Cladistics 15, 351–361 (1999)
Article ID clad.1999.0105, available online at http://www.idealibrary.com on IDEAL®

Phylogeny of the Lemuridae: Effects of Character and Taxon Sampling on Resolution of Species Relationships within *Eulemur*

Anne D. Yoder*,†,[1] and Jodi A. Irwin*

**Department of Cell and Molecular Biology, Northwestern University Medical School, Chicago, Illinois 60611; and †Department of Zoology, Field Museum of Natural History, Roosevelt Road at Lake Shore Drive, Chicago, Illinois 60605*

Accepted for publication October 12, 1998

DNA sequences from three mitochondrial genes and one nuclear gene were analyzed to determine the phylogeny of the Malagasy primate family Lemuridae. Whether analyzed separately or in combination, the data consistently indicate that *Eulemur* species comprise a clade that is sister to a *Lemur catta* plus *Hapalemur* clade. The genus *Varecia* is basal to both. Resolution of cladogenic events within *Eulemur* was found to be extremely problematic with a total of six alternative arrangements offered by various data sets and weighting regimes. We attempt to determine the best arrangement of *Eulemur* taxa through a variety of character and taxon sampling strategies. Because our study includes all but one *Eulemur* species, increased taxon sampling is probably not an option for enhancing phylogenetic accuracy. We find, however, that the combined genetic data set is more robust to changes in taxon sample than are any of the individual data sets, suggesting that increased character sampling stabilizes phylogenetic resolution. Nonetheless, due to the difficult nature of the problem, we may have to accept certain aspects of *Eulemur* interrelationships as uncertain.

Key Words: Strepsirrhini; lemurs; congruence; taxon sampling; combined data analysis.

INTRODUCTION

The Lemuridae is the most taxonomically diverse of the five lemuriform families, all of which are endemic to the island of Madagascar. At present, 13 subspecies within 10 species and 4 genera are commonly recognized within this family (Mittermeier *et al.*, 1994). The phylogenetic relationships among taxa, however, are not well understood. The lack of understanding does not reflect lack of investigative energy. In fact, over the past decade, this single family has been the subject of more systematic scrutiny, by far, than the other six strepsirrhine families combined (Crovella *et al.*, 1993; Crovella and Rumpler, 1992; Eaglen, 1980; Groves and Eaglen, 1988; Groves and Trueman, 1995; Macedonia and Stanger, 1994; Randria, 1998; Simons and Rumpler,

[1]To whom correspondence should be addressed at Department of Cell and Molecular Biology, 303 E. Chicago Avenue, Ward Building, Chicago, IL 60611-3008. Fax: 312-503-7912. E-mail: ayoder@nwu.edu.

0748-3007/99 $30.00

1988; Stanger-Hall, 1997; Tattersall, 1988, 1993; Tattersall and Koopman, 1989; Tattersall and Schwartz ,1991; Yoder, 1994). Interest in lemurid taxonomy lay relatively quiescent until 1988, at which time three papers were published, nearly simultaneously, that questioned previously accepted ideas of both phylogeny and taxonomy (Groves and Eaglen, 1988; Simons and Rumpler, 1988; Tattersall, 1988). In particular, the authors of these papers were interested in clarifying the relationship of *Lemur catta* (the ring-tailed lemur) to other lemurid species. Following decades of reports (summarized in Groves and Eaglen, 1988) of special similarities between *L. catta* and the genus *Hapalemur* (bamboo lemurs), the three papers independently recommended new genus-level taxonomy for *Lemur* species other than *L. catta*. *Prosimia* (Tattersall, 1988), *Petterus* (Groves and Eaglen, 1988), and *Eulemur* (Simons and Rumpler, 1988) were the suggested alternatives. As reviewed by Groves and Trueman (1995), the genus *Eulemur* was ruled by the International Commission on Zoological Nomenclature to be the valid taxon.

Eulemur has been widely accepted and now permeates the relevant literature. Even so, the phylogenetic underpinnings for the new taxonomy are not secure. The taxonomy was originally proposed more as a repository for those species orphaned by the recognition of *L. catta* / *Hapalemur* affinities than as a distinct phylogenetic unit (Simons and Rumpler, 1988). In other words, these authors did not explicitly address the issue of *Eulemur* monophyly. Subsequently, numerous studies employing a variety of character sets have investigated the question directly and have found support for a *Eulemur* clade (Crovella *et al.*, 1993; Groves and Trueman, 1995; Macedonia and Stanger, 1994; Stanger-Hall, 1997; Yoder, 1994). Ironically, however, the relatedness of *L. catta* to *Hapalemur* has proven to be questionable. Whereas some studies have found support for the sister group relationship of the two taxa (Adkins and Honeycutt, 1994; Crovella *et al.*, 1995; Jung *et al.*, 1992; Macedonia and Stanger, 1994; Stanger-Hall and Cunningham, 1998; Yoder *et al.*, 1996a; Yoder *et al.*, 1996b), others have not (Groves and Trueman, 1995; Stanger-Hall, 1997; Tattersall, 1993; Tattersall and Schwartz, 1991). An additional complication for phylogenetic (and thus taxonomic) resolution relates to the position of the genus *Varecia* (the ruffed lemur). Among other distinctions, *Varecia* is the only member of the Lemuridae to routinely deliver multiple offspring per birth and, perhaps as a consequence, the only lemurid to nest infants rather than carry them as do other lemurids. These life history and other anomalies have resulted in a virtually exhaustive number of alternative phylogenetic placements for *Varecia*. The genus has been placed outside of the Lemuridae (Macedonia and Stanger, 1994; Stanger-Hall, 1997), basal to other lemurids (Adkins and Honeycutt, 1994; Crovella *et al.*, 1995; Yoder, 1994; Yoder *et al.*, 1996b), sister to *Eulemur* (Groves and Eaglen, 1988; Groves and Trueman, 1995; Tattersall and Schwartz, 1991), and unresolved relative to *Lemur*, *Hapalemur*, and *Eulemur* (Crovella *et al.*, 1993; Stanger-Hall and Cunningham, 1998).

Taxonomy that is congruent with phylogeny has always been of importance to phylogenetic systematists. Even so, calls for phylogenetic taxonomy have become increasingly precise and well reasoned, both in theory (de Queiroz and Gauthier, 1990, 1992, 1994; Lee, 1998; Schander and Thollesson, 1995) and in practice (Bryant, 1996; Cantino *et al.*, 1997; Wyss and Meng, 1996). Our paper is an attempt to unambiguously resolve lemurid phylogeny as a first step towards establishing a phylogenetic taxonomy for this group of primates. Specifically, we employ a variety of genetic markers to assess the relative placement of the genera *Eulemur*, *Lemur*, *Hapalemur*, and *Varecia*. For those genera containing multiple species and / or subspecies, we also wish to confirm their monophyletic status and interrelationships.

