

Inferring macro-evolutionary patterns using an adaptive peak model of evolution

Jeroen B. Smaers and Lucio Vinicius

*Leverhulme Centre for Human Evolutionary Studies,
Department of Biological Anthropology,
University of Cambridge, Cambridge, UK*

ABSTRACT

Background: Currently available methods of ancestral reconstruction are built either on Brownian Motion (BM) or Ornstein-Uhlenbeck (OU) models of evolution. Results from these methods sometimes do not agree with the fossil record.

Aim: Develop a new method, the method of Independent Evolution (IE), built on an adaptive peak model of evolution.

Key assumptions: In evolution, population phenotypes are affected by the wandering adaptive peaks of adaptive surfaces. Branch-specific rates of evolution depend on weighted, relative distances between an ancestral adaptive peak and a new adaptive surface that descendant populations attempt to climb.

Methods: We defined an eight-step algorithm to incorporate the assumptions, and then applied it recursively for each node in the tree to produce the phylogenetic histories. We performed two studies: a simulation study of directional selection on particular branches in a model primate phylogeny, and a case study of primate brain and body size in which we reconstructed the ancestral states of primate brain sizes and body sizes and compared them with 28 fossil data points. In each study, we used strong inference – that is, we employed six different methods of ancestral reconstruction to determine how well each succeeds. Four of the methods of ancestral reconstruction are based on BM, one on the OU, and one is IE.

Results: The method based on an adaptive peak model of evolution (IE) significantly outperformed both BM-based and OU-based methods of ancestral reconstruction in the simulation study. Independent Evolution also yielded much more accurate estimates of ancestral (i.e. fossil) primate brain and body sizes than the other methods of ancestral reconstruction.

Keywords: adaptation, adaptive peak, ancestral reconstruction, Brownian motion, comparative method, Ornstein-Uhlenbeck, phylogenetic analysis.

Correspondence: J.B. Smaers, Leverhulme Centre for Human Evolutionary Studies, Department of Biological Anthropology, University of Cambridge, The Henry Wellcome Building, Fitzwilliam Street, Cambridge CB2 1QH, UK. e-mail: jbs32@cam.ac.uk

Consult the copyright statement on the inside front cover for non-commercial copying policies.

INTRODUCTION

Methods of ancestral reconstruction (MARs) take into account phylogenies and estimate patterns of change along the branches of the phylogeny based on extant variation in order to investigate timing and mode of character acquisition and infer trait co-evolution (Pagel, 1999). Ancestral reconstruction of character states can help to fill in the gaps of the palaeontological record, expand the range of variables that can be included in the analysis, and increase the level of detail by which historical patterns of character evolution can be revealed (Harvey and Pagel, 1991). Methods of ancestral reconstruction thus provide a valuable contribution to traditional palaeontological approaches (e.g. Finarelli and Flynn, 2006; Soligo, 2006), especially in cases where the fossil record is patchy or incomplete and where biological traits that do not fossilize are under investigation (e.g. soft tissue).

To assess the influence of historical and biological processes on evolution, methods of ancestral reconstruction assume that: historical patterns of evolutionary change along the branches of a phylogeny are preserved in the variation among the tips of the phylogeny; the phylogeny used is correct; and the model of evolution used to ‘count back in time’ is a valid approximation of how the trait under investigation has actually evolved (and hence that the data must comply to the assumptions of the model) (cf. Pagel, 1999).

Because methods of ancestral reconstruction aim to estimate the (most often unobservable) past, a great deal of attention must be given to the question of how we test their accuracy. Two main approaches have been used: comparing what is known from the palaeontological record with the estimates of particular methods of ancestral reconstruction (e.g. Oakley and Cunningham, 2000; Polly, 2001; Webster and Purvis, 2002a, 2002b) and assessing the accuracy of MAR estimates of simulated patterns of evolution (e.g. Martins and Garland, 1991; Gittleman and Luh, 1992; Diaz-Uriarte and Garland, 1996, 1998; Martins, 1999; Martins *et al.*, 2002; Housworth *et al.*, 2004; Hansen *et al.*, 2008). These studies, however, have shown that accurately estimating ancestral states of quantitative characters is a real challenge, leading some authors to suggest that no currently available method yields acceptable results (e.g. Oakley and Cunningham, 2000; Webster and Purvis, 2002a, 2002b). The root of the problems associated with ancestral estimation is commonly attributed to unrealistic assumptions regarding the model of evolution (Martins and Garland, 1991; Westoby *et al.*, 1995; Diaz-Uriarte and Garland, 1996; Price, 1997). Methods of ancestral reconstruction are generally built around either a Brownian Motion (Felsenstein, 1985; Maddison, 1991; Schluter *et al.*, 1997) or an Ornstein-Uhlenbeck model of evolution (Hansen, 1997; Butler and King, 2004; Hansen *et al.*, 2008).

Brownian Motion (BM) is a model originally used to describe the evolution of gene frequencies (Edwards and Cavalli-Sforza, 1964), incorporating a purely gradual evolutionary scenario assuming that rates of change are constant throughout time and along all branches and that the probability of trait change is independent from both prior and current character states and from changes elsewhere in the tree (Schluter *et al.*, 1997; Webster and Purvis, 2002b). Felsenstein (1985), Maddison (1991), and Schluter *et al.* (1997) developed methods of ancestral reconstruction that are directly, or indirectly, based on the BM model of evolution. The main advantage of applying BM to estimating the evolution of quantitative characters is that it is mathematically tractable, since change is assumed to be proportional to branch length. The disadvantage of a purely gradual model of evolution is that it may be inappropriate to study adaptive evolution (Westoby *et al.*, 1995; Price, 1997), in particular when a trait is thought to be evolving under selection (Felsenstein, 1988; Harvey and Pagel, 1991; Harvey and Rambaut, 2000) or when two species are under different selective regimes (Butler and King, 2004). The underlying reason for this is that BM does not allow a trait to be ‘pulled’ towards a more

optimal value with any force other than one assumed to be normally distributed with mean zero and a variance proportional to the time over which the change occurs. In other words, the character state of the descendant is not expected to match a more adaptive optimal value if the character state of the ancestor does not match it. Brownian Motion-based methods are thus expected to ‘carry over’ error from ancestor to descendant when a burst of selection occurs on one particular branch [referred to as ‘the problem of inherited maladaptation’ by Hansen and Orzack (2005)].

Felsenstein (1988) originally proposed the use of the Ornstein-Uhlenbeck (OU) process (Uhlenbeck and Ornstein, 1930) in the context of ancestral estimation. This process models change towards an adaptive optimum and provides a good approximation of a population that is wandering back and forth on a selective peak under the influence of genetic drift. The change around the adaptive optimum varies in a way typical of BM. Methods built around the OU process thereby overcome the main disadvantage of BM-based methods (i.e. not taking into account selection) by modelling adaptive change towards selective peaks. Hansen and colleagues (Hansen, 1997; Hansen *et al.*, 2008) and Butler and King (2004) further developed OU-based approaches by developing methods to incorporate multiple selective regimes with different selective peaks and different selective ‘powers’ in order to assess correlated evolution of biological traits. The OU-based methods thus take into account selection and formalize selection as a combination of inertia and adaptation given a specific optimum (Hansen, 1997; Butler and King, 2004; Hansen and Orzack, 2005; Hansen *et al.*, 2008). The disadvantage of the OU model of evolution is that it gradually ‘forgets’ past history because the steady pull towards an adaptive optimum gradually erases older historic information. However, it is generally accepted that morphology often allows us access to information on ancient evolutionary events, thus limiting the explanatory value of a model of evolution that erases all record of ancient phenotypes (Felsenstein, 1988, p. 465).

In general, BM-based methods do not take into account selection and assume that traits change at a rate proportional to the time over which the change occurs, while OU-based methods take into account selection by predefining a number of selective optima (‘niches’) that can be mapped onto a phylogeny (Butler and King, 2004) or by allowing the selective optima to vary at random over time (Hansen *et al.*, 2008).

In this paper, we propose a novel method of ancestral reconstruction (the method of Independent Evolution, IE) that models character evolution using an ‘adaptive peak (AP) model’ [first suggested in the context of ancestral reconstruction by Felsenstein (1988)]. Such a model assumes that, at each moment in time, a population is attempting to climb an adaptive peak whose location wanders with time (Felsenstein, 1988, p. 453). As the adaptive peak wanders through phenotypic space, it pulls the population along with it. The adaptive peak may sometimes favour higher, sometimes lower phenotype (incorporating variation in selection into the model) and may differ for each internal node. Using this model, the main factor affecting the population’s movements is the covariance matrix of the peak’s movements, which can only be obtained by directly observing the process of phenotypic change in palaeontological time, or inferring it from the differences between species (Felsenstein, 1988, p. 454). Theoretically, the AP model of evolution includes more specific models of evolution such as BM and OU by allowing adaptive peaks to differ for each internal node. When adaptive peaks favour static higher or lower phenotypes in a particular sequence of descendant nodes, they have a similar function as ‘selective optima’ in the OU model of evolution, ‘pulling’ the population towards a static selective peak within the confines of a particular ‘selective regime’. When, however, the adaptive peaks favour static values across

all nodes, the adaptive peak model closely resembles BM. Taking into account that under an adaptive peak model of evolution selection will cause the population to move in each generation towards the peak of the adaptive surface, the IE model recalculates the value of the adaptive peak for each internal node and formalizes the ‘pull towards the adaptive peak’ as the discrepancy between the value of the adaptive peak for internal node x and the values of the descendant nodes of x . The main advantage of the IE algorithm based on an adaptive peak model of evolution is that it allows one to incorporate purely gradual evolutionary scenarios (such as BM) and deviations from them (such as OU-inferred scenarios) as special cases.

