

Cytochrome *b* Phylogeny and the Taxonomy of Great Apes and Mammals

Jose Castresana

European Molecular Biology Laboratory (EMBL), Biocomputing Unit, Heidelberg, Germany

In the Linnaean system of classification, the generic status of a species is part of its binomial name, and it is therefore important that the classification at the level of genus is consistent at least in related groups of organisms. Using maximum-likelihood phylogenetic trees constructed from a large number of complete or nearly complete mammalian cytochrome *b* sequences, I show that the distributions of intrageneric and intergeneric distances derived from these trees are clearly separated, which allows the limits for a more rational generic classification of mammals to be established. The analysis of genetic distances among hominids in this context provides strong support for the inclusion of humans and chimpanzees in the same genus. It is also of interest to decipher the main reasons for the possible biases in the mammalian classification. I found by correlation analysis that the classification of mammals of large body size tends to be oversplit, whereas that of small mammals has an excess of lumping, which may be a manifestation of the larger difficulty in finding diagnostic characters in the classification of small animals. In addition, and contrary to some previous observations, there is no correlation between body size and rate of cytochrome *b* evolution in mammals, which excludes the difference in evolutionary rates as the cause of the observed body size taxonomic bias.

Introduction

The cytochrome *b* gene has been used in numerous studies of phylogenetic relationships within mammals, and it is the gene for which the most sequence information from different mammalian species is available (Irwin, Kocher, and Wilson 1991; Meyer 1994; Johns and Avise 1998). The sequence variability of cytochrome *b* makes it most useful for the comparison of species in the same genus or the same family. The results obtained in many of the phylogenetic studies in which this gene has been used led to the proposition of new classification schemes that better reflected the phylogenetic relationships among the species studied (Arnason et al. 1995; Lara, Patton, and da Silva 1996; Faulkes et al. 1997; LeDuc, Perrin, and Dizon 1999; Matthee and Robinson 1999). In some works, the new classifications were proposed to preserve monophyletic taxa, but in many others, they were also intended to maintain comparable levels of divergence in groups with the same taxonomic rank in such a way that the new classifications convey more useful comparative information (Avise and Johns 1999). Furthermore, cytochrome *b* phylogenies can help in the genus assignment of newly described species (Giao et al. 1998). However, there are no clear guidelines about which genetic distances between species of the same or different genera are the most representative in mammals, and the use of different methods of tree reconstruction, different models of evolution, and different parameters for these models (Swofford et al. 1996) makes it difficult to compare the phylogenetic proximity among species in different groups.

To determine whether some guidelines can be established, I conducted a comprehensive analysis of complete or nearly complete cytochrome *b* sequences using

a single method of tree reconstruction and genetic distance estimation. In particular, I measured genetic distances from maximum-likelihood trees, and among-site rate variation was allowed in the model of evolution to avoid underestimation of distances among the most divergent species (Golding 1983; Yang 1996). Understanding the general trends and, very importantly, the main reasons for the possible biases in the mammalian classification will facilitate the proposition of more consistent taxonomic revisions from a molecular perspective.

Materials and Methods

Sequences and Alignments

Complete cytochrome *b* sequences (with 1,140 nt) or fragments larger than 1,000 nt and with no more than 12 ambiguous nucleotides were retrieved from the EMBL database (Baker et al. 2000) with sequences released prior to February 7, 2000. When available, cytochrome *b* sequences were extracted from complete mitochondrial genomes. This included all hominid sequences (accession numbers X93334, D38113–D38116). Only one sequence per species was selected, and for every family, one alignment was generated. When sequences of different lengths were included in the same family, only the common part was used in the phylogenetic analysis. No gaps were necessary in the alignments. All sequences were given the binomial name of the corresponding species as it appears in the EMBL database so that all subsequent steps could be automated. Alignments are available at <http://www.molbioevol.org> or <http://www.embl-heidelberg.de/~castresana>.

Phylogenetic Analysis

Phylogenetic analyses were performed using PAUP*, version 4.0b4a (Swofford 1998). The model of evolution used in the different calculations was the HKY model (Hasegawa, Kishino, and Yano 1985) with among-site rate variation assuming a discrete gamma

Key words: cytochrome *b*, body size, mammals, hominids, maximum likelihood, genetic distance.