MATERIALS AND METHODS

Tissues (liver, spleen, kidney, muscle) for all study taxa were acquired from animals that died of natural causes at the Duke University Primate Center (DUPC). Taxon sampling at the genus level within the Lemuridae is exhaustive. Species-level sampling is nearly complete except for the omission of two of three *Hapalemur* species, *H. simus* and *H. aureus*, and one *Eulemur* species, *E. coronatus*. Total genomic DNA was extracted with a standard phenol / chloroform technique after digesting overnight in a SDS-based extraction buffer. Amplification and sequencing conditions for the entire 1140-bp cytochrome *b* gene are as reported in Yoder *et al.* (1996b). A portion of the mitochondrial control

region (D-loop), homologous with the hypervariable 1 (HV1) region found in humans, was amplified and sequenced with primers L15926 (TCA AAG CTT ACA CCA GTC TTG TAA ACC), L16540 (CCA TCG TGA TGT CTT ATT TAA GGG GAA CGT), and H16498 (CCT GAA GTA GGA ACC AGA TG). The entire cytochrome oxidase subunit II gene (COII) was amplified and directly sequenced using primers described in Adkins and Honeycutt (1994). A 1067-bp fragment of exon 1 of the interphotoreceptor retinoid binding protein (IRBP) was amplified with primers p141 (CTG GTC ATC TCC TAT GAG CCC AGC A) and m1208 (TCA GCA AAG CTG TCG AAG CGC AGG TA) and sequenced with these and two internal primers (p555—CTG GGA GAG AGG TAT GGT GCC GAC AA and m697—ACG GTG AGG AAG AAG TTG GAT TGG). PCR products were cycle sequenced using a dye-terminator sequencing kit (Applied Biosystems, Foster City, CA) and then analyzed by gel electrophoresis with an Applied Biosystems automated DNA sequencer Model 377. Sequences were edited and compiled with AutoAssembler 1.3.0 (Applied Biosystems).

The complete gene sequences are the consensus of at least two different double-stranded PCR amplification reactions for which both strands were sequenced. All protein-coding sequences (cytochrome *b*, COII, and IRBP) were easily aligned by eye due to the lack of insertions and deletions (indels). D-loop sequences within the Lemuridae were also aligned by eye, although CLUSTAL (Higgins and Sharp, 1988; Thompson *et al.*, 1994) was employed for the alignment of the *Microcebus murinus* (the mouse lemur) and *Propithecus tattersalli* (Tattersall's sifaka) outgroup sequences. The D-loop sequences show multiple indels among the different taxa compared; resulting gaps were treated as missing data rather than recoded as present/absent states at the end of the matrix. The alignment is presented in Appendix 1 and has been deposited in TreeBASE. Sequences for all four genes are available from GenBank under the Accession Nos. listed in Table 1.

The branch and bound algorithm in PAUP* 4.0b1 (Swofford, 1998) was employed for parsimony analysis. For bootstrap tests, 100 replicates were run with the random addition option (one addition replicate per bootstrap replicate) selected from the heuristic search menu. Parsimony analyses of each individual gene, and of the combined genetic data set, were conducted with all characters either equally or differentially weighted. In the latter case, characters were weighted according to *a priori* assumptions of character informativeness: transversions only for D-loop, transversions weighted 10 times more than transitions for COII and cytochrome *b*, and third positions only (i.e., Class 1 sites) for IRBP. PAUP* was also employed to generate uncorrected pairwise distance matrices for each gene. These values were considered separately for outgroup

TABLE 1
GenBank Accession Numbers for Study Taxa

Binomial	HV1	COII	cyt b	IRBP
Microcebus murinus	AF081026	AF081039	U53572†	AF081054
Propithecus tattersalli	AF081027	L22782*	U53573†	AF081053
Varecia varieagata varieagata	AF081029	AF081040	AF081047	AF081056
Varecia varieagata rubra	AF081028	L22785*	U53578†	AF081055
Hapalemur griseus	AF081030	L22778*	U53574†	AF081057
Lemur catta	AF081031	L22780*	U53575†	AF081058
Eulemur fulvus collaris	AF081032	AF081041	U53576†	AF081059
Eulemur fulvus rufus	AF081033	AF081042	U53577†	AF081060
Eulemur fulvus albifrons	AF081034	AF081043	AF081048	AF081061
Eulemur macaco macaco	AF081035	L22777*	AF081049	AF081062
Eulemur macaco flavifrons	AF081036	AF081044	AF081050	AF081063
Eulemur rubriventer	AF081038	AF081046	AF081052	AF081065
Eulemur mongoz.	AF081037	AF081045	AF081051	AF081064

Note. Accession nos. marked with an asterisk are from Adkins and Honeycutt (1994) and those marked with a dagger are from Yoder *et al.* (1996a).

to ingroup comparisons (e.g., the distance between *Microcebus* and *E. f. rufus*), for interspecies comparisons (e.g., the distance between *L. catta* and *Hapalemur*), and intersubspecies comparisons (e.g., the distance between *Varecia variegata variegata* and *V. v. rubra*) and are taken to represent a gross estimate of the relative rates of evolution among genetic markers. In other words, the observation that pairwise distances from one marker are higher than those from another is taken as evidence that the former evolves more rapidly than the latter. Maximum likelihood trees were also calculated by heuristic search with PAUP*. For these analyses, estimation of gamma distribution of variable sites and transition / transversion rate ratio (kappa) were allowed; default settings were maintained for all other options, thus yielding the equivalent of the HKY model (Hasegawa *et al.*, 1985). This model is frequently employed as it attempts to correct for multiple substitutions by accounting for differential rates of transitions and transversions and by considering nucleotide frequencies in estimating the likelihood of specific base-pair changes (Swofford *et al.*, 1996, and references therein).

RESULTS AND DISCUSSION

Congruence Among Markers

The questions that motivated this study—Is *Eulemur* monophyletic? Do *L. catta* and *Hapalemur* form a clade? What is the relative position of *Varecia*?—are clearly and consistently resolved in the parsimony analyses of the individual genes (Fig. 1), whether the characters are equally weighted (top row) or differentially weighted (bottom row). Repeatedly, the answers are: *Eulemur* is monophyletic, *L. catta* and *Hapalemur* do form a clade, and *Varecia* is basal to a *Eulemur* plus *L. catta* / *Hapalemur* clade. As indicated by the bootstrap values, support for these results is typically robust. The only consistent exception to these results is seen in the IRBP trees, neither of which resolve the relative placement of the three primary lineages. Nonetheless, it can be appreciated that despite the significantly lower rate of evolution exhibited by this nuclear exon, the majority of primary nodes for the Lemuridae are well supported.

Branching Inconsistencies within Eulemur

Although the overall observed congruence among data sets is gratifying, one still hopes to find a single fully resolved tree that is supported by all of the data. Such is not the case for this study with respect to branching patterns within *Eulemur*. Specifically, the placement of the species *E. mongoz* and *E. rubriventer* varies not only from one data set to another (as in the comparison of Fig. 1a and Fig. 1e) but also within data sets depending on the weighting scheme employed (as in the comparison of Fig. 1a and Fig. 1b). Given the otherwise perfect topological stability of the phylogeny (at least for the mitochondrial data), it might seem curious that the placement of these two taxa would be so consistently inconsistent. In fact, such patterns imply that internal branches within the *Eulemur* radiation are proportionally too short to provide robust resolution. To investigate this idea, we combined the data to maximize the potential for internal branch resolution and performed both a maximum parsimony and a maximum likelihood analysis of the combined data. Figure 2, in which branch lengths are drawn proportionally, illustrates the results. The phylograms for both analyses confirm the suspicion that internal branches among the four *Eulemur* species are proportionally much shorter than other branches. This effect is particularly notable in the maximum likelihood tree (Fig. 2b). Given that we are interested in discovering the one true completely resolved phylogeny, is there anything to be done to overcome the short internal branch problem?

Sampling Effects

A variety of solutions have been proposed for overcoming difficult phylogenetic problems. Many authors, beginning with Felsenstein (1978), have suggested that phylogenetic algorithms that incorporate a model of character-state change can accurately resolve short internal branches (although enhanced performance is usually expected for cases in which parsimony is inconsistent due to long-branch attraction). Others have suggested that increased sampling, of either characters or taxa, can improve accuracy. Thus, we are left with three general approaches to the problem: employing a model of sequence evolution, adding characters, or adding taxa.