Because the underlying process of evolution of many traits is mostly unknown, testing the validity of a single hypothesis (i.e. a specific model of evolution proposed to approximate the underlying process of evolution) may not be the most beneficial approach to evaluating the analytical power of different methods of ancestral reconstruction. Here we use the method of strong inference (Platt, 1964) by comparing a set of alternative hypotheses (different methods of ancestral reconstruction), each of which proposes a particular model of evolution (BM, OU or AP) to best describe the underlying evolutionary process. We use a simulation study and a case study of primate brain and body size to evaluate which of the alternative hypotheses (i.e. which of the different methods of ancestral reconstruction) yields the most accurate results. In the simulation study, we simulate directional selection on particular branches in the phylogeny and investigate the ability of each method of ancestral reconstruction to accurately detect phenotypic changes in the nodes surrounding the ‘selection branch’. Ancestral states of primate brain size and body size are then reconstructed for different methods of ancestral reconstruction and compared with fossil values. The BM-based methods used in the current analysis include: Independent Contrasts (Felsenstein, 1985), weighted for branch lengths (IC) or not so weighted (UIC, for unweighted independent contrasts); squared change parsimony [SCP (Swofford and Maddison, 1987; Maddison, 1991)]; and the one-parameter maximum likelihood method [ML-BM (Schluter *et al.*, 1997)]. The OU-based method used in the current analysis is a two-regime model using four different ‘selective powers’ [$\alpha = 0.5, 1, 1.5$ or 2 ; incorporated using COMPARE (Martins, 2004)]. In both the simulation study and the case study, the performance of these methods is compared with that of IE.

THE METHOD OF INDEPENDENT EVOLUTION

The IE distance metric

When considering the divergence of a trait between two species, gradual models of evolution assume that it evolved by a similar absolute amount per unit time along the two branches from a last common ancestor (cf. Felsenstein, 1985). However, since a 50-g change in body size represents much more relative trait divergence or trait evolution in a mouse than in an elephant, absolute size is generally seen as a potentially confounding variable (e.g. McMahon and Bonner, 1983). We therefore prefer to use a metric that directly measures proportional change. By dividing the amount of trait change (or Euclidean distance between two species) by the average magnitude of the trait, we obtain an estimate of ‘relative change’. Consider the example of an ancestor with a trait value of 100 and a descendant with a trait value of 50 (Fig. 1). The estimate of relative change is therefore -0.66 ($50 - 100/75$). Suppose the same ancestor also produces a descendant with a trait

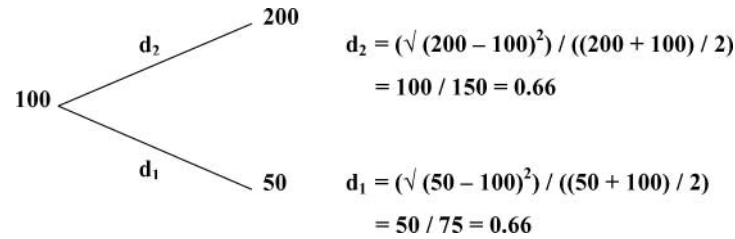


Fig. 1. The IE distance metric in sister species. d represents the distance between nodes.

value of 200. The relative change here will be 0.66 $(200 - 100/150)$. The relative change along a branch is thus calculated as the change between ancestor and descendant divided by their average. Note that the absolute value of the estimate of relative change is the same when the ancestor either decreases or increases in size with an equal factor, while the absolute change is quite different. This distance metric is therefore a useful way to quantify relative or proportional change, as it provides equal values for equal relative change, independently of the absolute value of the trait.

Estimating IE ancestral states

Independent Evolution estimates trait evolution and ancestral states directly from the trait distribution across species. Figure 2 considers a given continuous trait x and the estimation of the common ancestor ($A_{1,2}$) of two extant species 1 and 2. The IE algorithm consists of eight steps:

1. We select a group of species including species 1 and 2 with resolved phylogeny, known branch lengths at every node, and data on trait x for all terminal taxa (Fig. 2.1).
2. An 'adaptive peak' (AP) is calculated (as described below) at internal node $A_{1,2}$ from data for all terminal taxa (Fig. 2.2).
3. The adaptive peak replaces all branches of the tree ancestral to $A_{1,2}$. The tree is then unrooted creating a star tree including AP, x_1 , and x_2 (Fig. 2.3).
4. The star tree is considered as a triangle whereby the barycentre of the triangle represents $A_{1,2}$, and values x_1 , x_2 , and AP are represented by the tips of the triangle (Fig. 2.4).
5. We use the IE distance metric between x_1 , x_2 , and AP to calculate the triangle sides (S_1 , S_2 , and S_3 ; Fig. 2.5).
6. Using the Ptolemean property of triangle inequality, we calculate the distance from $A_{1,2}$ to the three tips x_1 , x_2 , and AP (T_1 , T_2 , and T_3) based only on the triangle sides (Fig. 2.6) [for a more in-depth discussion on using Ptolemy's theorem on triangle inequality in estimating trees from distance matrices, see Farris (1970, 1972)].
7. The resulting distances from $A_{1,2}$ to tips AP, x_1 , and x_2 represent the relative phenetic distances between ancestor ($A_{1,2}$) and descendants (x_1 and x_2), taking into account the AP (Fig. 2.7).
8. The distances between $A_{1,2}$ and x_1 and x_2 (T_1 and T_2) are weighted for their phylogenetic branch lengths, creating a rate of change (' R -value') for the branch of each descendant of ancestor $A_{1,2}$. R -values, representing the relative branch-specific evolutionary change of a trait, are then used to estimate $A_{1,2}$ according to Felsenstein's (1985) algorithm (Fig. 2.8). Note that with equal phylogenetic branch lengths in sister branches, the weight for both phenetic branch lengths equals unity.

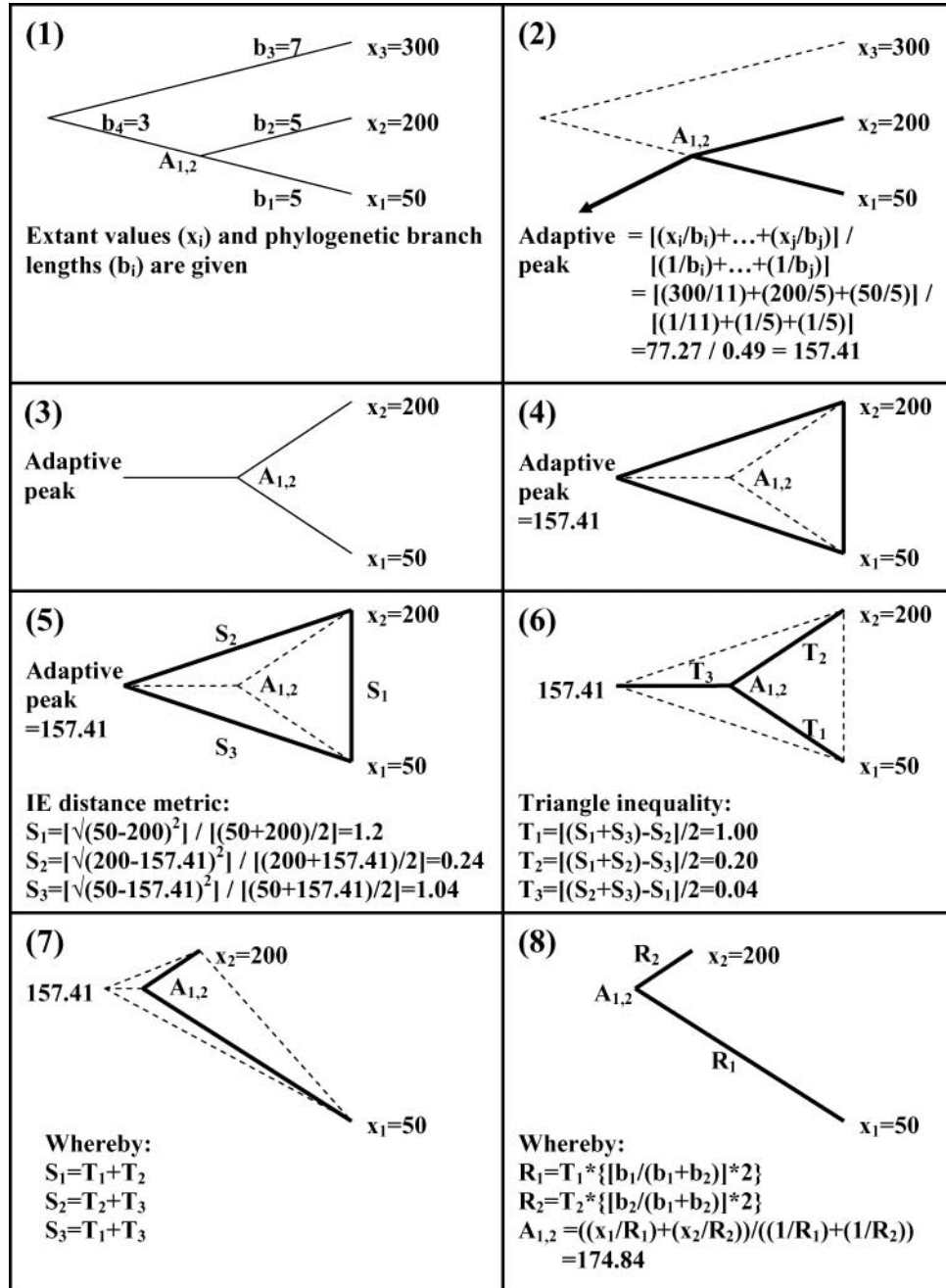


Fig. 2. The IE procedure for estimating ancestral states. x represents a continuously defined biological trait; x_i represents the extant value of trait x for species i ; b_i represents the phylogenetic branch length for species i ; $A_{i,j}$ represents the ancestral value for species x_i and x_j ; S represents the relative phenetic distances between the tips of the triangle (thus representing the sides of the triangle); T represents the relative phenetic distance from A to the tips of the triangle; R represents relative branch-specific evolutionary change of trait x .

To calculate ancestral values further down the tree, the eight-step algorithm is repeated recursively starting at the terminal taxa and proceeding towards some internal node of interest. At each step in the recursion, the estimated ancestral value at each internal node is stored and used as a descendant value to calculate ancestral values further down the tree.