Address for correspondence and reprints: Jose Castresana, European Molecular Biology Laboratory (EMBL), Biocomputing Unit, Meyerhofstrasse 1, D-69117 Heidelberg, Germany. E-mail: jose.castresana@embl-heidelberg.de.

Mol. Biol. Evol. 18(4):465–471. 2001

© 2001 by the Society for Molecular Biology and Evolution. ISSN: 0737-4038

distribution with six rate categories. It has been shown that the estimation of parameters is more reliable for large trees (Excoffier and Yang 1999). Thus, the parameters necessary for this model were estimated from an alignment of 632 full-length mammalian sequences, and values of 4.2 for the transition/transversion ratio and 0.54 for the gamma rate parameter were obtained. Randomly chosen sets of smaller numbers of sequences produced successively lower values for both parameters and indicated that the values obtained from the complete set were close to the saturation point. For families with only two representatives, the maximum-likelihood distance for the single pair was calculated. For families with three or more representatives, a neighbor-joining tree was first calculated and branch lengths were further optimized by maximum likelihood. Then, distances among all pairs of species were measured as the sum of the branch lengths separating them in the tree (patristic distances).

Statistical Analysis

Pairwise distances were divided into intragenetic and intergeneric distances for their examination. In the histograms of unweighted distances, genera and families with large numbers of species are overwhelmingly represented, since the number of pairwise distances in each genus or family grows in proportion to the square of the number of species. On the other hand, taking averages of pairwise distances in every genus and family, as was previously done (Johns and Avise 1998), means that the information about the most biased values—one of the interests in the present work—gets lost. Thus, a weight was given to each distance so that every genus or family contributes to the histogram proportionally to the number of species and not to the square of this number. Specifically, intragenetic distances were given a weight of $2/(S - 1)$, where S is the number of species in the corresponding genus, and intergeneric distances were given a weight of $1/D^{1/2}$, where D is the number of intergeneric distances in the corresponding family. This ensures that all intragenetic distances in a genus contribute to the histogram with a value equal to the number of species in this genus, and all intergeneric distances in a family contribute with a value equal to the square root of the number of intergeneric distances, which grows approximately in proportion to the number of species in the family. Other weighting schemes with the same objective produced similar results.

Body masses of the species analyzed here were taken from Silva and Downing (1995). Correlations of the logarithm of the average body mass and the maximum intragenetic distance in each genus were calculated with the program JMP, version 3.2.6 (SAS Institute, Cary, N.C.). Phylogenetically independent comparisons (Harvey and Pagel 1991) were not used for this calculation because the maximum intragenetic distance in each genus depends on a taxonomic decision.

For the computation of the correlation of body mass and evolutionary rates, I used phylogenetically independent comparisons (Harvey and Pagel 1991). All

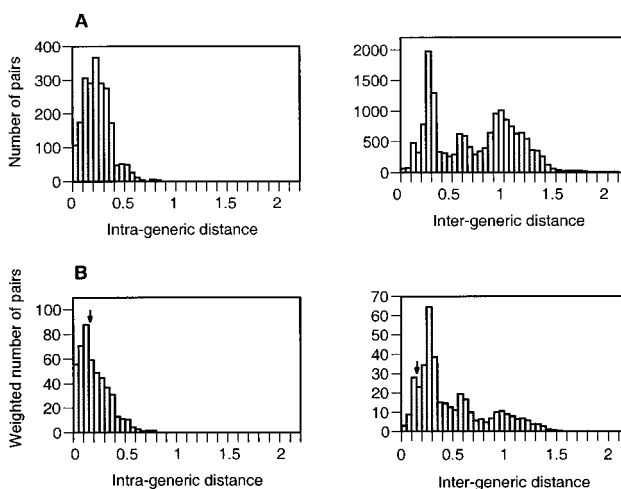


FIG. 1.—Histograms of unweighted (A) and weighted (B) intra-genetic and intergeneric pairwise distances, in substitutions per site, measured from cytochrome *b* trees of mammals. The arrows show the distance between humans and chimpanzees (0.150–0.160 substitutions per site) in both distributions.