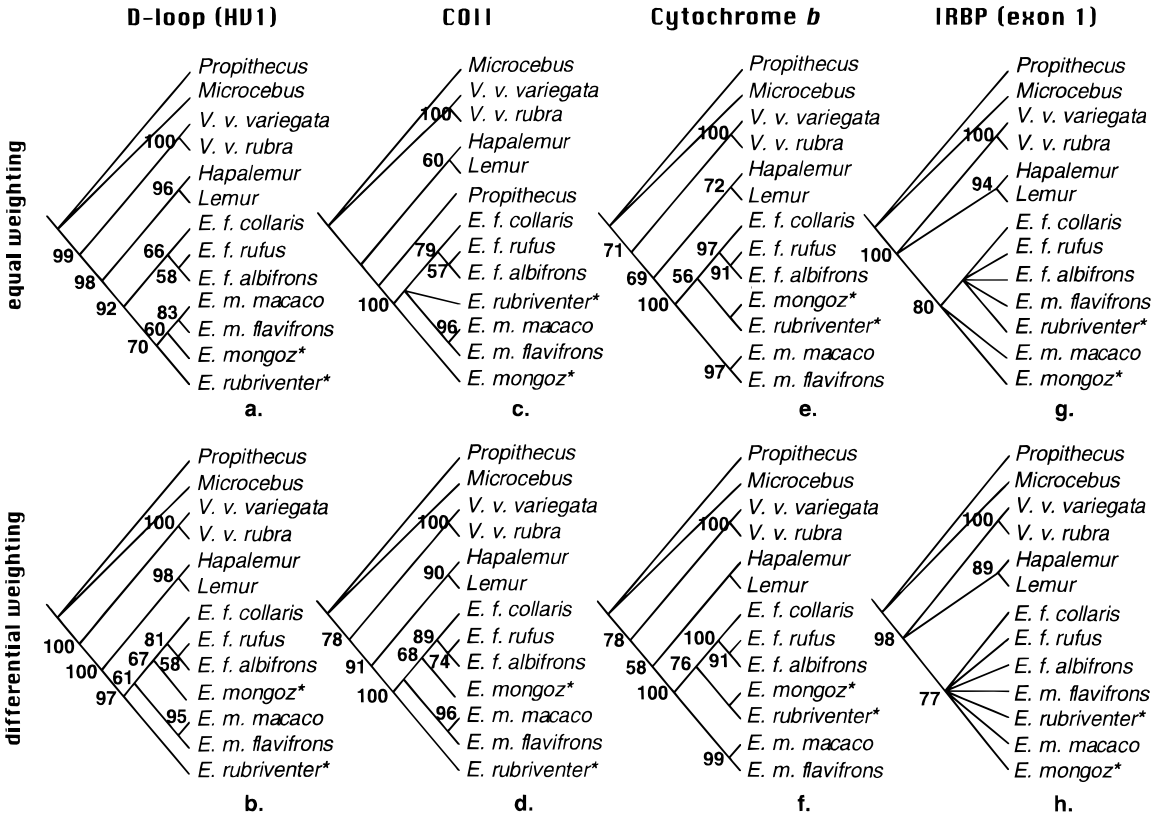


FIG. 1. Maximum parsimony trees for individual genes. Trees in top row derive from analyses in which all characters were equally weighted. Trees in bottom row derive from analyses in which characters were differentially weighted (see Materials and Methods for details of differential weighting). Tree statistics are as follows: (a) 1 tree of length 665, CI = 0.774, RI = 0.631; (b) 1 tree of length 315, CI = 0.819, RI = 0.750; (c) strict consensus of 2 trees of length 492, CI = 0.610, RI = 0.614; (d) 1 tree of length 1237, CI = 0.737, RI = 0.729; (e) 1 tree of length 946, CI = 0.579, RI = .532; (f) 1 tree of length 2629, CI = 0.708; RI = 0.683; (g) strict consensus of 38 trees of length 153, CI = 0.889, RI = 0.790; (h) strict consensus of 156 trees of length 103, CI = 0.864, RI = 0.770. Numbers indicate bootstrap values greater than 50%. *E. mongoz* and *E. rubriventer* are highlighted with asterisk to draw attention to their problematic placement relative to other *Eulemur* taxa.

The hypothesized advantages of the first alternative are presumably not applicable given the results presented in Fig. 2b. Moreover, the results of individual analysis of each gene are discouraging in that, as with parsimony, different resolutions for *Eulemur* interrelationships are recovered for each data set. These results are not illustrated, but are identical to, or are a subset of, the parsimony results. Thus, even with a more sophisticated algorithm and model of sequence evolution, we find the same branching inconsistencies among data sets as observed with parsimony.

A time-honored solution to phylogenetic deficiency is the addition of characters to the analysis. This topic has received much attention in the simulation literature (e.g., Hillis *et al.*, 1994), most recently from investigators concerned with the issue of combined data analysis (Bull *et al.*, 1993; Huelsenbeck *et al.*, 1996). If weak and/or competing resolutions are observed for small data sets, and significant heterogeneity is not detected among them, then the probable explanation is sampling error (i.e., too few characters). Assuming that the phylogenetic method is consistent and appropriate to the question, one can expect the analysis to converge on the correct tree as more data are included (Hillis *et al.*, 1994). Moreover, empirical studies of mitochondrial data indicate that more power is gained by drawing small character sets from numerous genes than by drawing the same absolute number of characters from a single gene or contiguous genetic region (Cao *et al.*, 1994; Cummings *et al.*, 1995).

Accordingly, as with the branch-length test, we combined the four gene-specific data sets into a single data

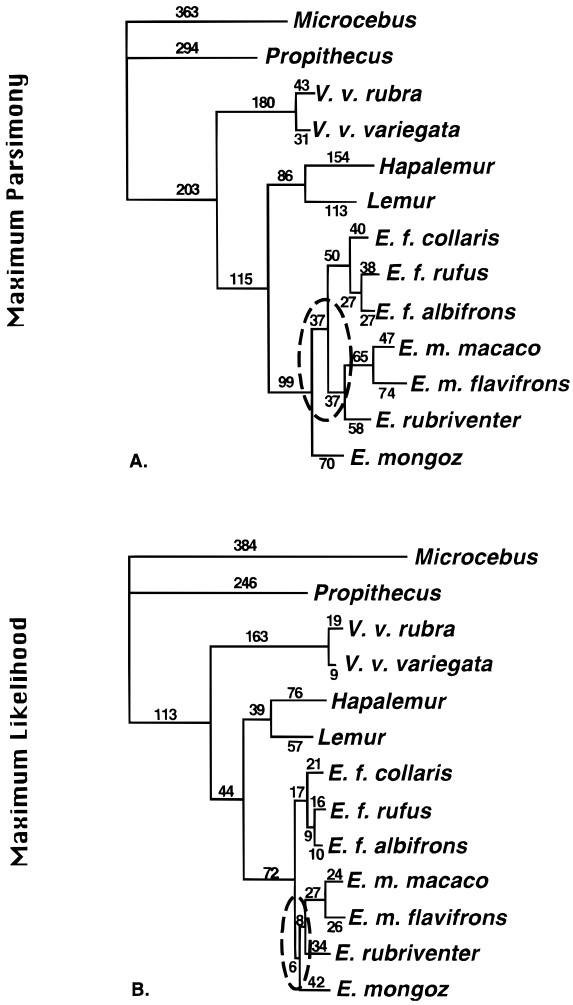


FIG. 2. Phylograms for equally weighted maximum parsimony (a) and maximum likelihood (b) analyses of a combined data set that contains sequences for all four genes (3303 bp total). Assigned branch lengths estimated by PAUP* using the ACCTRAN character optimization are indicated for the parsimony tree. Maximum likelihood expected branch lengths were converted from percentage values to estimated number of substitutions, rounded to the nearest integer. Tree statistics are as follows: (a) 1 tree of length 2251, CI = 0.654, RI = 0.581; (b) -Ln likelihood = 14299.989, estimated kappa = 9.3, estimated gamma shape parameter = 0.205. Dashed ovals illustrate the relatively short internal branches that resolve the placement of *E. rubriventer* and *E. mongoz*.