Following Felsenstein's (1988) suggestion to assume that the adaptive peak changes through time wandering according to a BM, the IE adaptive peak for a particular internal node is calculated based on the BM-based Independent Contrasts algorithm to estimate ancestral states: the sum of the sister nodes divided by their branch lengths, divided by the sum of their inverse branch lengths. However, instead of only taking into account direct descendants when assigning nodal states, the IE adaptive peak extends this approach to all extant values. In other words, the IE adaptive peak value is calculated as the sum of all terminal nodes divided by the patristic difference between the terminal node and the node in question, divided by the sum of the inverse patristic differences of all terminal nodes (Fig. 2.2). The patristic difference is defined as the sum of the lengths of the branches lying on the path connecting the two nodes (Farris, 1967). The IE algorithm to estimate the adaptive peak thus takes into account all the available (i.e. extant) biological information provided by the taxa in the phylogeny in order to increase the reliability of character optimization (Ryan, 1996; Cunningham *et al.*, 1998; Lee and Shine, 1998), with the additional advantage of giving more weight to a species the closer it is to the ancestor being estimated. The assumption is that change among closely related species is likely to be directed towards an optimum that is most apparent in the trait's distribution among its closest relatives. As the adaptive peak is considered to represent the trait's distribution at the node in question, it replaces all branches of the tree apart from the descendant branches of $A_{1,2}$ (Figs. 2.2 and 2.3).

Properties of IE's adaptive peak model of evolution

By estimating an adaptive peak for each internal node of the tree and calculating the distance between the value of the adaptive peak for that particular internal node and the value of its descendant nodes to reflect different degrees of selective power, IE models a population moving in each generation towards the peak of the adaptive surface. The adaptive peak is inferred from the differences between extant species and is thus only constrained by extant variation. By allowing the adaptive peaks to differ for each internal node, variation in selection is incorporated into the model sometimes favouring higher, sometimes lower phenotypes.

The combination of these properties allows the inclusion of specific models of evolution such as BM and OU as special cases of IE's underlying adaptive peak model. Formally, IE's adaptive peak model collapses into a BM model of evolution when S_2 equals S_3 in Fig. 2.5, because the IE algorithm for calculating $A_{1,2}$ is then reduced to the estimation of ancestral states via Independent Contrasts under an explicit BM model of evolution. When $S_2 \neq S_3$, directional trends can be recognized indicating the wanderings of the adaptive peak towards a new adaptive surface. The larger the inequality between S_2 and S_3 , the larger the directional trend that is inferred. Within IE, adaptive peak values have a similar function as the 'selective optima' used in OU-based methods, pulling the population towards a particular phenotype. Formally, the possibility of a 'pull' towards a selective optimum as modelled by an OU model of evolution is incorporated into IE's algorithm when a great disparity exists between S_2 and S_3 in a particular branch directly followed by a sequence of branches in which $S_2 \approx S_3$. Such variation in selection is taken into account by recalculating

S_2 and S_3 for each internal node depending on the adaptive peak for that node. In this way, the properties of IE's algorithm allow the possibility to incorporate fundamental aspects of the OU model of evolution. The difference is that IE infers the value of the selective optimum and where the selective regime takes place based on extant variation, while OU-based methods either *a priori* define the topological location of the selective regime and the value of the selective optimum (Butler and King, 2004) or allow the selective optima to vary at random over time (Hansen *et al.*, 2008). In the case where a valid argument can be made towards predefining selective optima in particular selective regimes, IE can accommodate this by adjusting the value of the adaptive peaks accordingly. Selective regimes are then incorporated into IE in much the same way as Butler and King (2004) suggest incorporating selective optima when using an OU model of evolution. Independent Evolution thus provides the algorithms needed to formally include specific models of evolution such as BM and OU into a theoretically more inclusive adaptive peak model of evolution when reconstructing ancestral states.

SIMULATIONS

Simulation studies are commonly used to investigate the statistical properties and analytical power of comparative methods (Martins and Garland, 1991; Gittleman and Luh, 1992, 1994; Diaz-Uriarte and Garland, 1996; Garland and Diaz-Uriarte, 1999; Martins *et al.*, 2002; Housworth *et al.*, 2004; Hansen *et al.*, 2008). Simulation procedures usually assign a specific value at the root of the phylogeny and add random change to the previous value at each step in the phylogeny until trait values are obtained for all terminal nodes (cf. Martins and Garland, 1991; Garland *et al.*, 1999; Martins, 1999). Different models of evolution are commonly simulated by using BM as a template of change and adjusting the rate of change or the length of particular branches (e.g. Martins and Garland, 1991; Martins, 1993; Diaz-Uriarte and Garland, 1996; Butler and King, 2004).

The current study employs a strong inference approach to evaluate the effect on performance of ancestral reconstruction between different alternative hypotheses (i.e. different methods of ancestral reconstruction built on different models of evolution) when simulating directional selection on particular ancestral branches in the phylogeny. Our simulation uses the topology of a model primate phylogeny and simulates phenotype values for each node in the tree. A primate phylogeny is used as a model phylogeny because primates are the most well-studied clade in this context, providing a realistic topology with resolved branch lengths at every node and a large sample of terminal taxa. To simulate phenotype values for every node in the phylogeny, BM is used as a template by randomly assigning each branch with an index of percentage of change (C_i) varying between -0.1% and 0.1% . Internal node values are then reconstructed from the root upwards according to the following algorithm: descendant = ancestor $[(C_i/100) + 1]^{\text{branch length}}$. Directional selection is simulated by adjusting the rate of change (i.e. C_i value) on particular branches to 10% , 20% , -10% or -20% (representing positive and negative directional selection on the one hand and strong versus moderate selective power on the other). Each 'simulation-run' thus reconstructs values for all internal and terminal node values of the phylogeny. Reconstructed values can be considered to represent a hypothetical biological trait (e.g. body size). Simulated terminal node values are subsequently used as input for each method of ancestral reconstruction, providing MAR estimates for all internal nodes, which are compared to the simulated values of the internal nodes to assess accuracy. Figure 3 indicates the topology of the primate phylogeny used (Smith and Cheverud, 2002),

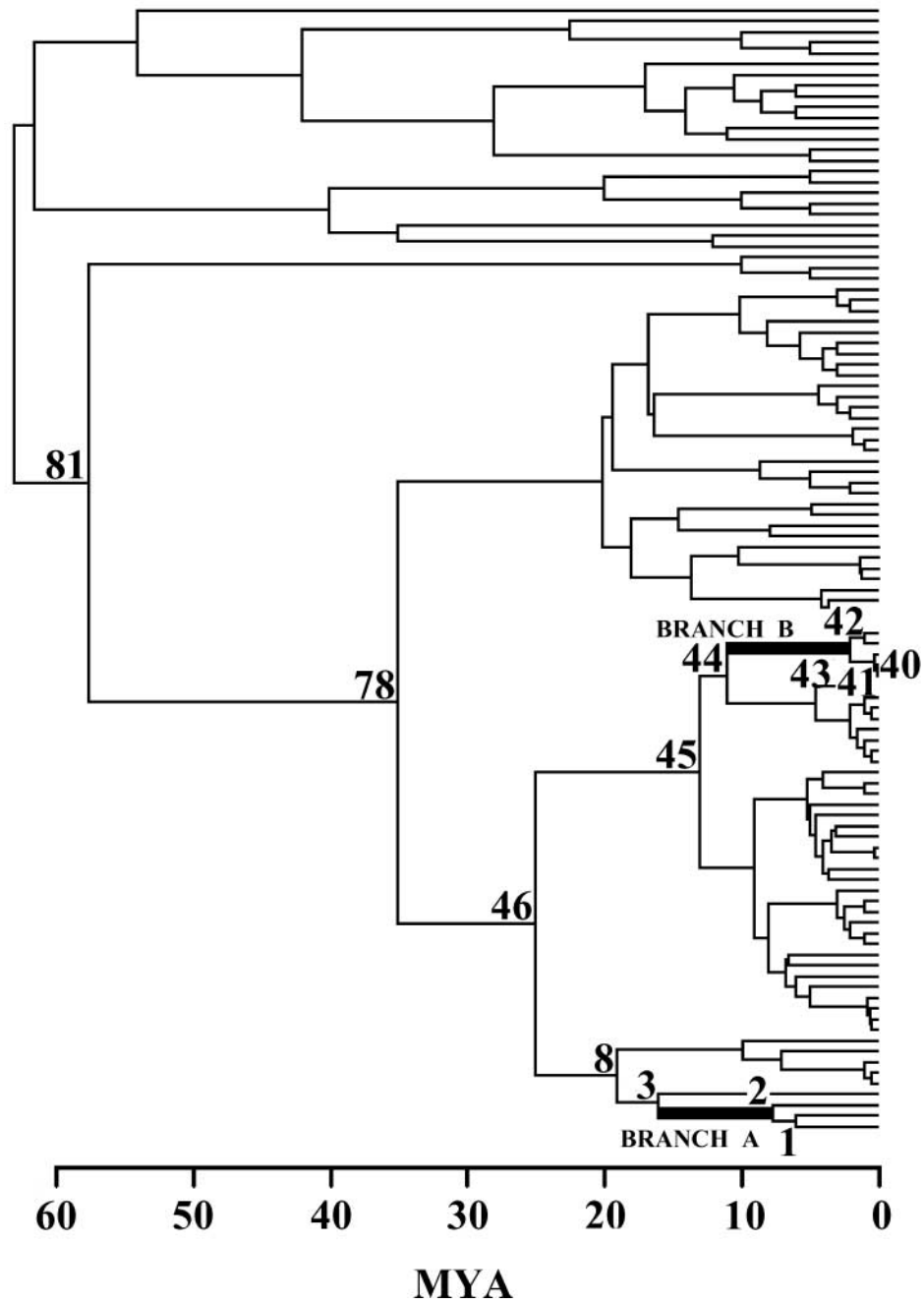


Fig. 3. Topology of the primate phylogeny used in the simulation study. Directional selection is simulated on 'Branch A' and 'Branch B'. Numbers indicate the ancestral nodes for which accuracy is assessed using different methods of ancestral reconstruction.

where directional selection is simulated and the ancestral nodes are assessed in terms of accuracy.