trees were rooted with midpoint rooting, and terminal pairs of sister taxa (of the same genus or a different genus) for which the weights were available in Silva and Downing (1995) were compared. If the longest branch in each pair is nonrandomly associated with the biggest (or smallest) species, then a positive (or negative) correlation will be found. Thus, for each pair of species, the difference of the logarithm of the body mass was compared with the difference of the logarithm of the branch length from their last common ancestor, as was described before (Bromham, Rambaut, and Harvey 1996), and correlations of these two variables were calculated. Both the Pearson product-moment correlation and the Spearman rank correlation produced very similar results, and only the former is reported.

Results and Discussion

Intragenetic and Intergeneric Distance Distributions

All complete or nearly complete cytochrome *b* sequences available from mammals were extracted from the DNA databases. A total of 688 mammalian species, distributed in 310 genera and 52 families, were found. This number represents 15% of the known extant species (Wilson and Reeder 1993). Sequence alignments and phylogenetic trees were constructed separately for each family. A large proportion of the genera were rendered monophyletic, showing an overall good agreement between the generic classification and the molecular tree topology. To measure more precisely the divergence levels, genetic distances were computed as the sum of branch lengths separating pairs of species in the trees. Pairwise distances were divided into intragenetic (between species of the same genus) and intergeneric (between species of different genera but same family) distances, and their distributions were analyzed separately.

In the histograms of unweighted distances (fig. 1A), a few genera and families with a large number of species were overwhelmingly represented, since the number of

Table 1
Genera that Show Intrageneric Distances Among Some of Their Members that are Longer than the Limit of the Modal Class of the Intergeneric Distances (excess of lumping)

Family	Genera
Didelphidae (American opossums)	Monodelphis
Dasyuridae (marsupial mice and cats, Tasmanian devil).	Antechinus, Dasyurus, Ningai, Planigale, Pseudantechinus, Sminthopsis
Soricidae (shrews).	Sorex
Phyllostomidae (New World leaf nosed bats)	Artibeus
Sciuridae (squirrels, marmots).	Spermophilus
Muridae (rats, mice, hamsters, voles, gerbils)	Acomys, Akodon, Apodemus, Bolomys, Eliurus, Microtus, Mus, Oxymycterus, Peromyscus, Rattus, Reithrodontomys, Thomasomys
Geomyidae (pocket gophers)	Cratogeomys, Thomomys
Echimyidae (spiny rats)	Echimyus, Trinomys
Bathyergidae (African mole rats)	Cryptomys

pairwise distances in each genus or family grows in proportion to the square of the number of species. To overcome this problem, which can make comparisons difficult, a weight was given to each distance so that every genus or family contributed to the histograms proportionally to the number of species and not to the square of this number (see *Materials and Methods*). Now there was an important relative reduction of the bands at long distances (at ~ 0.22 substitutions per site in the intrageneric distance distribution and at ~ 1 substitution per site in the intergeneric distribution; fig. 1B), which contained mainly pairwise distances belonging to rodent families and dasyurids, making the histograms more suitable for comparative purposes.

These histograms with weighted distances revealed that the modal classes of the intrageneric and intergeneric distance distributions were well separated, at 0.10–0.15 and 0.25–0.30 substitutions per site, respectively (fig. 1B). The difference in the averages of both distributions was highly significant, indicating that, overall, species classified in the same genus are indeed phylogenetically closer than species of different genera but of the same family. In a previous analysis of cytochrome *b* genetic distances in vertebrates (Johns and Avise 1998), the same significant difference was found, although the distributions of intrageneric and intergeneric distances were not so clearly separated in mammals, probably due to the inclusion of small cytochrome *b* fragments in the analysis and the estimation of distances without rate heterogeneity. The high intrageneric and intergeneric distances found for rodent families and dasyurids already suggests that the classification of these groups follows very different criteria than that for the rest of mammals.