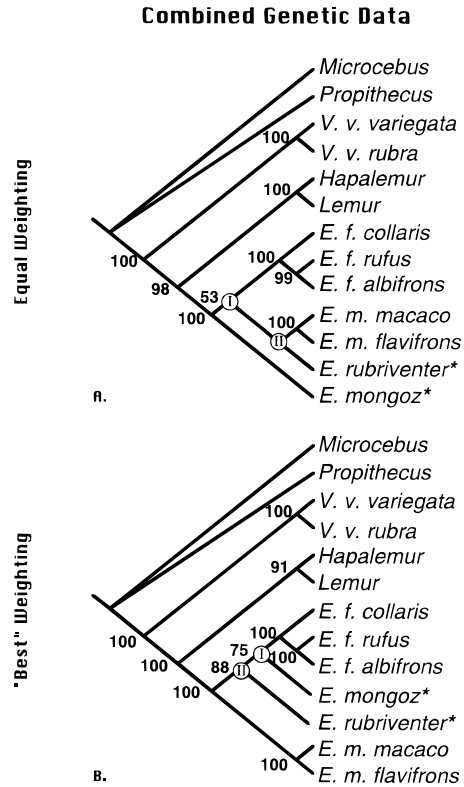


FIG. 3. Parsimony trees for combined mitochondrial and nuclear data set. (a) Analysis in which all characters were equally weighted, tree length = 2251, CI = 0.654, RI = 0.581. (b) Analysis in which characters were differentially weighted (as described under Materials and Methods), tree length = 4266, CI = 0.725, RI = 0.701. Numbers indicate bootstrap values greater than 50%. Nodes indicating the relative placement of *E. mongoz* and *E. rubriventer* are marked with I and II, respectively, in both trees.

set for parsimony analysis. The results are illustrated in Fig. 3. When characters are equally weighted, bootstrap support for virtually every lemurid node is ≥98%—except for those nodes resolving the placement of *E. rubriventer* and *E. mongoz*. On the other hand, when characters are differentially weighted according to *a priori* notions of character consistency and informativeness (i.e., transversions only for the rapidly evolving D-loop data, transversions weighted 10× transitions for the protein-coding mitochondrial data, and Class 1 sites only for the slowly evolving IRBP data), bootstrap values for the placement of these taxa go up to ≥75%—although their relative placement is different than with the equally weighted analysis of the same data set. In fact, both trees show unique solutions to *Eulemur* interrelationships which, when all data sets and weighting schemes are considered, leaves us in the uncomfortable position of choosing among six distinct topologies of *Eulemur* interrelationships.

Before making such a choice, another aspect of analytical design must be considered. The effect of taxon sampling on phylogenetic accuracy has become a topic

TABLE 2
Effects of Reduced Taxon Sampling

D-loop (HV1)	COII	Cytochrome b	IRBP	Combined
Lower resolution	Lower resolution	Different resolution	Lower resolution	Identical
Lower resolution	Different resolution	Different resolution	Different resolution	Identical

Note. "Lower resolution" defined as multiple EP trees whose strict consensus is less resolved than tree(s) derived from complete taxon sample; "different resolution" defined as tree or trees whose strict consensus differs in hierarchical arrangement of nodes; "identical" defined as hierarchical arrangement of relevant nodes identical to tree(s) derived from complete taxon sample. Top row shows resolution with equal weighting; bottom row shows resolution with differential weighting.

of much investigation and discussion (Graybeal, 1998; Hillis, 1996, 1998; Kim, 1996, 1998; Poe, 1998; Soltis *et al.*, 1998). Specifically, Hillis (1996) and others have found that simply by adding taxa to a parsimony analysis, seemingly intractable phylogenetic problems can sometimes become tractable. Nuances of sampling strategy and effect have been debated, but Hillis's (1998) conclusion that "the addition of taxa can have a highly beneficial effect" seems noncontroversial and, indeed, is consistent with a subset of results from the Kim (1996) study. The theoretical foundation for Hillis's assertion rests on the observation that densely sampled phylogenies tend to have proportionally fewer long branches. Given that short internal branches are the likely cause of the *Eulemur* dilemma, we must ask if enhanced taxon sampling could potentially alleviate the problem.

Unfortunately, the answer is "probably not." As mentioned under Materials and Methods, we have sampled all but one of the extant *Eulemur* species. Although it is conceivable that the addition of this single taxon could improve resolution, it is doubtful that the effects would be significant. Nonetheless, we can ask if the results from any or all of the data sets are robust to taxon sampling. In other words, if we were to severely reduce taxon sampling within the Lemuridae, would the hierarchical arrangement of nodes be the same as with the complete taxon sample? To investigate this question, we removed *V. v. rubra, Hapalemur, E. f. albifrons, E. f. rufus, and E. m. flavifrons* from the ingroup taxa and ran parsimony analyses, either with both outgroups or with one removed. Table 2 summarizes the results. Only the combined data set, both equally and differentially weighted, was robust to the reduced taxon sample. This supports the expectation that a larger character sample, particularly one comprising portions of several genes, is better able to consistently resolve the hierarchical order of internal nodes than are the smaller data sets. Nonetheless, even if the combined data set is to be preferred, we are still left with two distinct resolutions of intra-*Eulemur* relationships. If we choose the equally weighted analysis, the conclusion is that *E. mongoz* is the basal taxon (Fig. 3a—node I) with *E. rubriventer* forming a clade with *E. macaco* (Fig. 3a—node II). If, on the other hand, we choose the "best" weighting scheme (described above), then *E. macaco* is the basal taxon (Fig. 3b—node II) with *E. mongoz* forming a clade with *E. fulvus* (Fig. 3b—node I).

The power of differential character weighting in parsimony analysis has been frequently demonstrated (Chippindale and Wiens, 1994; Hillis *et al.*, 1993; Miyamoto *et al.*, 1994; Naylor and Brown, 1997; Yoder *et al.*, 1996b), although arguments for equal weighting are also compelling (Allard and Carpenter, 1996; Kluge, 1997). Thus, it is difficult to choose *a priori* between the two phylogenies illustrated in Fig. 3. Given this quandary, a question occurs to us: How do these trees, derived from a strictly genetic data set, compare with phylogenies derived from either morphological (Groves and Trueman, 1995; Stanger-Hall, 1997) or behavioral characters (Macedonia and Stanger, 1994)? The comparison of the published nongenetic trees reveals considerably more disagreement among them than among the individual gene trees, making it difficult to choose one as definitive. Even so, as with the genetic data, we wish to include all available data for the determination of this problematic phylogeny. Consequently, we combined the nonoverlapping morphological and behavioral characters from the three studies cited above with the combined genetic data analyzed in Fig. 3 (matrix deposited in TreeBASE).

Again, we conducted equally weighted and differentially weighted analyses. (For the latter, genetic characters were weighted as previously described; genetic and morphological characters were otherwise equally weighted.) The hope was that, regardless of weighting scheme, the combined genetic and morphological signal would converge on the same phylogeny, thereby allowing us to unequivocally choose a single phylogeny.