Accuracy is assessed by considering MAR estimated ancestral values relative to simulated ancestral values for particular internal nodes (MAR estimate/simulated value). This results in a value indicating a percentage of residual error (a value of 1.20 indicates that the MAR estimate overestimates the simulated value by 20%). When MAR estimates underestimate simulated values ($0 < \text{residual error} < 1$), residual errors are inversed and given a negative sign so that the order of magnitude is comparable to when simulated values are overestimated. For each simulation, 10 simulation runs are considered. Such a small number of simulation runs is justified because standard deviations of simulated values over 10 simulation runs do not surpass 0.02 while averaging around 100 within each simulation.

Results indicate that the baseline BM simulation (in which no directional selection is assigned) yields accurate results for all methods of ancestral reconstruction (average of residual errors within MARs = 1.00, standard deviations of residual errors within MARs < 0.001, average of ancestral values ~100). This confirms the ability of IE and OU to collapse into BM when its premises are not violated.

Figures 4 and 5 show results when directional selection is simulated on the ancestral branches indicated in Fig. 3. The choice as to which branch is used to simulate directional selection on is trivial because similar trends can be recognized for all branches (similarities between Branches A and B can be recognized when comparing Figs. 4 and 5). Four variants of directional change are considered: positive and negative selection with a moderate or a strong selective regime. Figure 4 compares the accuracy of the indicated ancestral nodes between IE and BM-based methods, while Fig. 5 compares results between IE and OU-based methods.

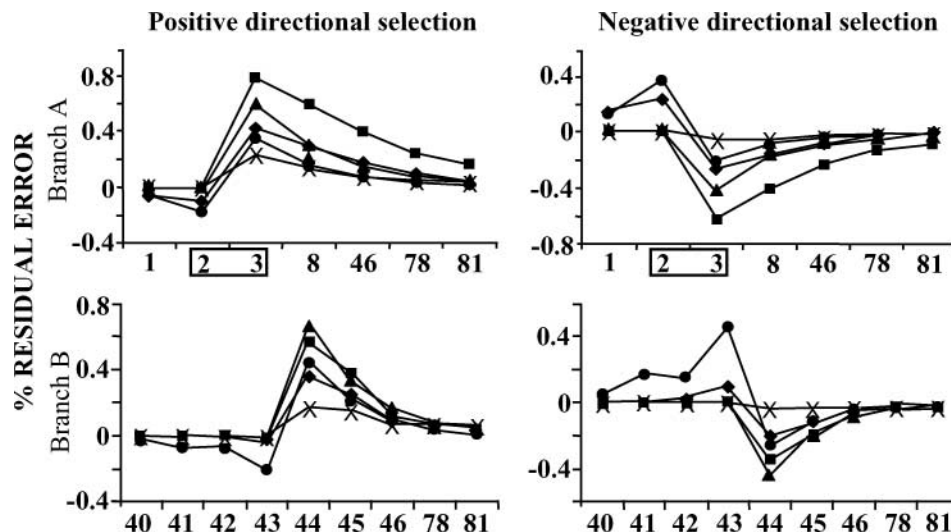


Fig. 4. Results for IE and BM-based methods of ancestral reconstruction when considering positive and negative directional selection under a selective regime with moderate power. For Branch A, directional selection is simulated on the branch connecting ancestral nodes 2 and 3; for Branch B, directional selection is simulated on the branch connecting ancestral nodes 43 and 44. Crosses indicate the results for IE, squares for IC, triangles for UIC, circles for SCP, and diamonds for ML-BM.

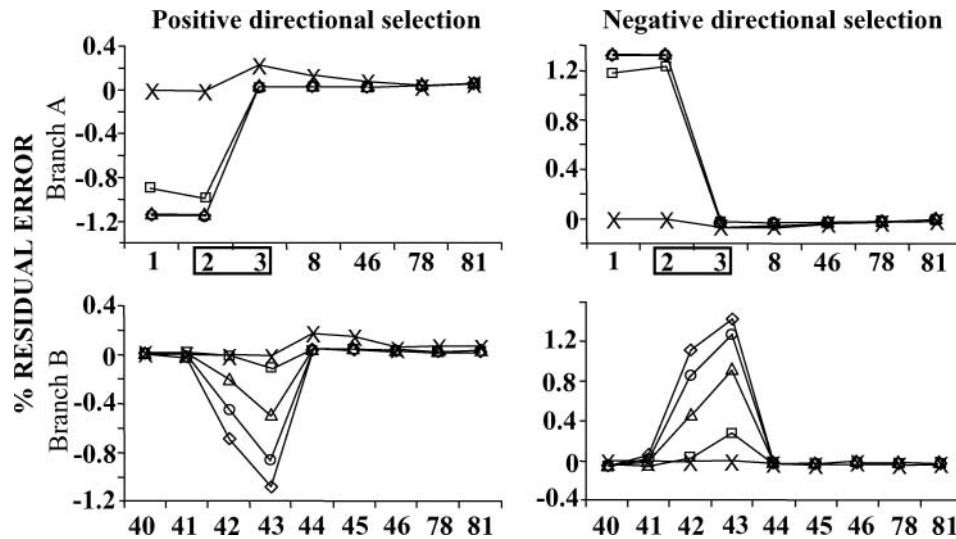


Fig. 5. Results for IE and OU-based methods of ancestral reconstruction when considering positive and negative directional selection under a selective regime with moderate power. For Branch A, directional selection is simulated on the branch connecting ancestral nodes 2 and 3; for Branch B, directional selection is simulated on the branch connecting ancestral nodes 43 and 44. Crosses indicate results for IE, squares for $OU_{0.5}$, triangles for OU_1 , circles for $OU_{1.5}$, and diamonds for OU_2 .

When considering the results for positive directional selection in Fig. 4, IE and BM-based methods pile up error in nodes ancestral to the branch on which directional selection is simulated, indicating most error in the node directly ancestral to the directional branch, and error gradually disappearing in higher nodes. More specifically, an overestimation of ancestral states is indicated. Independent Evolution, however, indicates more accurate results than any BM-based method for every ancestral branch considered. When considering negative directional selection in Fig. 4, SCP and ML-BM indicate an overestimation of the ancestral states in the nodes descendant from the directional branch and an underestimation of the nodes ancestral to it, while only the latter trend occurs for IC and UIC, but in a more pronounced way. Also here, IE outperforms all BM-based methods. This result accurately reflects what has been referred to as the problem of ‘inherited maladaptation’ when using BM to model selection (Hansen and Orzack, 2005).

Comparing the performance of IE to OU-based methods, Fig. 5 indicates that only nodes descendant from the directional branch indicate significant error in OU-based methods. More specifically, nodes descendant from the directional branch are significantly underestimated in the case of positive directional selection and significantly overestimated when negative directional selection is simulated. For Branch A this is true for all descendant nodes of the directional branch, while for Branch B this is only true for some descendant nodes (42 and 43), as others are accurately reconstructed (40 and 41).

When considering ancestral nodes that are not descendant or ancestral to the directional branch, IE and all BM-based methods indicate no error, thus accurately reflecting the simulated BM template. The OU-based methods, however, consistently over- or underestimate all ancestral nodes depending on whether the selective regime is positive or

negative. This trend differs between moderate and strong selective powers only in order of magnitude and is more pronounced when considering positive directional selection (2–3% for positive moderate, 6–9% for positive strong, 1–2% for negative moderate, and 2–3% for negative strong directional selection). Furthermore, Figs. 4 and 5 only consider directional selection on ancestral branches. When directional selection is simulated on terminal branches, the same trends occur (data not shown). Note that in this case OU-based methods do not indicate a burst of error in any particular ancestral node, because the error in these methods is especially concentrated on nodes descendant from directional branches. As the power of the selective regime increases, OU-based methods, however, still consistently over- or underestimate all ancestral nodes depending on whether the selective regime is positive or negative. Finally, comparing results between moderate and strong selective regimes reveals that residual error for all methods of ancestral reconstruction only differs in order of magnitude (Fig. 6).

Results thus indicate that IE significantly outperforms both BM-based and OU-based methods of ancestral reconstruction when simulating different degrees of directional selection on specific branches in the currently considered phylogeny. When considering either moderate or strong selective power, results differ only in order of magnitude (*mutatis mutandis* when considering directional selection on small vs. long branches). These results confirm what is expected from the theoretical development, which suggests that IE accommodates variation in selection more accurately than methods that are built on either a BM or an OU model of evolution.

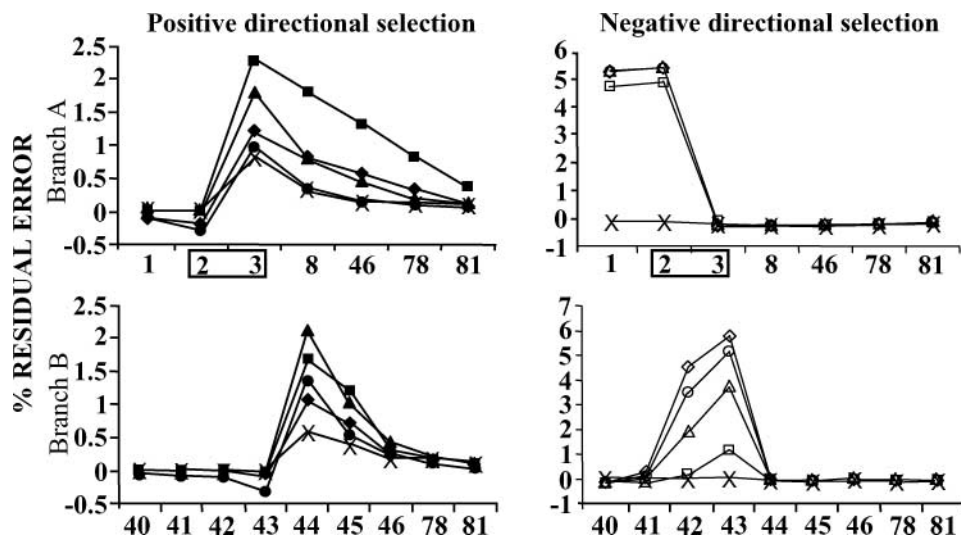


Fig. 6. Results for IE and BM-based methods when positive directional selection is simulated and for IE and OU-based methods when negative selection is simulated, both under a selective regime with strong power. Comparing the former with results from Fig. 4 and the latter with results from Fig. 5 clearly indicates that the difference in results when simulating directional selection under a selective regime with moderate power differs from a selective regime with strong power only in order of magnitude.