The distinct separation of the intrageneric and intergeneric distance distributions allows the detection of groups of species whose classification is highly deviant from the main trends. The most extreme cases can be identified using the alternate weighted distribution (fig. 1B). Thus, all genera showing intrageneric distances among some of their species that are higher than the limit of the modal class of the intergeneric distance dis-

tribution (i.e., larger than 0.30 substitutions per site) are listed in table 1, and their classification would clearly be more consistent with the split of these genera. In fact, many of these genera are known to be differentiated in distinct clades or subgenera. The analysis of the other distribution, that of intergeneric distances, reveals species separated in different genera (in some cases up to five genera) but genetically very similar (i.e., with distances smaller than 0.10 substitutions per site; table 2). In these species, the classification in a single genus would better reflect their phylogenetic proximity. Furthermore, the highest intergeneric distances, mostly due to rodent and dasyurid genera, would support the split of some of the corresponding families.

Finally, the analysis of the phylogenetic trees revealed that many of the species whose classification is in substantial disagreement with their phylogenetic relatedness are involved in the formation of nonmonophyletic taxa. Of the 27 genera listed in table 1, 10 (Antechinus, Dasyurus, Pseudantechinus, Sminthopsis, Spermophilus, Akodon, Eliurus, Oxymycterus, Trinomys, and Cryptomys) are not monophyletic according to the trees calculated here. In some of these genera, the correction of clearly misclassified species would reduce the high intrageneric distances. On the other hand, a few genera of table 2, such as *Phoca* and *Ursus*, would remain nonmonophyletic if only the closest genera listed in the table were joined. Thus, considering only the species analyzed here, *Phoca* would need the addition of *Cystophora* together with *Halichoerus* to keep the genus monophyletic, whereas *Ursus* would require the inclusion of *Helarctos* and *Selenarctos* together with *Thal-arctos* for the same purpose. Obviously, the taxonomic revision of all of these groups would need a case-by-case analysis, correcting for nonmonophyly issues and taking the taxonomic spread into account. Furthermore, it would be desirable to examine additional molecular data to ensure, for example, that accelerated evolutionary rates are not inducing divergences that are too high in some particular case or that nuclear inserts are not responsible for some of the distances that are too short.

Table 2
Groups of Genera that Show Intergeneric Distances Shorter than the Modal Class of the Intrageneric Distances (excess of splitting)

Family	Groups of Genera
Otariidae (sea lions)	Eumetopias/Zalophus
Phocidae (seals)	Halichoerus/Phoca Hydrurga/Leptonychotes
Ursidae (bears)	Thalarcos/Ursus
Elephantidae (elephants)	Elephas/Loxodonta/Mammuthus
Balaenopteridae (rorquals)	Balaenoptera/Megaptera
Delphinidae (dolphins)	Orcaella/Orcinus Sotalia/Steno Delphinus/Sousa/Stenella/Lagenodelphis/Tursiops Cephalorhynchus/Lagenorhynchus/Lissodelphis Feresa/Globicephala/Grampus/Peponocephala/Pseudorca
Monodontidae (beluga, narval)	Delphinapterus/Monodon
Bovidae (cattle, antelopes, bison, goats, sheep)	Adax/Oryx Alcelaphus/Sigmoceros Bison/Bos Capra/Hemitragus
Camelidae (camels, llamas)	Lama/Vicugna
Cervidae (deer)	Megamuntiacus/Muntiacus
Hippopotamidae (hippopotamuses)	Hexaprotodon/Hippopotamus
Tayassuidae (peccaries)	Catagonus/Tayassu