Such was the not the case. For the equally weighted combined genetic, morphological, and behavioral analysis (Fig. 4a), the hierarchical arrangement of nodes is identical to that in Fig. 3a. Significantly, however, bootstrap support for the placement of nodes I and II increases dramatically: from 53 to 75% for node I (these being the values that support the basal positioning of *E. mongoz* in this phylogeny) and from <50 to 93% for node II. For the differentially weighted analysis (Fig. 4b), node hierarchy again remains stable, although bootstrap support in this case either decreases (from 88 to 66% for node II) or does not change significantly (75 to 72% for node I). The results from the addition of morphological and behavioral characters are puzzling in that the individual analysis of the combined morphological and behavioral characters provides yet another distinct arrangement of *Eulemur* taxa (((Efc,Emm),Er),Emon) (See Appendix 1 for taxon identification). Although the bootstrap results suggest that the signal from the nongenetic characters is strongly supportive of the equally weighted phylogeny (Fig. 3a), it also seems that these characters do not significantly conflict with the differentially weighted phylogeny (Fig. 3b).

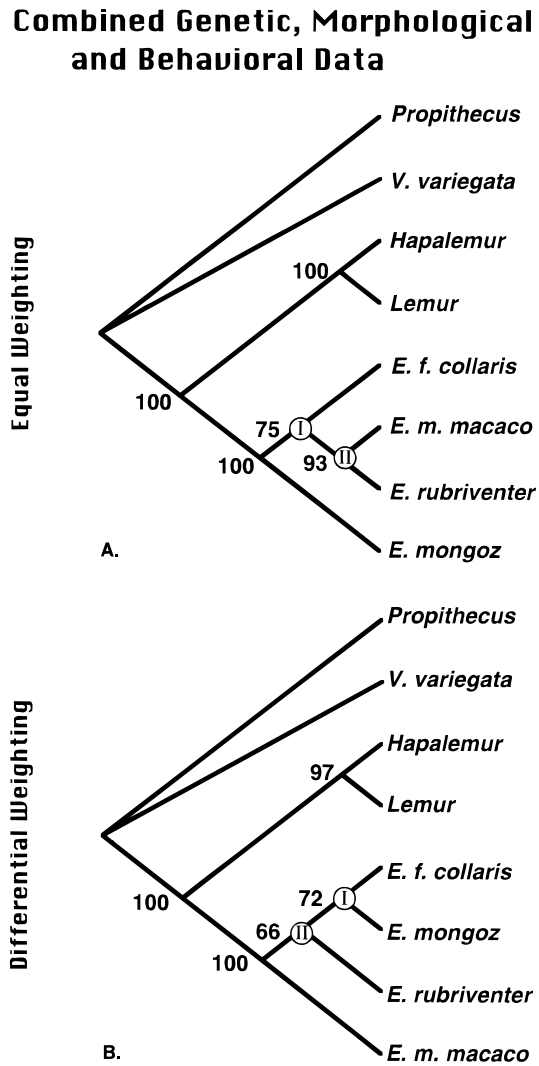
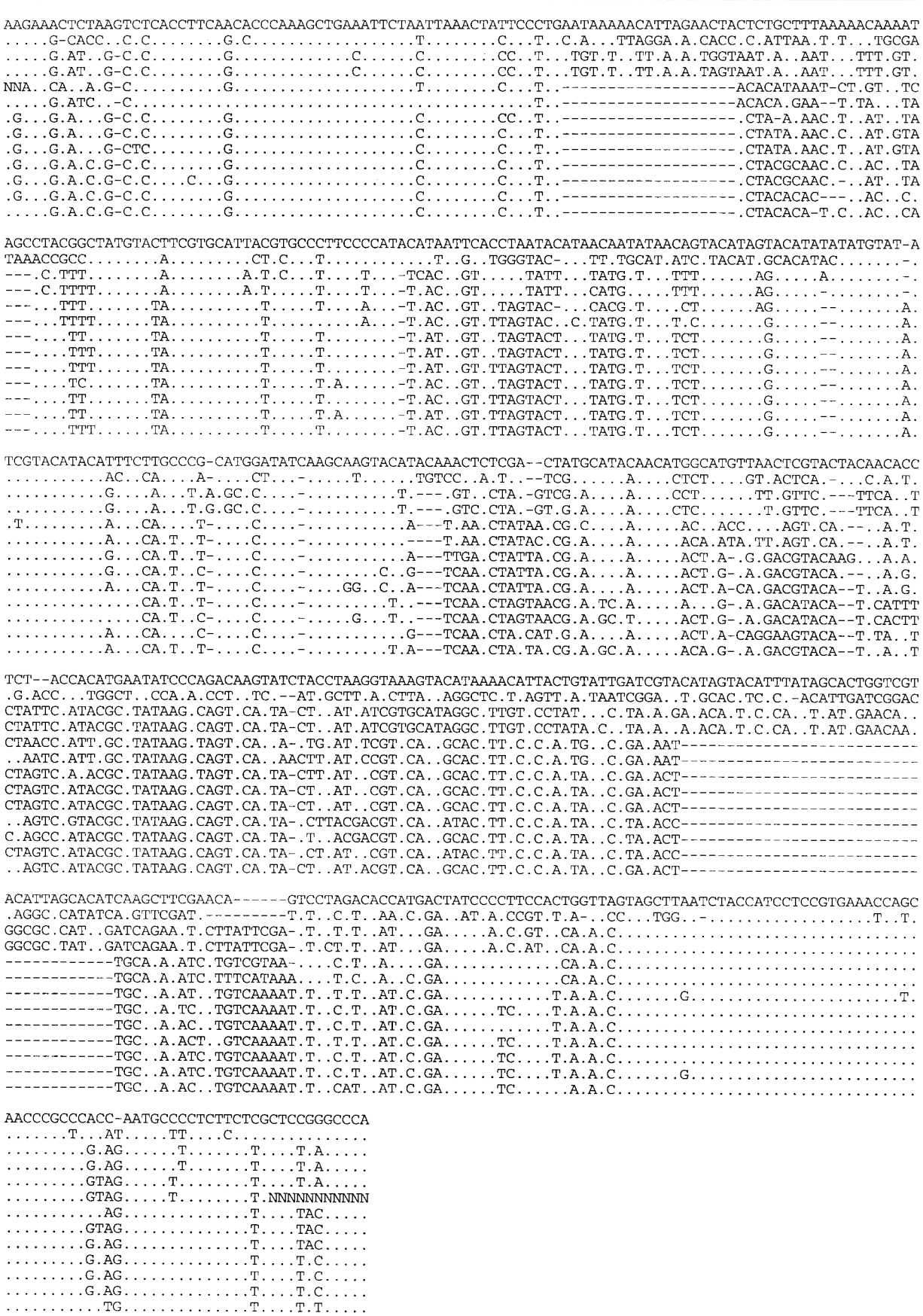


FIG. 4. Maximum parsimony trees of combined genetic data (as in Figs. 3 and 4) along with combined morphological and behavioral data taken from Groves and Trueman (1995), Stanger-Hall (1997), and Macedonia and Stanger (1994). (a) Analysis in which all characters were equally weighted; analysis resulted in 1 tree of length 1772, CI = 0.740, RI = 0.505. (b) Analysis in which genetic characters were differentially weighted (as described under Materials and Methods); morphological and behavioral characters were equally weighted; analysis resulted in 1 tree of length 3079, CI = 0.790, RI = 0.648. Numbers indicate bootstrap values greater than 50%.