ESTIMATING ANCESTRAL PRIMATE BRAIN AND BODY SIZE

Applying the strong inference approach, we also use primate brain and body size as a worked example to evaluate the accuracy of six methods of ancestral reconstruction by comparing their estimates with fossil data. We focus more on body size, as fossil brain data are less abundant.

Phylogeny and data set

The phylogeny of Smith and Cheverud (2002) is used, adjusted for 105 extant species (23 Strepsirrhini, 32 Platyrrhini, 47 Catarrhini; see Fig. 4). We use this phylogeny because it is fully resolved for all nodes and provides divergence date estimates for each node. The body mass data are from Smith and Cheverud (2002); brain size data are from Bauchot and Stephan (1969), H. Stephan (unpublished data), and Kaplan *et al.* (2003); and fossil data are from Aiello and Dunbar (1993), Fleagle (1999), and Moyà-Solà *et al.* (2004) (see Appendix 1).

Mapping fossils onto the phylogeny

To overcome some of the uncertainties related to mapping fossils onto particular topological locations, a consistent three-step procedure is used based on the age of the fossil and its proposed phyletic relationships (see Appendix 2): (1) a specific age or epoch is assigned to every fossil based on the dating of the site in which it was recovered; (2) based on the proposed phyletic relationships of a particular fossil, a specific phylogenetic branch or node is assigned onto which the fossil should be mapped; (3) if the specific age attributed to every fossil (in step 1) overlaps with its assigned branch or node (in step 2), the corresponding topological location is used for that fossil (applies to 22 fossils used). If, however, there is no overlap between information from steps 1 and 2 (applies to three fossils used: *Dolichocebus gaimanensis*, *Homunculus*, and *Aegyptopithecus*), the topological location of the fossil is attributed to that point on the assigned branch (from step 2) that is closest to its attributed specific age (from step 1). For the latter three fossils, the difference between its assigned branch and its attributed specific age is no more than 2 Mya (see Appendix 2).

Webster and Purvis (2002b) provided the first comparison of a wide range of methods of ancestral estimation against fossil data using estimates of eight fossil Platyrrhine species. We extend this approach by assessing 28 primate fossil data points that can be considered to be in a direct line of ancestry to at least some of the extant species of the phylogeny (see Fig. 7 and Appendix 1), including six of eight fossil species that Webster and Purvis assessed [data on *Mohanamico herskovitzi* are not included in the current analysis because its extant descendant *Callimico goeldii* is not included in Smith and Cheverud's (2002) phylogeny; data on *Protopithecus brasiliensis* are not included because it constitutes 'the first fossil primate that was recognized to be unlike any living species' (Fleagle, 1999, p. 441)].

Assessing accuracy

We use four measures to assess accuracy of ancestral estimates. Three of these measures are absolute, in the sense that they provide a test on whether a method is accurate or not (e.g. on level $P < 0.05$). The fourth method is a relative measure of accuracy in the sense that it indicates the relative accuracy of one method over another (e.g. method 1 is three times more accurate than method 2).

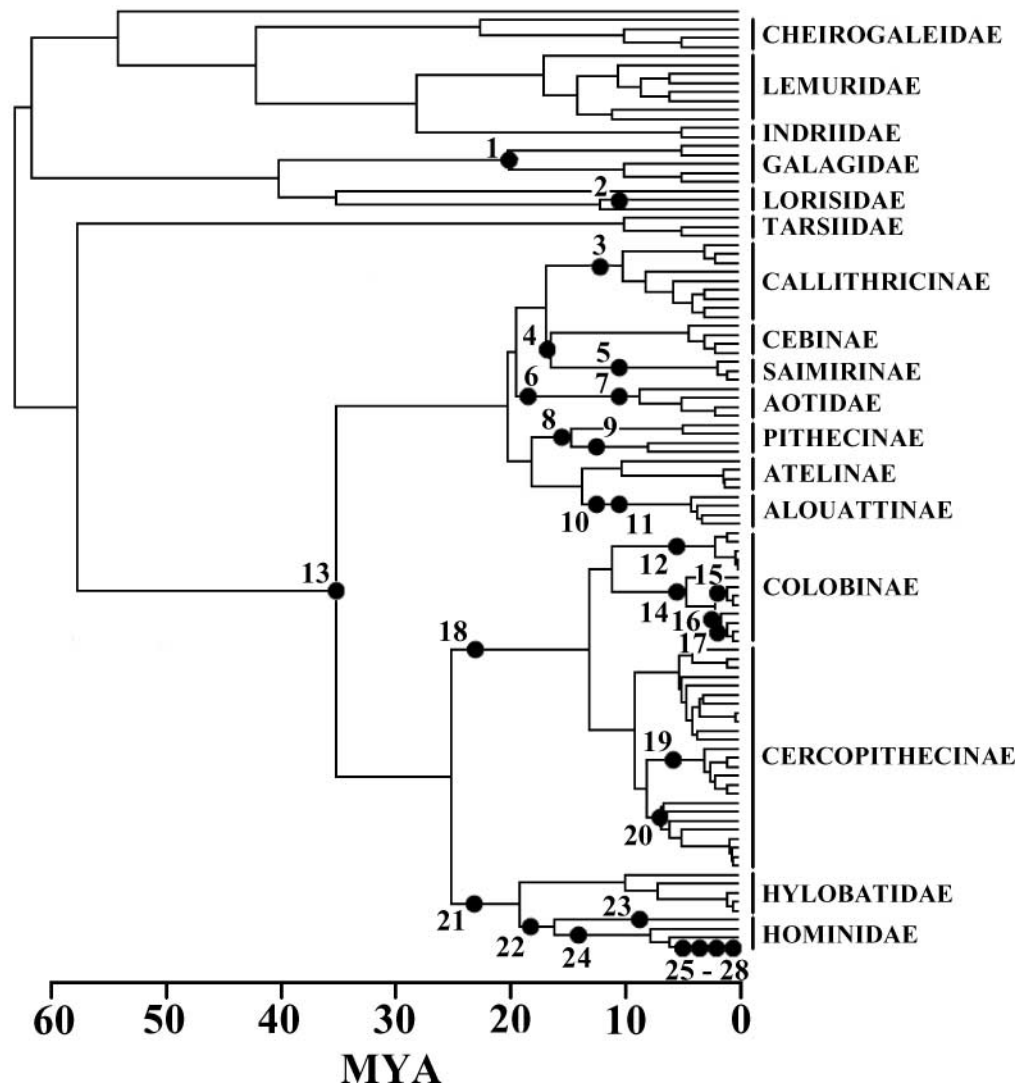


Fig. 7. The primate phylogeny used in the estimation of fossil data points by different methods of ancestral reconstruction. To estimate ancestral primate brain and body size, fossil data points are mapped onto the phylogeny and MAR estimates of the fossil data points are calculated based on extant values.

Absolute measures

First, overall r^2 is the proportion of variance among ancestral values that is explained by the estimates. However, r^2 should not be assessed alone, since the correlation r (standardized coefficient) between observation and estimate may be high even if the slope (non-standardized coefficient) relating the two is not unity and/or the intercept is not zero (a regression with slope 1 and intercept 0 is expected if predicted values are accurate, i.e. similar to observed values). Since both variables in the regression are subject to error,

reduced major axis regression is used (Sokal and Rohlf, 1995) and implemented using RMA for JAVA (Bohonak and van der Linde, 2004). As a second measure of accuracy, we divide the estimate by the fossil value. Since deviation from unity represents error, we used a one-sample t -test (with a set value of 1) to determine whether the average of estimations deviates from unity. A weakness of this method is that the average can equal unity while the standard deviation around the mean is large. Using this measure of accuracy, a method with good accuracy would therefore not only have a mean value of unity, but also a low standard deviation around the mean. A third absolute measure of accuracy consists of a paired-samples t -test between the residual errors (calculated as the absolute difference between the actual fossil value and the estimate) of different methods.

Relative measure

A fourth measure (E -value) calculates the absolute difference between the estimate of each method and the fossil data point (residual error) for each node and divides results for IC, UIC, ML-BM, and SCP by that of IE to determine to what extent IE outperforms other methods. The E -value for each method is averaged across nodes to produce the overall E -value for a particular method.

Results

A reduced major axis regression between actual fossil body size values and estimates of them (for a full list of the estimates of all fossil data points, see Appendix 3) indicates accurate results for IE, SCP, and ML-BM in terms of intercept and slope (see Table 1). Both IC and UIC indicate intercepts significantly lower than zero and slopes significantly higher than unity. A one-sample t -test reveals that IE, IC, and UIC are the only methods in which

Table 1. Results of four accuracy measures for the estimates of fossil data points by different methods

	Reduced major axis			One-sample t -test		Paired-sample t -test		E -value	
	r^2_{rma}	Intercept _{rma}	Slope _{rma}	E	σ	Average	σ	E_{all}	$E_{\text{excl 4}}$
IE	0.956	0.151	0.964	1.074	0.346	3523	5315	1	1
ML	0.929	0.343	0.936	1.326**	0.461	6048 [#]	7880	4.07	4.14
SCP	0.940	0.211	0.964	1.243**	0.406	5660 [#]	8502	3.35	3.51
IC	0.932	-0.452	1.129*	1.230*	0.648	6653 [#]	9119	4.97	2.10
UIC	0.944	-0.567*	1.151*	1.135	0.513	6163	10157	4.44	1.91
OU _{0.5}	0.487	2.163***	0.470***	2.383*	3.007	6541 [#]	7989	17.62	6.22
OU ₁	0.428	2.203***	0.458***	2.424*	3.093	6942 [#]	8151	18.17	6.28
OU _{1.5}	0.415	2.218***	0.454***	2.468*	3.170	6975 [#]	8137	18.26	6.23
OU ₂	0.409	2.226***	0.453***	2.494*	3.218	6975 [#]	8131	18.33	6.19

Note: All data points are described in Appendix 3. $N=28$ for all methods. Reduced major axis regression, one-sample t -test, and paired-sample t -test represent absolute measures of accuracy; E is a relative measure of accuracy. E_{all} indicates the E -values based on all fossil data points; $E_{\text{excl 4}}$ indicates E -values based on all fossil values excluding 4 fossils with exceptionally high E -values (see text for explanation). Values marked *, **, and *** are significantly different from 0 (for Intercept_{rma}) and 1 (for Slope_{rma} and E of the one-sample t -test) at $P < 0.10$, $P < 0.05$, and $P < 0.01$ respectively. Values marked with [#] and ^{##} are significantly different from the IE value at $P < 0.10$ and $P < 0.05$ respectively.

the average of the estimated value divided by the fossil value does not significantly differ from unity (see Table 1), although the standard deviation around the mean is lower with IE than any other method. Furthermore, a paired-samples *t*-test between the residual errors of different methods indicates that the average residual error per fossil data point is significantly smaller in IE compared with any other method apart from UIC (see Table 1).