Great Apes Taxonomy

Other genera with distance values not as extreme as those in tables 1 and 2 may also need to change their generic name to make it more consistent with their phylogeny. Recent classification schemes (Groves 1993), as well as the taxonomy adopted by the EMBL and GenBank databases, include the great apes (chimpanzee, bonobo or pygmy chimpanzee, gorilla, and orangutan) in the same family as humans, the Hominidae. Although the phylogeny of hominids is well known (Ruvolo 1997; Satta, Klein, and Takahata 2000), all species except the two chimpanzees are classified in separate genera, and it has been suggested that at least humans and chimpanzees should be classified in the same genus (Diamond 1992, p. 25; Eastal, Collet, and Betty 1995, p. 131; Goodman et al. 1998). The cytochrome *b* genetic distances between humans and both chimpanzee species are 0.150 and 0.160 substitutions per site, respectively. Although these distances are not as small as the distances between different genera reported in table 2, they are much closer to the modal class of the intrageneric distance distribution than to the modal class of the intergeneric distribution (fig. 1B), which would support the inclusion of humans, chimpanzees, and bonobos in the same genus, *Homo* (Diamond 1992, p. 25; Eastal, Collet, and Betty 1995; Goodman et al. 1998). Distances between gorillas and the chimpanzees/humans clade range between 0.166 and 0.190 substitutions per site, closer to an equidistant point from both distributions (fig. 1B), indicating that the generic status of gorillas (*Homo* or *Gorilla*) is more difficult to discern with the analysis of a single locus. Finally, distances between orangutans and the other hominids range between 0.221 and 0.264, close to the modal class of the intergeneric distance distribution, justifying its inclusion in the genus *Pongo*, different from the other hominids.

Toward a More Objective Classification

Although the genetic distance distributions of the cytochrome *b* gene shown in this work can be most easily used to detect groups with highly inconsistent classification in the context of the mammalian relationships, they can also help to estimate an approximate limit of divergence from which two species should be separated into different genera in a biological classification aided by a standardized temporal scheme (Avice and Johns 1999). The minimum disruption of the current mammalian taxonomy would occur by establishing the limit at a genetic distance, when measured with the complete cytochrome *b* gene, of around 0.2 substitutions per site (i.e., in the middle of the intrageneric and intergeneric distance distributions; fig. 1B), or 0.1 substitutions per site per lineage. If the human-chimpanzee divergence happened 5 MYA (see Yoder and Yang 2000), this limit would correspond very approximately to 6–7 MYA, although analyses of more loci should be performed to obtain a better estimate (Avice and Johns 1999). In addition, certain flexibility around this limit (in the form of a pre-established interval) that allows generic divisions to be placed on the longer edges of the tree would be desirable.

Body Size Taxonomic Bias

It is also of interest to decipher the main reasons for the biases in the mammalian classification. Examination of the mammalian genera that are too lumped (table 1) or too split (table 2) indicates that the former are mostly small animals (e.g., rodents or shrews), whereas the latter are bigger ones (e.g., elephants or dolphins). To determine whether this bias is more general, I plotted the maximum intrageneric distance in each genus (as a measure of the genetic diversity in the genus) versus the average body mass of the species in the

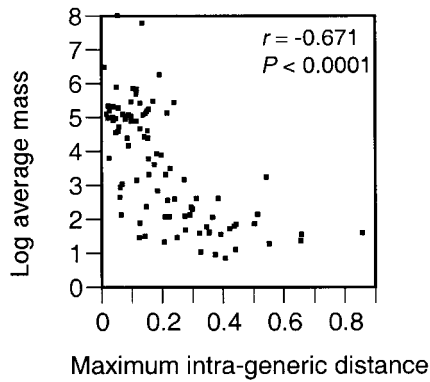


FIG. 2.—Plots of the logarithm of the average body mass, in grams, versus maximum divergence (maximum intrageneric distance) for 94 genera of mammals. The correlation of both variables is highly significant.

genus (fig. 2). For example, the genus *Marmota* contains 14 species (Hoffmann et al. 1993; Steppan et al. 1999), of which 13 had complete cytochrome *b* sequences available in the databases; the maximum genetic distance (0.155 substitutions per site) occurred between *Marmota caligata* and *Marmota himalayana*, two species belonging, respectively, to the two earliest divergent groups, and this distance was plotted against the logarithm of the average body mass in grams of this genus, which was 3.704 (Silva and Downing 1995). The plot of all genera for which at least two species were available shows the tendency of genera of big animals to contain species genetically similar and of genera of smaller animals to embrace species genetically more diverse (fig. 2). In fact, there is a strong negative correlation between the two variables. The use of the average instead of the maximum intrageneric distance as a measure of the genetic diversity or the use of the logarithm of the maximum distance yielded basically the same results.