CONCLUSIONS

Despite the problematic nature of the species interrelationships within the genus *Eulemur*, all of the questions that originally motivated this study have been answered consistently and with robust support: *Eulemur* describes a clade, *L. catta* and *Hapalemur* form a clade that is sister to *Eulemur*, and *Varecia* is basal to both. The results have been confirmed with mitochondrial and nuclear DNA data, as well as with a select set of morphological and behavioral characters taken from the literature. Given the strength of the results, the diverse character support, and nearly complete taxon sampling, the lemurids are good candidates for the derivation of a stable phylogenetic taxonomy. Currently, the majority of significant nodes are named and define complete clades. The notable exception is that the *Lemur* plus *Hapalemur* node has no designation. We


Mm AAGAAACTCTAAGTCTCACCTTCAACACCCAAAGCTGAAATTCTAATTAAACTATTCCCTGAATAAAAACATTAGAACTACTCTGCTTTAAAAACAAAAT
Pt .....G-CACC..C.C.....G.C.....T.....C...T..C.A...TTAGGA.A.CACC.C.ATTAA.T.T...TGCGA
Vvr .....G.AT..G-C.C.....G.....C.....C.....CC..T...TGT.T..TT.A.A.TGGTAAT.A..AAT...TTT.GT.
Vvv .....G.AT..G-C.C.....G.....C.....C.....CC..T...TGT.T..TT.A.A.TAGTAAT.A..AAT...TTT.GT.
Hg NNA..CA..A.G-C.....G.....T.....C...T..----------ACACATAAAT-CT.GT..TC
Lc .....G.ATC..-C.....T..----------ACACA.GAA--T.TA...TA
Efc .G...G.A...G-C.C.....G.....C.....CC..T..----------.CTA-A.AAC.T..AT..TA
Efr .G...G.A...G-C.C.....G.....C.....C...T..----------.CTATA.AAC.C..AT.GTA
Efa .G...G.A...G-CTC.....G.....C.....C...T..----------.CTATA.AAC.T..AT.GTA
Emm .G...G.A.C.G-C.C.....G.....C.....C...T..----------.CTACGCAAC.C..AC..TA
Emf .G...G.A.C.G-C.C....C...G.....C.....C...T..----------.CTACGCAAC.-..AT..TA
Er .G...G.A.C.G-C.C.....G.....C.....C...T..----------.CTACACAC----.AC..C.
Emon .....G.A.C.G-C.C.....G.....C.....C...T..----------.CTACACA-T.C..AC..CA
Mm AGCCTACGGCTATGTACTTCGTGCATTACGTGCCCTTCCCCATACATAATTCACCTAATACATAACAATATAACAGTACATAGTACATATATATGTAT-A
Pt TAAACCGCC.....A.....CT.C...T.....T..G..TGGGTAC-...TT.TGCAT.ATC.TACAT.GCACATAC.....-.
Vvr ---.C.TTT.....A.....A.T.C...T....T...-TCAC..GT.....TATT...TATG.T...TTT.....AG.....A.....-.
Vvv ---.C.TTTT.....A.....A.T.....T....T...-T.AC..GT.....TATT...CATG.....TTT.....AG.....-.....-.
Hg ---...TTT.....TA.....T.....T....A...-T.AC..GT..TAGTAC-...CACG.T....CT.....AG.....--.....A.
Lc ---...TTTT.....TA.....T.....A...-T.AC..GT.TTAGTAC..C.TATG.T...T.C.....G.....--.....A.
Efc ---....TT.....TA.....T.....T.....-T.AT..GT..TAGTACT...TATG.T...TCT.....G.....--.....A.
Efr ---....TTT.....TA.....T.....T.....-T.AT..GT..TAGTACT...TATG.T...TCT.....G.....--.....A.
Efa ---....TTT.....TA.....T.....T.....-T.AT..GT.TTAGTACT...TATG.T...TCT.....G.....--.....A.
Emm ---....TC.....TA.....T.....T.A.....-T.AC..GT..TAGTACT...TATG.T...TCT.....G.....--.....A.
Emf ---....TT.....TA.....T.....T.....-T.AC..GT.TTAGTACT...TATG.T...TCT.....G.....--.....A.
Er ---....TT.....TA.....T.....T.A.....-T.AT..GT.TTAGTACT...TATG.T...TCT.....G.....--.....A.
Emon ---....TTT.....TA.....T.....T.....-T.AC..GT.TTAGTACT...TATG.T...TCT.....G.....--.....A.
Mm TCGTACATACATTTCTTGCCCG-CATGGATATCAAGCAAGTACATACAAACTCTCGA--CTATGCATACAACATGGCATGTTAACTCGTACTACAACACC
Pt .....AC..CA....A-....CT...-.....T.....TGTCC..A.T..--TCG.....A....CTCT....GT.ACTCA.-...C.A.T.
Vvr .....G....A...T.A.GC.C....-.....T.---.GT..CTA.-GTCG.A....A....CCT.....TT.GTTC.---TTCA..T
Vvv .....G....A...T.G.GC.C....-.....T.---.GTC.CTA.-GT.G.A....A....CTC.....T.GTTC.---TTCA..T
Hg .T.....A...CA....T-....C....-.....A---T.AA.CTATAA.CG.C....A.....AC..ACC....AGT.CA.--..A.T.
Lc .....A...CA.T..T-....C....-.....----T.AA.CTATAC.CG.A....A.....ACA.ATA.TT.AGT.CA.--..A.T.
Efc .....G...CA.T..T-....C....-.....A---TTGA.CTATTA.CG.A....A.....ACT.A-.G.GACGTACAAG...A.A.
Efr .....G...CA.T..C-....C....-.....C..G---TCAA.CTATTA.CG.A....A.....ACT.G-.A.GACGTACA.--..A.G.
Efa .....A...CA.T..T-....C....-....GG..C..A---TCAA.CTATTA.CG.A....A.....ACT.A-CA.GACGTACA--T..A.G.
Emm .....CA.T..T-....C....-.....T..---TCAA.CTAGTAACG.A.TC.A.....A...G-.A.GACATACA--T.CATTT
Emf .....CA.T..C-....C....-.....G...T..---TCAA.CTAGTAACG.A.GC.T.....ACT.G-.A.GACATACA--T.CACTT
Er .....A...CA....C-....C....-.....G---TCAA.CTA.CAT.G.A....A.....ACT.A-CAGGAAGTACA--T.TA..T
Emon .....A...CA.T..T-....C....-.....T.A---TCAA.CTA.TA.CG.A.GC.A.....ACA.G-.A.GACGTACA--T..A..T
Mm TCT--ACCACATGAATATCCCAGACAAGTATCTACCTAAGGTAAAGTACATAAAACATTACTGTATTGATCGTACATAGTACATTTATAGCACTGGTCGT
Pt .G.ACC...TGGCT..CCA.A.CCT..TC.--AT.GCTT.A.CTTA..AGGCTC.T.AGTT.A.TAATCGGA..T.GCAC.TC.C.-ACATTGATCGGAC
Vvr CTATTC.ATACGC.TATAAG.CAGT.CA.TA-CT..AT.ATCGTGCATAGGC.TTGT.CCTAT...C.TA.A.GA.ACA.T.C.CA..T.AT.GAACA..
Vvv CTATTC.ATACGC.TATAAG.CAGT.CA.TA-CT..AT.ATCGTGCATAGGC.TTGT.CCTATA.C..TA.A..A.ACA.T.C.CA..T.AT.GAACAA.
Hg CTAACC.ATT.GC.TATAAG.TAGT.CA..A-.TG.AT.TCGT.CA..GCAC.TT.C.C.A.TG..C.GA.AAT----------
Lc ..AATC.ATT.GC.TATAAG.CAGT.CA..AACTT.AT.CCGT.CA..GCAC.TT.C.C.A.TG..C.GA.AAT----------
Efc CTAGTC.A.ACGC.TATAAG.TAGT.CA.TA-CTT.AT..CGT.CA..GCAC.TT.C.C.A.TA..C.GA.ACT----------
Efr CTAGTC.ATACGC.TATAAG.CAGT.CA.TA-CT..AT..CGT.CA..GCAC.TT.C.C.A.TA..C.GA.ACT----------
Efa CTAGTC.ATACGC.TATAAG.CAGT.CA.TA-CT..AT..CGT.CA..GCAC.TT.C.C.A.TA..C.GA.ACT----------
Emm ..AGTC.GTACGC.TATAAG.CAGT.CA.TA-.CTTACGACGT.CA..ATAC.TT.C.C.A.TA..C.TA.ACC----------
Emf C.AGCC.ATACGC.TATAAG.CAGT.CA.TA-.T..ACGACGT.CA..GCAC.TT.C.C.A.TA..C.TA.ACT----------
Er CTAGTC.ATACGC.TATAAG.CAGT.CA.TA-.CT.AT..CGT.CA..ATAC.TT.C.C.A.TA..C.TA.ACC----------
Emon ..AGTC.ATACGC.TATAAG.CAGT.CA.TA-CT..AT.ACGT.CA..GCAC.TT.C.C.A.TA..C.GA.ACT----------
Mm ACATTAGCACATCAAGCTTCGAACA------GTCCTAGACACCATGACTATCCCCTTCCACTGGTTAGTAGCTTAATCTACCATCCTCCGTGAAACCAGC
Pt .AGGC.CATATCA.GTTCGAT.---------T.T..C.T..AA.C.GA..AT.A.CCGT.T.A-..CC...TGG..-.....T..T.
Vvr GGCGC.CAT..GATCAGAA.T.CTTATTCGA-.T..T.T..AT...GA.....A.C.GT..CA.A.C.....
Vvv GGCGC.TAT..GATCAGAA.T.CTTATTCGA-.T.CT.T..AT...GA.....A.C.AT..CA.A.C.....
Hg ----------TGCA.A.ATC.TGTCGTAA-....C.T..A....GA.....CA.A.C.....
Lc ----------TGCA.A.ATC.TTTCATAAA....T.C..A..C.GA.....CA.A.C.....
Efc ----------TGC..A.AT..TGTCAAAAT.T..T.T..AT.C.GA.....T.A.A.C.....G.....T.
Efr ----------TGC..A.TC..TGTCAAAAT.T..C.T..AT.C.GA.....TC....T.A.A.C.....
Efa ----------TGC..A.AC..TGTCAAAAT.T..C.T..AT.C.GA.....T.A.A.C.....
Emm ----------TGC..A.ACT..GTCAAAAT.T..T.T..AT.C.GA.....TC....T.A.A.C.....
Emf ----------TGC..A.ATC.TGTCAAAAT.T..C.T..AT.C.GA.....TC....T.A.A.C.....
Er ----------TGC..A.ATC.TGTCAAAAT.T..C.T..AT.C.GA.....TC....T.A.A.C.....G.....
Emon ----------TGC..A.AC..TGTCAAAAT.T..CAT..AT.C.GA.....TC.....A.A.C.....
Mm AACCCGCCCACC-AATGCCCCTCTTCTCGCTCCGGGCCCA
Pt .....T...AT.....TT....C.....
Vvr .....G.AG.....T.....T....T.A.....
Vvv .....G.AG.....T.....T....T.A.....
Hg .....GTAG.....T.....T....T.A.....
Lc .....GTAG.....T.....T.NNNNNNNNNNN
Efc .....AG.....T....TAC.....
Efr .....GTAG.....T....TAC.....
Efa .....G.AG.....T....TAC.....
Emm .....G.AG.....T....T.C.....
Emf .....G.AG.....T....T.C.....
Er .....G.AG.....T....T.C.....
Emon .....TG.....T....T.T.....