The *E*-values confirm that IE outperforms other methods, indicating that IE is consistently around between 3 and 18 times more accurate than the other methods (see E_{all} in Table 1). More detailed analysis of specific data points reveals that *E*-values are exceptionally high for four fossil data points in IC, UIC, SCP, and ML-BM: *Progalago – Komba*, *Lagonimico conclutatus*, *Proconsulidae*, and *Dryopithecus* (*E*-values averaged across methods of ancestral reconstruction: 34.97, 17.77, 28.76, and 14.39 respectively). For the OU-based methods, four other fossil data points reveal exceptionally high *E*-values: *Progalago – Komba*, *Nycticebus simpsoni*, *Lagonimico conclutatus*, and *Aotus dindensis* (*E*-values averaged across methods of ancestral reconstruction: 1062.31, 674.73, 249.73, and 71.20 respectively). These exceptionally high *E*-values may artificially inflate the *E*-values. The *E*-values are therefore also calculated excluding these four fossil data points and results indicate that IE still provides estimates that are consistently 2–6 times more accurate than other methods (see $E_{\text{excl 4}}$ in Table 1).

Overall, these results indicate that for the current analysis, IE yields more accurate results than the other methods considered. Furthermore, Table 1 suggests that IC and UIC underestimate lower actual fossil values and overestimate higher actual fossil values (intercept < 0 and slope > 1), while ML-BM and SCP consistently overestimate all actual fossil values ($r^2 < 1$ and the average of the estimate divided by the actual fossil value > 1, while intercept = 0 and slope = 1). The OU-based methods do not yield accurate results for any of the accuracy measures considered.

The two-regime OU model of evolution used in the current analysis may be considered to be inappropriate for modelling primate body size evolution because more than two selective regimes are expected to occur. However, when considering *R*-values as produced by the only method that yields accurate estimates of fossils (IE), results indicate that selective regimes are gradually distributed throughout phenotypic space (Fig. 8) rather than distinctly characterized on particular topological locations. When considering the evolution of primate body size, this calls into question the validity of predetermining a distinct amount of selective regimes on specific topological locations, such as is done in OU-based methods that allow the incorporation of multiple selective regimes (e.g. Butler and King, 2004; Hansen *et al.*, 2008).

Furthermore, *R*-values, which can be interpreted as branch-specific rates of change, can be plotted to investigate evolutionary change within and between species or subfamilies of species within and between traits. Comparing primate brain size and body size using IE, the *R*-values reflect what is commonly known about brain-to-body allometry and species dwarfism and gigantism. When regressing *R*-values of brain size onto *R*-values of body size, a slope of 0.770 is obtained. This scaling parameter is in line with earlier studies (e.g. Eisenberg, 1981; Armstrong, 1982a, 1982b, 1983; Martin, 1982). In terms of gigantism, *Homo sapiens* yields the highest *R*-value for brain size, while *Gorilla* yields the highest for body size. The ancestral branch of great apes indicates the shift in brain size and body size growth in this subfamily. In terms of dwarfism, the highest values are indicated by the ancestral branch of tarsiers (Tarsiidae), the ancestral branch of dwarf and mouse lemurs (Cheirogaleidae), the ancestral

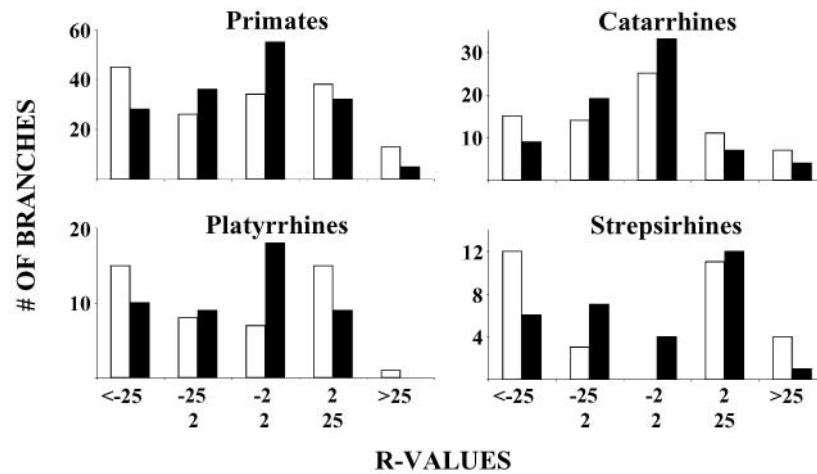


Fig. 8. *R*-value distribution for body size (white bars) and brain size (black bars). On the x-axis, upper and lower values represent the range of *R*-values presented.

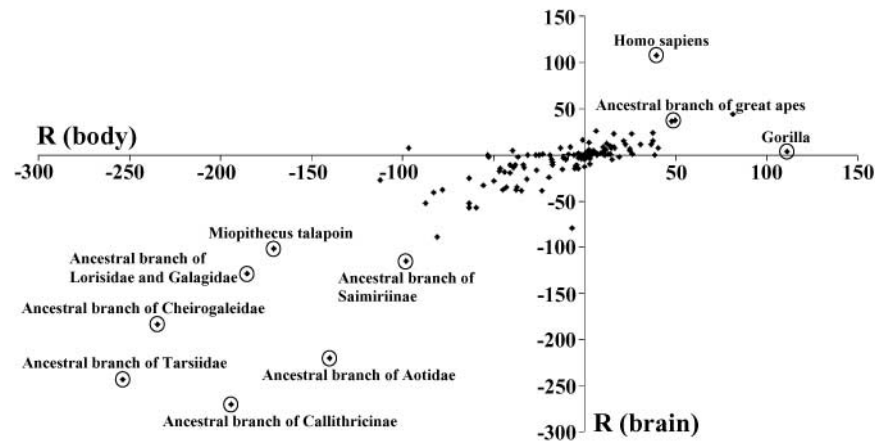


Fig. 9. *R*-values for brain size and body size in primates.

branch of marmosets and tamarins (Callithricinae), the ancestral branch of the lorises and bushbabies (Lorisidae and Galagidae), followed by *Miopithecus talapoin* (the smallest Old World monkey), the ancestral branch of owl monkeys (Aotidae), and the ancestral branch of squirrel monkeys (Saimiriinae) (Fig. 9).

These results thus suggest that in terms of primate body size and brain size evolution, IE is able to accurately reconstruct macro-evolutionary processes that occurred on ancestral branches of extant subfamilies and families of primates. A good example is the human–chimpanzee ancestral point, for which only IE yields realistic estimates of brain size and body size (Table 2).

Table 2. The human–chimp ancestral state for brain size and body size as estimated by different methods

	<i>H. sapiens</i>	<i>P. troglodytes</i>	<i>A. afarensis</i>	IE	ML	SCP	IC	UIC	OU _{0.5}	OU ₁	OU _{1.5}	OU ₂
Brain size	1334	378	433	444	704	744	856	856	171	95	92	92
Male body size	60200	38200	44600	43931	73310	68901	51450	51450	14188	8626	8478	8527
Female bodysize	53600	33700	29300	35965	47250	46707	43650	43650	9896	5673	5548	5581

Note: Brain size estimates are computed with data from Kaplan *et al.* (2003), body size estimates are computed with data from Smith and Cheverud (2002). Data on *Homo sapiens*, *Pan troglodytes*, and *Australopithecus afarensis* come from Aiello and Dunbar (1993), Bauchot and Stephan (1969), Fleagle (1999), and Smith and Cheverud (2002).

DISCUSSION

The main contribution of incorporating an adaptive peak model of evolution into comparative research of quantitative characters lies in its applicability to different data sets with different processes of evolution. Its main advantage is that it includes more specific models of evolution such as BM and OU as special cases. The advantage of the IE algorithm proposed to incorporate the assumptions of an AP model of evolution is that it allows investigation of evolution on a branch-specific level, which in turn allows exploration of directional trends and rates of changes within and between species and subfamilies within and between traits. This may not only increase the accuracy of ancestral estimation but may also be, in itself, of particular interest to many comparative researchers.

The theoretical development underlying IE is, however, yet to incorporate computing of standard errors of estimates, and taking account of uncertainty of character mapping and of the phylogeny. Moreover, IE is expected to be subject to the parametric limitations of data sampling. Because the proposed procedure for estimating adaptive peak values cannot be considered to be independent of the specific phylogeny that is used, it is expected that larger samples will yield progressively more accurate results. A metric to estimate when the sample is large enough for IE to be valid is currently unavailable (although preliminary results indicate adaptive peak values reach an asymptote when 15–20 species are taken into account). To explore this parametric limitation for the current analysis of primate body size, the original data set of 105 was downsized to 59 by averaging all branches that diverged in the last 3.5 million years. Results indicate that using the smaller sample still yields accurate results for every accuracy measure discussed, and IE still significantly outperforms all other methods.

Furthermore, because the value of the adaptive peaks is directly deduced from extant variation, trends that are not apparent in the extant distribution cannot be estimated without prior knowledge. Moreover, IE will estimate the character state at the root of the tree within the range of the observed values in the data. Directional GLS models (Pagel, 1997, 1999) can overcome this limitation. If species that have diverged more from the root also tend to have changed more in a given direction, then directional GLS models can use the trend to reconstruct the character state at the root of the tree to lie outside of the range of observed values. A precondition of this method is that tips differ in their distance from the root. Many samples, however, consist of observed extant values only, in which case all values are equidistant to the root.