To test whether some systematic error was causing this correlation, several variations of it were also calculated. First, in repeated samplings in which only one genus was randomly selected per family to avoid any possible overrepresentation of some families in the comparisons, the correlation was significant ($P < 0.001$; mean P measured from 100 samplings). Furthermore, the correlation was significant ($P < 0.0001$) after excluding those genera with $<50\%$ of the species sampled, in which the genetic diversity of the whole genus might not be properly reflected by the measure used when only closely related species were sequenced. The correlation was also significant ($P < 0.0001$) after excluding all dasyurids and rodents, whose classification is the most deviant. Finally, when the maximum intrageneric distances were taken from a single tree of all mammals, and maximum-likelihood branch lengths were estimated forcing the molecular clock in order to compensate for different evolutionary rates, the correlation was also highly significant ($P < 0.0001$). Taking these correlation analyses together with the obvious body mass differences between the genera with the

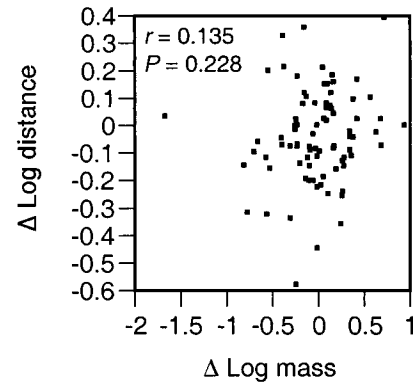


FIG. 3.—Plots that measure whether changes in cytochrome *b* evolutionary rates are correlated with body mass. In the plots, the difference of the logarithm of body mass is compared with the difference of the logarithm of the branch length from the last common ancestor in 82 terminal pairs of mammalian species taken from cytochrome *b* trees. There is no significant correlation between both variables.

most biased classification (tables 1 and 2), it seems that taxonomic practice has tended, in general, to oversplit the classification of big mammals, as well as to group small ones in the same genus even if they are genetically very divergent, perhaps due to the larger difficulty in finding diagnostic characters in the classification of small animals.

Cytochrome *b* Evolutionary Rates

Finally, it was also necessary to test whether there was a systematic variation of the cytochrome *b* rates with body size, since it is well known that there are differences in the evolutionary rates of the cytochrome *b* gene among different lineages (Kocher et al. 1989; Andrews, Jermin, and Easteal 1998). Some authors have postulated that the metabolic rate (Martin and Palumbi 1993; Nunn and Stanley 1998) or the generation time (Bromham, Rambaut, and Harvey 1996; Li et al. 1996), which, in turn, are correlated with the body mass, could be responsible for the different rates. Although these hypotheses were put forward using a limited number of species (Bromham, Rambaut, and Harvey 1996), the possibility exists that the observed correlation between maximum cytochrome *b* divergence and body mass (fig. 2) only reflects that the cytochrome *b* sequences of small animals have faster evolutionary rates. Now this can be tested with the large sample of cytochrome *b* sequences used here. Using phylogenetically independent comparisons (Bromham, Rambaut, and Harvey 1996) of 82 terminal pairs of mammals taken from the cytochrome *b* trees (see *Materials and Methods*), I have shown that no correlation exists between body mass and maximum-likelihood distance since the last common ancestor of each pair (fig. 3), at least at the divergence levels analyzed in this work. Therefore, the observed correlation between body mass and genetic diversity in mammalian genera (fig. 2) cannot be explained by an evolutionary rate acceleration in small an-

imals; it is, rather, a reflection of a body size taxonomic bias.

Conclusions

Thus, despite the body size taxonomic bias in the mammalian classification, which causes most of the overlap between the intrageneric and the intergeneric distance distributions of the cytochrome *b* gene, the quality of the classification is good enough that these distributions are distinctly separated. In addition, these distance distributions can be used as objective references to improve the taxonomy of particular groups, such as the great apes and humans, in the proper context of the mammalian relationships.

Acknowledgment

I thank Toby Gibson for useful help during the work.