suggest that the results of this study demonstrate the reality of this clade and that it should be appropriately named.

Although the average reader of this journal will have little to no acute interest in the exact sequence of cladogenic events within the genus *Eulemur*, it nonetheless seems that our struggle to resolve this problem has general implications. Mammalian phylogeny is rife with problematic areas relating to short internal branches, many of which cannot be addressed with the addition of taxa (for the simple reason that additional taxa do not exist). Thus, the only alternative lies in the addition of characters.

APPENDIX 1

The 540-bp Alignment of D-Loop (HVl) Sequences

See preceding page. Taxon identifiers: Mm, *Microcebus murinus*; Pt, *Propithecus tattersalli*; Vvv, *Varecia varieagata varieagata*; Vvr, *Varecia varieagata rubra*; Hg, *Hapalemur griseus*; Lc, *Lemur catta*; Efc, *Eulemur fulvus collaris*; Efr, *Eulemur fulvus rufus*; Efa, *Eulemur fulvus albifrons*; Emm, *Eulemur macaco macaco*; Emf, *Eulemur macaco flavifrons*; Er, *Eulemur rubriventer*; Em, *Eulemur mongoz*. *Microcebus* (Mm) sequence serves as reference with dot (·) indicating base identical to reference sequence at that site. Dashes (-) indicate position of inferred indels. "N" indicates missing data.

ACKNOWLEDGMENTS

This project was supported by grants from NSF and the Leakey Foundation to A.D.Y. We thank Marc Allard for inviting this submission and for his helpful editorial recommendations. The manuscript was improved thanks to comments from two anonymous reviewers. This is DUPC Publication No. 674.

REFERENCES

Adkins, R. M., and Honeycutt, R. L. (1994). Evolution of the primate cytochrome *c* oxidase subunit II gene. *J. Mol. Evol.* **38,** 215–231.

Allard, M. W., and Carpenter, J. M. (1996). On weighting and congruence. *Cladistics* **12,** 183–198.

Bryant, H. N. (1996). Explicitness, stability, and universality in the phylogenetic definition and usage of taxon names: A case study of the phylogenetic taxonomy of the Carnivora (Mammalia). *Syst. Biol.* **45,** 174–189.

Bull, J. J., Huelsenbeck, J. P., Cunningham, C. W., Swofford, D. L., and Waddell, P. J. (1993). Partitioning and combining data in phylogenetic analysis. *Syst. Biol.* **42,** 384–397.

Cantino, P. D., Olmstead, R. G., and Wagstaff, S. J. (1997). A comparisons of phylogenetic nomenclature with the current system: A botanical case study. *Syst. Biol.* **46,** 313–331.

Cao, Y., Adachi, J., Janke, A., Paabo, S., and Hasegawa, M. (1994). Phylogenetic relationships among eutherian orders estimated from inferred sequences of mitochondrial proteins: Instability of a tree based on a single gene. *J. Mol. Evol.* **39,** 519–527.

Chippindale, P. T., and Wiens, J. J. (1994). Weighting, partitioning, and combining characters in phylogenetic analysis. *Syst. Biol.* **43,** 278–287.

Crovella, S., Montagnon, D., and Rumpler, Y. (1993). Highly repeated DNA analysis and systematics of the Lemuridae, a family of Malagasy prosimians. *Primates* **34,** 61–69.

Crovella, S., Montagnon, D., and Rumpler, Y. (1995). Highly repeated DNA sequences and systematics of Malagasy primates. *Human Evol.* **10,** 35–44.

Crovella, S., and Rumpler, Y. (1992). Confirmation of the specific status of Hapalemur simus (Primates, Strepsirrhini) by restriction genomic DNA banding patterns. *Human Evol.* **7,** 63–67.