There is reason to question the general applicability of any one method to any data set. Schultz *et al.* (1996) showed that more rapidly evolving characters are estimated with less accuracy, and Schluter *et al.* (1997) confirmed the dependence of accuracy on rates by using maximum likelihood approaches. Similarly, Alroy (1998, 2000) suggested that the evolutionary dynamics of rapid changing traits may well be beyond the scope of any of the models considered, since evolutionary trends will be more likely and more abundant. Other studies have also indicated the serious effect of such evolutionary trends (Garland *et al.*, 1999; Oakley and Cunningham, 2000; Webster and Purvis, 2002a). Although the current analysis suggests preliminarily that IE accommodates effects of selection on trait evolution in a more straightforward and accurate way than established methods, some trends (e.g. strong overall directional trends) are not apparent in extant species and thus cannot be estimated based only on extant data. A possible solution here may be to incorporate known fossil values into the estimation process (e.g. Oakley and Cunningham, 2000; Finarelli and Flynn, 2006).

Overall, our results suggest that using an adaptive peak model of evolution that includes both BM and OU processes of evolution as special cases significantly improves accuracy when reconstructing macro-evolutionary patterns in a primate phylogeny.

ACKNOWLEDGEMENTS

We thank M.L. Rosenzweig and a number of anonymous referees for their very useful comments on a previous version of our manuscript. We would also like to thank A. Friday and R.W. Wrangham for useful comments and discussion. This work was funded by the Cambridge European Trust and the Leverhulme Trust.

REFERENCES

- Aiello, L.C. and Dunbar, R.I.M. 1993. Neocortex size, group size, and the evolution of language. *Curr. Anthropol.*, **34**: 184–193.
- Alroy, J. 1998. Cope's Rule and the dynamics of body mass evolution in North American fossil mammals. *Science*, **280**: 731–734.
- Alroy, J. 2000. Understanding the dynamics of trends within evolving lineages. *Paleobiology*, **26**: 319–329.
- Armstrong, E. 1982a. An analysis of brain allometry: consideration of the cerebral metabolic demand. *Am. J. Phys. Anthropol.*, **57**: 167–168.
- Armstrong, E. 1982b. A look at relative brain size in mammals. *Neurosci. Lett.*, **34**: 101–104.
- Armstrong, E. 1983. Metabolism and relative brain size. *Science*, **220**: 1302–1304.
- Bauchot, R. and Stephan, H. 1969. Encephalisation et niveau évolutif chez les simiens. *Mammalia*, **3**: 225–275.
- Bohonak, A.J. and van der Linde, K. 2004. *RMA: Software for Reduced Major Axis Regression, Java Version* (available at: <http://www.kimvdlinde.com/professional/rma.html>).
- Butler, M.A. and King, A.A. 2004. Phylogenetic Comparative Analysis: a modeling approach for adaptive evolution. *Am. Nat.*, **164**: 683–695.
- Cunningham, C.W., Omland, K.O. and Oakley, T.H. 1998. Reconstructing ancestral character states: a critical reappraisal. *Trends Ecol. Evol.*, **13**: 361–366.
- Diaz-Uriarte, R. and Garland, T. 1996. Testing hypotheses of correlated evolution using phylogenetically independent contrasts: sensitivity to deviations from Brownian motion. *Syst. Biol.*, **45**: 27–47.
- Diaz-Uriarte, R. and Garland, T. 1998. Effects of branch lengths errors on the performance of phylogenetically independent contrasts. *Syst. Biol.*, **47**: 654–672.

- Edwards, A.W.F. and Cavalli-Sforza L.L. 1964. Reconstruction of evolutionary trees. *Phenetic and Phylogenetic Classification*, **6**: 67–76.
- Eisenberg, J.F. 1981. *The Mammalian Radiations*. Chicago, IL: University of Chicago Press.
- Farris, J.S. 1967. The meaning of relationship and taxonomic procedure. *Syst. Zool.*, **16**: 44–51.
- Farris, J.S. 1970. Methods for computing Wagner trees. *Syst. Zool.*, **19**: 83–92.
- Farris, J.S. 1972. Estimating phylogenetic trees from distance matrices. *Am. Nat.*, **951**: 645–668.
- Felsenstein, J. 1985. Phylogenies and the comparative method. *Am. Nat.*, **125**: 1–15.
- Felsenstein, J. 1988. Phylogenies and quantitative characters. *Annu. Rev. Ecol. Syst.*, **19**: 445–471.
- Finarelli, J.A. and Flynn, J.J. 2006. Ancestral state reconstruction of body size in the caniformia (Carnivora, Mammalia): the effects of incorporating data from the fossil record. *Syst. Biol.*, **55**: 301–313.
- Fleagle, J.G. 1999. *Primate Adaptation and Evolution*, 2nd edn. San Diego, CA: Academic Press.
- Garland, T. and Diaz-Uriarte, R. 1999. Polytomies and phylogenetically independent contrasts: examination of the bounded degrees of freedom approach. *Syst. Biol.*, **48**: 547–558.
- Garland, T., Midford, P.E. and Ives, A.R. 1999. An introduction to phylogenetically-based statistical methods with a new method for confidence intervals on ancestral values. *Am. Zool.*, **39**: 374–388.
- Gittleman, J.L. and Luh, H.-K. 1992. On comparative methods. *Annu. Rev. Ecol. Syst.*, **23**: 383–404.
- Gittleman, J.L. and Luh, H.-K. 1994. Phylogeny, evolutionary models and comparative methods: a simulation study. In *Phylogenetics and Ecology* (P. Eggleton and D. Vane-Wright, eds.), pp. 103–122. London: Academic Press.
- Hansen, T.F. 1997. Stabilizing selection and the comparative analysis of adaptation. *Evolution*, **51**: 1341–1351.
- Hansen, T.F. and Orzack, S.H. 2005. Assessing current adaptation and phylogenetic inertia as explanation of trait evolution: the need for controlled comparisons. *Evolution*, **59**: 2063–2072.
- Hansen, T.F., Pienaar, J. and Orzack, S.H. 2008. A comparative method for studying adaptation to a randomly evolving environment. *Evolution*, **62**: 1965–1977.
- Harvey, P.H. and Pagel, M. 1991. *The Comparative Method in Evolutionary Biology*. Oxford: Oxford University Press.
- Harvey, P.H. and Rambaut, A. 2000. Comparative analyses for adaptive radiations. *Phil. Trans. R. Soc. Lond. B*, **355**: 1599–1605.
- Housworth, E.A., Martins, E.P. and Lynch, M. 2004. The phylogenetic mixed model. *Am. Nat.*, **163**: 84–96.
- Kaplan, H., Mueller, T., Gangestad, S. and Lancaster, J.B. 2003. Neural capital and life span evolution among primates and humans. In *Brain and Longevity* (C.E. Finch, J.M. Robine and Y. Christen, eds.), pp. 69–98. Berlin: Springer.
- Lee, M.S.Y. and Shine, R. 1998. Reptilian viviparity and Dollo's law. *Evolution*, **52**: 1441–1450.
- Maddison, W.P. 1991. Squared-change parsimony reconstructions of ancestral states for continuous valued characters on a phylogenetic tree. *Syst. Zool.*, **40**: 304–314.
- Martin, R.D. 1982. Allometric approaches to the evolution of the primate nervous system. In *Primate Brain Evolution: Methods and Concepts* (E. Armstrong and D. Falk, eds.), pp. 39–56. New York: Plenum Press.
- Martins, E.P. 1993. Comparative studies, phylogenies, and predictions of co-evolutionary relationships. *Behav. Brain Sci.*, **16**: 714–716.
- Martins, E.P. 1999. Estimation of ancestral states of continuous characters: a computer simulation study. *Syst. Biol.*, **48**: 642–650.
- Martins, E.P. 2004. *COMPARE, version 4.6b*. Computer programs for the statistical analysis of comparative data (available at: <http://compare.bio.indiana.edu/>). Bloomington, IN: Department of Biology, Indiana University.

- Martins, E.P. and Garland, T. 1991. Phylogenetic analyses of the correlated evolution of continuous characters: a simulation study. *Evolution*, **45**: 534–557.
- Martins, E.P., Diniz, J.A.F. and Housworth, E.A. 2002. Adaptive constraints and the phylogenetic comparative method: a computer simulation test. *Evolution*, **56**: 1–13.
- McMahon, T. and Bonner, J.T. 1983. *On Size and Live*. New York: Scientific American Books.
- Moyà-Solà, S., Köhler, M., Alba, D.M., Casanovas-Vilar, I. and Galindo, J. 2004. *Pierolapithecus catalaunicus*, a new middle Miocene great ape from Spain. *Science*, **306** (5700): 1339–1344.
- Oakley, T.H. and Cunningham, C.W. 2000. Independent contrasts succeed where ancestor reconstruction fails in a known bacteriophage phylogeny. *Evolution*, **54**: 397–405.
- Pagel, M. 1997. Inferring evolutionary processes from phylogenies. *Zool. Scripta*, **26**: 331–348.
- Pagel, M. 1999. Inferring the historical patterns of biological evolution. *Nature*, **401**: 877–884.
- Platt, J.R. 1964. Strong inference. *Science*, **146** (3642): 347–353.
- Polly, P.D. 2001. Paleontology and the comparative method: ancestral node reconstructions versus observed values. *Am. Nat.*, **157**: 596–609.
- Price, T. 1997. Correlated evolution and independent contrasts. *Phil. Trans. R. Soc. Lond. B*, **352**: 519–529.
- Ryan, M.J. 1996. Phylogenetics in behavior: some cautions and expectations. In *Phylogenies and the Comparative Method in Animal Behavior* (E.P. Martins, ed.), pp. 1–21. New York: Oxford University Press.
- Schluter, D., Price, T., Mooers, A.O. and Ludwig, D. 1997. Likelihood of ancestor states in adaptive radiation. *Evolution*, **51**: 1699–1711.
- Schultz, T.R., Cocroft, R.B. and Churchill, G.A. 1996. The reconstruction of ancestral character states. *Evolution*, **50**: 504–511.
- Smith, R.J. and Cheverud, J.M. 2002. Scaling of sexual dimorphism in body mass: a phylogenetic analysis of Rensch's rule in primates. *Int. J. Primatol.*, **5**: 1095–1135.
- Sokal, R.R. and Rohlf, J.F. 1995. *Biometry*, 3rd edn. San Francisco, CA: W.H. Freeman.
- Soligo, C. 2006. Correlates of body mass evolution in primates. *Am. J. Phys. Anthropol.*, **130**: 283–293.
- Swofford, L. and Maddison, W.P. 1987. Reconstructing ancestral states under Wagner parsimony. *Math. Biosci.*, **87**: 199–229.
- Uhlenbeck, G.E. and Ornstein, L.S. 1930. On the theory of Brownian motion. *Phys. Rev.*, **36**: 823–841.
- Webster, A.J. and Purvis, A. 2002a. Testing the accuracy of methods for reconstructing ancestral states of continuous characters. *Proc. R. Soc. Lond. B*, **269**: 143–149.
- Webster, A.J. and Purvis, A. 2002b. Ancestral states and evolutionary rates of continuous characters. In *Morphology, Shape and Phylogenetics* (N. MacLeod and P. Forey, eds.), pp. 247–268. London: Taylor & Francis.
- Westoby, M., Leishman, M.R. and Lord, J.M. 1995. On misinterpreting the phylogenetic correction. *J. Ecol.*, **83**: 531–534.