LITERATURE CITED

- ANDREWS, T. D., L. S. JERMIIN, and S. EASTEAL. 1998. Accelerated evolution of cytochrome *b* in simian primates: adaptive evolution in concert with other mitochondrial proteins? *J. Mol. Evol.* **47**:249–257.
- ARNASON, U., K. BODIN, A. GULLBERG, C. LEDJE, and S. MOUCHATY. 1995. A molecular view of pinniped relationships with particular emphasis on the true seals. *J. Mol. Evol.* **40**:78–85.
- AVISE, J. C., and G. C. JOHNS. 1999. Proposal for a standardized temporal scheme of biological classification for extant species. *Proc. Natl. Acad. Sci. USA* **96**:7358–7363.
- BAKER, W., A. VAN DEN BROEK, E. CAMON, P. HINGAMP, P. STERK, G. STOESSER, and M. A. TULI. 2000. The EMBL nucleotide sequence database. *Nucleic Acids Res.* **28**:19–23.
- BROMHAM, L., A. RAMBAUT, and P. H. HARVEY. 1996. Determinants of rate variation in mammalian DNA sequence evolution. *J. Mol. Evol.* **43**:610–621.
- DIAMOND, J. M. 1992. *The third chimpanzee: the evolution and future of the human animal*. HarperCollins, New York.
- EASTEAL, S., C. COLLET, and D. BETTY. 1995. *The mammalian molecular clock*. R. G. Landes, Austin, Texas.
- EXCOFFIER, L., and Z. YANG. 1999. Substitution rate variation among sites in mitochondrial hypervariable region I of humans and chimpanzees. *Mol. Biol. Evol.* **16**:1357–1368.
- FAULKES, C. G., N. C. BENNETT, M. W. BRUFORD, H. P. O'BRIEN, G. H. AGUILAR, and J. U. JARVIS. 1997. Ecological constraints drive social evolution in the African mole-rats. *Proc. R. Soc. Lond. B Biol. Sci.* **264**:1619–1627.
- GIAO, P. M., D. TUOC, V. V. DUNG, E. D. WIKRAMANAYAKE, G. AMATO, P. ARCTANDER, and J. R. MACKINNON. 1998. Description of *Muntiacus truongsongensis*, a new species of muntjac (Artiodactyla: Muntiacidae) from central Vietnam, and implications for conservation. *Anim. Conserv.* **1**:61–68.
- GOLDING, G. B. 1983. Estimates of DNA and protein sequence divergence: an examination of some assumptions. *Mol. Biol. Evol.* **1**:125–142.
- GOODMAN, M., C. A. PORTER, J. CZELUSNIAK, S. L. PAGE, H. SCHNEIDER, J. SHOSHANI, G. GUNNELL, and C. P. GROVES. 1998. Toward a phylogenetic classification of Primates based on DNA evidence complemented by fossil evidence. *Mol. Phylogenet. Evol.* **9**:585–598.
- GROVES, C. P. 1993. Order primates. Pp. 243–277 in D. E. WILSON and D. M. REEDER, eds. *Mammal species of the world: a taxonomic and geographic reference*. Smithsonian Institution Press, Washington, D.C.
- HARVEY, P. H., and M. D. PAGEL. 1991. *The comparative method in evolutionary biology*. Oxford University Press, Oxford, England.
- HASEGAWA, M., H. KISHINO, and T. YANO. 1985. Dating of the human-ape splitting by a molecular clock of mitochondrial DNA. *J. Mol. Evol.* **22**:160–174.
- HOFFMANN, R. S., C. G. ANDERSON, R. W. THORINGTON, and L. R. HEANEY. 1993. Family sciuridae. Pp. 419–465 in D. E. WILSON and D. M. REEDER, eds. *Mammal species of the world: a taxonomic and geographic reference*. Smithsonian Institution Press, Washington, D.C.
- IRWIN, D. M., T. D. KOCHER, and A. C. WILSON. 1991. Evolution of the cytochrome *b* gene of mammals. *J. Mol. Evol.* **32**:128–144.