Cummings, M. P., Otto, S. P., and Wakeley, J. (1995). Sampling properties of DNA sequence data in phylogenetic analysis. *Mol. Biol. Evol.* **12,** 814–822.

de Queiroz, K., and Gauthier, J. (1990). Phylogeny as a central principle in taxonomy: Phylogenetic definitions of taxon names. *Syst. Zool.* **39,** 307–322.

de Queiroz, K., and Gauthier, J. (1992). Phylogenetic taxonomy. *Annu. Rev. Ecol. Syst.* **23,** 449–480.

de Queiroz, K., and Gauthier, J. (1994). Toward a phylogenetic system of biological nomenclature. *Trends Ecol. Evol.* **9,** 27–31.

Eaglen, R. H. (1980). "The Systematics of Living Strepsirhini, with Special Reference to the Lemuridae." Unpublished Ph.D. dissertation. Duke University.

Felsenstein, J. (1978). Cases in which parsimony or compatibility methods will be positively misleading. *Syst. Zool.* **27,** 401–410.

Graybeal, A. (1998). Is it better to add taxa or characters to a difficult phylogenetic problem? *Syst. Biol.* **47,** 9–17.

Groves, C. P., and Eaglen, R. H. (1988). Systematics of the Lemuridae (Primates, Strepsirhini). *J. Human Evol.* **17,** 513–538.

Groves, C. P., and Trueman, J. W. H. (1995). Lemurid systematics revisited. *J. Human Evol.* **28,** 427–437.

Hasegawa, M., Kishino, H., and Yano, T. (1985). Dating of the human-ape splitting by a molecular clock of mitochondrial DNA. *J. Mol. Evol.* **22,** 160–174.

Higgins, D. G., and Sharp, P. M. (1988). CLUSTAL: A package for performing multiple sequence alignment on a microcomputer. *Gene* **73,** 237–244.

Hillis, D. M. (1996). Inferring complex phylogenies. *Nature* **383,** 130–131.

Hillis, D. M. (1998). Taxonomic sampling, phylogenetic accuracy, and investigator bias. *Syst. Biol.* **47,** 3–8.

Hillis, D. M., Allard, M. W., and Miyamoto, M. M. (1993). Analysis of DNA sequence data: Phylogenetic inference. *Methods Enzymol.* **224,** 456–487.

Hillis, D. M., Huelsenbeck, J. P., and Cunningham, C. W. (1994). Application and accuracy of molecular phylogenies. *Science* **264,** 671–677.

Huelsenbeck, J. P., Bull, J. J., and Cunningham, C. W. (1996). Combining data in phylogenetic analysis. *Trends Ecol. Evol.* **11,** 152–158.

Jung, K. Y., Crovella, S., and Rumpler, Y. (1992). Phylogenetic relationships among lemuriform species determined from restriction genomic DNA banding patterns. *Folia Primatol.* **58,** 224–229.

Kim, J. (1996). General inconsistency conditions for maximum parsimony: Effects of branch lengths and increasing numbers of taxa. *Syst. Biol.* **45,** 363–374.

Kim, J. (1998). Large-scale phylogenies and measuring performance of phylogenetic estimators. *Syst. Biol.* **47,** 43–60.

Kluge, A. G. (1997). Sophisticated falsification and research cycles: Consequences for differential character weighting in phylogenetic systematics. *Zool. Scripta* **26,** 349–360.

Lee, M. S. Y. (1998). Ancestors and taxonomy. *Trends Ecol. Evol.* **13:**26.

Macedonia, J. M., and Stanger, K. F. (1994). Phylogeny of the Lemuridae revisited: Evidence from communication signals. *Folia Primatol.* **63,** 1–43.

Mittermeier, R. A., Tattersall, I., Konstant, W. R., Meyers, D. M., and Mast, R. B. (1994). "Lemurs of Madagascar." Conservation International, Washington, DC.

Miyamoto, M. M., Allard, M. W., Adkins, R. M., Janecek, L. L., and Honeycutt, R. L. (1994). A congruence test of reliability using linked mitochondrial DNA sequences. *Syst. Biol.* **43,** 236–249.

Naylor, G. J. P., and Brown, W. M. (1997). Structural biology and phylogenetic estimation. *Nature* **388,** 527–528.

Poe, S. (1998). Sensitivity of phylogeny estimation to taxonomic sampling. *Syst. Biol.* **47,** 18–31.

Randria, G. F. N. R. (1998). "Morphologie, Analyses Phenetique et Cladistique des Lemuridae (Gray, 1821)." Ph.D. Université d'Antananarivo.

Schander, C., and Thollesson, M. (1995). Phylogenetic taxonomy—Some comments. *Zool. Scripta* **24,** 263–267.

Simons, E. L., and Rumpler Y. (1988). Eulemur: New generic name for species of Lemur other than Lemur catta. *C.R. Acad. Sci. Paris* **307,** 547–551.

Soltis, D. E., Soltis, P. S., Mort, M. E., Chase, M. W., Savolainen, V., Hoot, S. B., and Morton, C. M. (1998). Inferring complex phylogenies using parsimony: An empirical approach using three large DNA data sets for angiosperms. *Syst. Biol.* **47,** 32–42.

Stanger-Hall, K., and Cunningham, C. W. (1998). Support for a monophyletic lemuriformes: overcoming incongruence between data partitions. *Mol. Biol. Evol.* **15,** 1572–1577.

Stanger-Hall, K. F. (1997). Phylogenetic affinities among the extant Malagasy lemurs (Lemuriformes) based on morphology and behavior. *J. Mamm. Evol.* **4,** 63–194.

Swofford, D. L. (1998). "PAUP*. Phylogenetic Analysis Using Parsimony (*and Other Methods)." Sinauer Associates, Sunderland, MA.

Swofford, D. L., Olsen, G. J., Waddel, P. J., and Hillis, D. M. (1996). Phylogenetic inference. *In* "Molecular Systematics" (D. M. Hillis, C. Moritz, and B. K. Mable, Eds.), pp. 407–514. Sinauer Associates, Sunderland, MA.

Tattersall, I. (1988). A note on the nomenclature in Lemuridae. *Phys. Anthropol News* 7:14.

Tattersall, I. (1993). "Speciation and Morphological Differentiation in the Genus *Lemur* Species, Species Concepts, and Primate Evolution" (W. H. Kimbel and L. B. Martin, Eds.). Plenum Press, New York.

Tattersall, I., and Koopman, K. (1989). A further note on nomenclature in Lemuridae. *J. Human Evol.* **18,** 499–500.

Tattersall, I., and Schwartz, J. (1991). Phylogeny and nomenclature in the "*Lemur*-group" of Malagasy strepsirhine primates. *Anthr. Papers Am. Mus. Nat. Hist.* **69,** 3–18.

Thompson, J. D., Higgins, D. G., and Gibson,T. J. (1994). CLUSTAL W: Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* **22,** 4673–4680.

Wyss, A. R., and Meng, J. (1996). Application of phylogenetic taxonomy to poorly resolved crown clades: A stem-modified node-based definition of Rodentia. *Syst. Biol.* **45,** 559–568.

Yoder, A. D. (1994). Molecular systematics of the Lemuridae: Is *Eulemur* necessary? *Am. J. Phy. Anthr.* **18,** 212–213.

Yoder, A. D., Cartmill, M., Ruvolo, M., Smith, K., and Vilgalys, R. (1996a). Ancient single origin of Malagasy primates. *Proc. Nat. Acad. Sci.* **93,** 5122–5126.

Yoder, A. D., Ruvolo, M., and Vilgalys, R. (1996b). Molecular evolutionary dynamics of cytochrome *b* in strepsirrhine primates: The phylogenetic significance of third position transversions. *Mol. Biol. Evol.* **13,** 1339–1350.