APPENDIX 1

Fossil data points

Data point in Fig. 4	Fossil	Epoch as proposed by Fleagle (1999)	Assigned age (Mya)	Body size (g)	No. of species
1	<i>Progalago – Komba</i>	Early Miocene	20	685	5
2	<i>Nycticebus simpsoni</i>	Late Miocene	10	500	1
3	<i>Lagonomico – Patasola</i>	Middle to late Miocene	12	1150	1
4	<i>Dolichocebus gainanensis</i>	Early Miocene	16.3	2700	1
5	<i>Neosaimiri fieldsi</i>	Middle to late Miocene	10	840	1
6	<i>Tremacebus harringtoni</i>	Early Miocene	19	1800	1
7	<i>Aotus dindensis</i>	Middle to late Miocene	10	1000	1
8	<i>Homunculus</i>	Early to middle Miocene	14.5	2700	1
9	<i>Cebupithecia sarmientoni</i>	Middle to late Miocene	12	2200	1
10	<i>Stirtonia victoriae</i>	Middle to late Miocene	12	10000	1
11	<i>Stirtonia tatacoensis</i>	Middle to late Miocene	10	5800	1
12	<i>Libypithecus</i>	Late Miocene	5	8400	1
13	<i>Aegyptopithecus</i>	Oligocene	35	6700	1
14	<i>Meso – Dolichopithecus</i>	Late Miocene	5	10333	3
15	<i>Presbytis</i>	Pleistocene	1	7000	1
16	<i>Sennopithecus</i>	Late Miocene to recent	1.5	8000	1
17	<i>Trachypithecus</i>	Pleistocene	1	7000	1
18	<i>Victoriapithecinae</i>	Early Miocene	23	13000	3
19	<i>Macaca</i>	Late Miocene	5	12125	4
20	<i>Parapapio</i>	Late Miocene to early Pleistocene	6.7	22250	4
21	<i>Proconsulidae</i>	Early Miocene	23	14707	13
22	<i>Dryopithecus</i>	Early to middle Miocene	18	27000	3
23	<i>Pierolapithecus catalaunicus</i>	Late Miocene	13	30000	1
24	<i>Sivapithecus</i>	Pliocene	8	68333	3
25	<i>Australopithecus afarensis</i>	Pliocene	4	36950	1
26	<i>Australopithecus africanus</i>	Pliocene	3	35500	1
27	<i>Homo habilis – rudolfensis</i>	Pleistocene	2	43275	2
28	<i>Homo erectus</i>	Pleistocene	1	57650	1

Note: Fossil data points (see Fig. 4) with their respective age, body size, and sample size. Data for *Pierolapithecus* was taken from Moyà-Solà *et al.* (2004); all other data from Fleagle (1999). When only broader age periods of fossils are available, exact fossil age is attributed based on the phylogenetic relationship for which most agreement exists within the range specified by the literature (for details of the procedure, see text).

APPENDIX 2

Overview of the consistent three-step procedure used to assign a specific branch and age to a fossil data point

Epoch	Fossil	Age based on broad period of time (epoch)	Specific age based on dating of site (Mya)	Phyletic relationship (age period of branch or node)	Final attributed age (Mya)
Oligocene	<i>Aegyptopithecus</i>	34–23		Node	35
Early Miocene	<i>Progalago – Komba</i>	23–18		Node	20
	<i>Victoriapithecinae</i>	23–18		Branch	23
	<i>Proconsulidae</i>	23–18		Branch	23
	<i>Tremacebus harringtoni</i>	23–18	21–18	Branch	19
	<i>Dolichocebus gainanensis</i>	23–18	21–18	Node	16.3
Early to middle Miocene	<i>Dryopithecus</i>	23–14		Branch	18
	<i>Homunculus</i>	23–14	18–15.5	Node	14.5
Middle to late Miocene	<i>Pierolapithecus catalaunicus</i>	14–5		Branch	13
	<i>Lagonomico – Patasola</i>	14–5	12	Branch	12
	<i>Cebupithecina sarmitoni</i>	14–5	12	Branch	12
	<i>Stirtonia victoriae</i>	14–5	12	Branch	12
	<i>Aotus dindensis</i>	14–5	10	Branch	10

(continued)

APPENDIX 2—Continued

Epoch	Fossil	Age based on broad period of time (epoch)	Specific age based on dating of site (Mya)	Phyletic relationship (age period of branch or node)	Final attributed age (Mya)
Late Miocene	<i>Nycticebus simpsoni</i>	10–5		Node	10
	<i>Neosaimiri fieldsi</i>	10–5	10	Branch	10
	<i>Stirtonia tatacoensis</i>	10–5	10	Branch	10
	<i>Sivapithecus</i>	10–5		Branch	8
	<i>Libypithecus</i>	10–5		Branch	5
	<i>Meso – Dolichopithecus</i>	10–5		Branch	5
	<i>Macaca</i>	10–5		Branch	5
Late Miocene to Pleistocene	<i>Parapapio</i>	10–1.7		Node	6.7
Pliocene	<i>Australopithecus afarensis</i>	5–1.7	4	Branch	4
	<i>Australopithecus africanus</i>	5–1.7	3	Branch	3
Late Miocene to recent	<i>Sennopithecus</i>	5–0		Node	1.5
Pleistocene	<i>Homo habilis – rudolfensis</i>	1.7–0	2	Branch	2
	<i>Homo erectus</i>	1.7–0		Branch	1
	<i>Presbytis</i>	1.7–0		Node	1
	<i>Trachypithecus</i>	1.7–0		Node	1

Note: For details of the procedure, see text.

APPENDIX 3

Actual values of fossil data points and estimates of them by different methods

Fossil	Fossil name	Fossil value	IE	ML-BM	SCP	IC	UIC	OU _{0.5}	OU ₁	OU _{1.5}	OU ₂
1	<i>Progalago – Komba</i>	685	679	474	543	953	893	6964	6939	7003	7054
2	<i>Nycticebus simpsoni</i>	500	508	461	461	747	866	5888	5891	5944	5987
3	<i>Lagonimico conclutatus</i>	1150	1173	853	610	1440	1667	6713	6937	7003	7054
4	<i>Dolichocebus gaimanensis</i>	2700	2872	1967	1891	2001	2809	6947	6939	7003	7054
5	<i>Neosaimiri fieldsi</i>	840	1971	1448	1410	1672	1970	4701	5757	6478	6833
6	<i>Tremacebus harringtoni</i>	1800	2718	1642	960	3117	3761	6949	6938	7003	7054
7	<i>Aotus dindensis</i>	1000	1083	904	810	1858	1819	6692	6935	7003	7054
8	<i>Homunculus</i>	2700	2562	1970	1833	3166	3395	6951	6939	7003	7054
9	<i>Cebupithecia sarmientoni</i>	2200	2617	2189	2105	3010	3195	6892	6937	7003	7054
10	<i>Stirtonia victoriae</i>	10000	6522	6976	6594	5972	5658	6885	6928	7001	7053
11	<i>Stirtonia tatacoensis</i>	5800	6489	6753	6391	5993	5735	6787	6915	7000	7053
12	<i>Libypithecus</i>	8400	8234	7784	8175	8695	8453	7221	7063	7049	7071
13	<i>Aegyptopithecus</i>	6700	9082	21561	13419	9016	12793	6965	6939	7003	7054
14	<i>Meso-dolichopithecus</i>	10333	8273	9829	10969	8184	8239	7248	6991	7009	7054
15	<i>Presbytis</i>	7000	6529	6328	6304	7953	8571	8842	8595	8177	7828
16	<i>Sennopithecus</i>	8000	7427	7376	7894	7273	7560	7676	7412	7189	7116
17	<i>Trachypithecus</i>	7000	7084	6783	6634	7819	7643	7732	7539	7333	7229
18	<i>Victoriapithecus</i>	13000	10810	30303	21575	15657	19861	6973	6939	7003	7054
19	<i>Macaca</i>	12125	8982	9749	9288	9586	9291	7724	7089	7035	7060
20	<i>Parapapio</i>	22250	10674	14942	13225	11814	12278	8420	7001	7005	7054
21	<i>Proconsul</i>	14707	14231	38559	29476	20470	25131	6972	6939	7003	7054
22	<i>Dryopithecus</i>	27000	28158	51117	50353	37235	35917	7005	6939	7003	7054
23	<i>Pierolapithecus</i>	30000	43545	60220	76093	63192	52122	8529	6971	7004	7054
24	<i>Sivapithecus</i>	68333	49637	57464	63123	54920	49315	31128	31073	31105	31131
25	<i>Australopithecus afarensis</i>	36950	46565	51700	51700	58536	60187	28028	24766	24675	24703
26	<i>Australopithecus africanus</i>	35500	49924	53775	53775	58902	60140	36021	33575	33506	33527
27	<i>Homo habilis</i>	43275	53282	55850	55850	59268	60093	44014	42383	42338	42351
28	<i>Homo erectus</i>	57650	56641	57925	57925	59634	60047	52007	51192	51169	51176