- JOHNS, G. C., and J. C. AVISE. 1998. A comparative summary of genetic distances in the vertebrates from the mitochondrial cytochrome *b*. *Mol. Biol. Evol.* **15**:1481–1490.
- KOCHER, T. D., W. K. THOMAS, A. MEYER, S. V. EDWARDS, S. PÄÄBO, F. X. VILLABLANCA, and A. C. WILSON. 1989. Dynamics of mitochondrial DNA evolution in animals: amplification and sequencing with conserved primers. *Proc. Natl. Acad. Sci. USA* **86**:6196–6200.
- LARA, M. C., J. L. PATTON, and M. N. DA SILVA. 1996. The simultaneous diversification of South American echimyid rodents (Hystricognathi) based on complete cytochrome *b* sequences. *Mol. Phylogenet. Evol.* **5**:403–413.
- LEDUC, R. G., W. F. PERRIN, and A. E. DIZON. 1999. Phylogenetic relationships among the delphinid cetaceans based on full cytochrome *b* sequences. *Mar. Mamm. Sci.* **15**:619–648.
- LI, W. H., D. L. ELLSWORTH, J. KRUSHKAL, B. H. CHANG, and D. HEWETT-EMMETT. 1996. Rates of nucleotide substitution in primates and rodents and the generation-time effect hypothesis. *Mol. Phylogenet. Evol.* **5**:182–187.
- MARTIN, A. P., and S. R. PALUMBI. 1993. Body size, metabolic rate, generation time, and the molecular clock. *Proc. Natl. Acad. Sci. USA* **90**:4087–4091.
- MATTHEE, C. A., and T. J. ROBINSON. 1999. Cytochrome *b* phylogeny of the family bovidae: resolution within the alcelaphini, antilopini, neotragini, and tragelaphini. *Mol. Phylogenet. Evol.* **12**:31–46.
- MEYER, A. 1994. Shortcomings of the cytochrome *b* gene as a molecular marker. *Trends Ecol. Evol.* **9**:278–280.
- NUNN, G. B., and S. E. STANLEY. 1998. Body size effects and rates of cytochrome *b* evolution in tube-nosed seabirds. *Mol. Biol. Evol.* **15**:1360–1371.
- RUVOLO, M. 1997. Molecular phylogeny of the hominoids: inferences from multiple independent DNA sequence data sets. *Mol. Biol. Evol.* **14**:248–265.
- SATTA, Y., J. KLEIN, and N. TAKAHATA. 2000. DNA archives and our nearest relative: the trichotomy problem revisited. *Mol. Phylogenet. Evol.* **14**:259–275.
- SILVA, M., and J. A. DOWNING. 1995. *CRC handbook of mammalian body masses*. CRC Press, Boca Raton, Fla.
- STEPAN, S. J., M. R. AKHVERDYAN, E. A. LYAPUNOVA, D. G. FRASER, N. N. VORONTSOV, R. S. HOFFMANN, and M. J. BRAUN. 1999. Molecular phylogeny of the marmots (Rodentia: Sciuridae): tests of evolutionary and biogeographic hypotheses. *Syst. Biol.* **48**:715–734.

- SWOFFORD, D. L. 1998. PAUP*: phylogenetic analysis using parsimony (*and other methods). Version 4. Sinauer, Sunderland, Mass.
- SWOFFORD, D. L., G. J. OLSEN, P. J. WADDELL, and D. M. HILLIS. 1996. Phylogenetic inference. Pp. 407–514 *in* D. M. HILLIS, C. MORITZ, and B. K. MABLE, eds. Molecular systematics. Sinauer, Sunderland, Mass.
- WILSON, D. E., and D. M. REEDER. 1993. Mammal species of the world: a taxonomic and geographic reference. 2nd edition. Smithsonian Institution Press, Washington, D.C.
- YANG, Z. 1996. Among-site rate variation and its impact on phylogenetic analysis. *Trends Ecol. Evol.* **11**:367–372.
- YODER, A. D., and Z. YANG. 2000. Estimation of primate speciation dates using local molecular clocks. *Mol. Biol. Evol.* **17**:1081–1090.

SIMON EASTEAL, reviewing editor

Accepted November 14, 2000