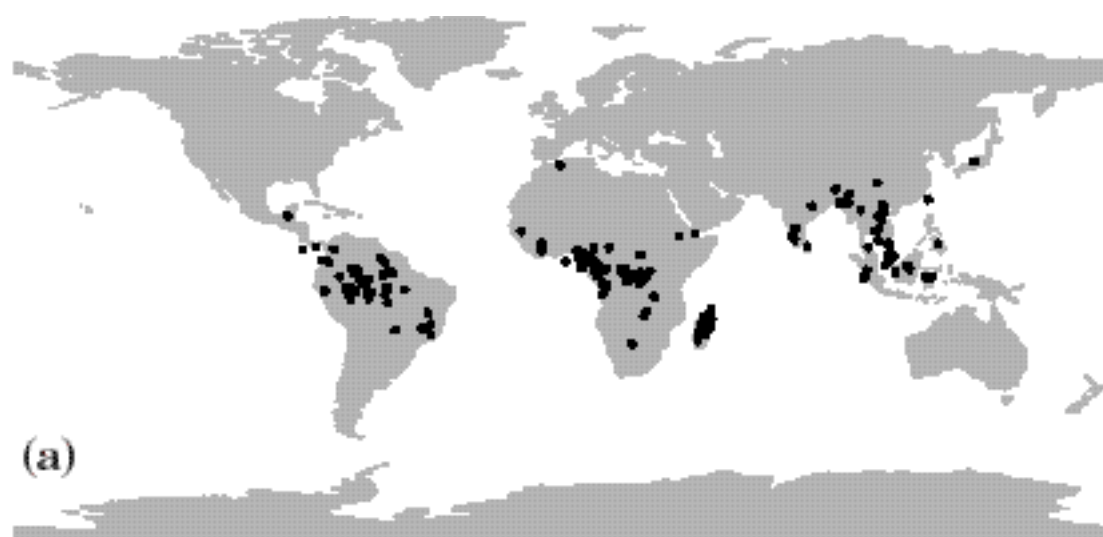
**Figure 1:** In a molecular phylogeny, the time of origin of taxon A (with living representatives  $A_1$ ,  $A_2$ , and  $A_3$ ) is indicated by node 1, the point of inferred divergence from the most closely related sister taxon with living representatives (B). The time of initial divergence of living representatives of taxon A from their last common ancestor may be considerably younger, as indicated by node 3. Molecular estimates can also be used to infer the date of node 3, in this case the time of divergence between  $A_1$  and  $(A_2+A_3)$ . Derived morphological features shared by the living representatives of taxon A may have developed at any time between node 1 and node 3. The earliest morphologically recognizable member of taxon A exhibiting derived diagnostic features shared with the living representatives is indicated by node 2. The first known fossil representative allocated to taxon A ( $A_F$ ), on the basis of derived features shared with living representatives, yields a minimum date for the origin of the taxon. It should be noted that  $A_F$  may be nested within the adaptive radiation leading to living representatives (as is widely presumed to be the case for Eocene adapiforms and omomyiforms), but it is also possible that  $A_F$  diverged at some time prior to the common ancestor of living representatives (i.e. prior to node 3).

**Figure 2:** Geographical distribution of individual modern and fossil euprimate species, taking the mid-range point in each case and plotting in relation to present-day continental positions (updated from Tavaré *et al.*, in press): (a) modern and subfossil primates; (b) fossil species between the Late Pleistocene and the Late Oligocene; (c) fossil species between the Early Oligocene and the Early Eocene. (b) and (c) are separated by a fossil-free gap of 6 million years. Note the progressive southward shift from (c) to (a).

**Figure 3:** Histogram showing the numbers of new species of fossil primates by year of publication, grouped into decades.









Soligo *et al.* Figure 3

