Development and application of a phylogenomic toolkit: Resolving the evolutionary history of Madagascar's lemurs

Julie E. Horvath,[1,7] David W. Weisrock,[2] Stephanie L. Embry,[1] Isabella Fiorentino,[2] James P. Balhoff,[3] Peter Kappeler,[4] Gregory A. Wray,[1,2] Huntington F. Willard,[1,2,5] and Anne D. Yoder[1,2,6]

[1]Institute for Genome Sciences & Policy, Duke University, Durham, North Carolina 27708, USA; [2]Department of Biology, Duke University, Durham, North Carolina 27708, USA; [3]National Evolutionary Synthesis Center, Durham, North Carolina 27705, USA; [4]Department of Behavioural Ecology and Sociobiology, German Primate Centre, Göttingen 37077, Germany; [5]Department of Molecular Genetics, Duke University Medical Center, Durham, North Carolina 27710, USA; [6]Duke Lemur Center, Duke University, Durham, North Carolina 27705, USA

Lemurs and the other strepsirrhine primates are of great interest to the primate genomics community due to their phylogenetic placement as the sister lineage to all other primates. Previous attempts to resolve the phylogeny of lemurs employed limited mitochondrial or small nuclear data sets, with many relationships poorly supported or entirely unresolved. We used genomic resources to develop 11 novel markers from nine chromosomes, representing ~9 kb of nuclear sequence data. In combination with previously published nuclear and mitochondrial loci, this yields a data set of more than 16 kb and adds ~275 kb of DNA sequence to current databases. Our phylogenetic analyses confirm hypotheses of lemuriform monophyly and provide robust resolution of the phylogenetic relationships among the five lemuriform families. We verify that the genus *Daubentonia* is the sister lineage to all other lemurs. The Cheirogaleidae and Lepilemuridae are sister taxa and together form the sister lineage to the Indriidae; this clade is the sister lineage to the Lemuridae. Divergence time estimates indicate that lemurs are an ancient group, with their initial diversification occurring around the Cretaceous-Tertiary boundary. Given the power of this data set to resolve branches in a notoriously problematic area of primate phylogeny, we anticipate that our phylogenomic toolkit will be of value to other studies of primate phylogeny and diversification. Moreover, the methods applied will be broadly applicable to other taxonomic groups where phylogenetic relationships have been notoriously difficult to resolve.

[Supplemental material is available online at www.genome.org. The sequence data from this study have been submitted to GenBank under accession nos. EU057196–EU057514 and EU342218–EU342345.]

In recent years, the increasing amount of genomic sequence data available for major evolutionary lineages has been paralleled by an increase in the use of multilocus phylogenetic approaches to resolve difficult and previously intractable evolutionary relationships (Rokas et al. 2003; Murphy et al. 2004; James et al. 2006; Weisrock et al. 2006). Some of these studies have been based exclusively on publicly available databases, mining DNA sequence information to yield large data sets, a number of which contain hundreds of thousands of orthologous nucleotide positions (Bapteste et al. 2002; Rokas et al. 2003; Pollard et al. 2006; Cannarozzi et al. 2007; Nikolaev et al. 2007). Still others have used comparative genomic patterns of insertions and deletions to reconstruct relationships, owing in part to the nonhomoplastic and conserved nature of such events (Kriegs et al. 2006; Murphy et al. 2007). Ideally, phylogenetic studies will sample broadly across a wide range of independently segregating loci distributed throughout the genome. Such a strategy conforms to a fundamental assumption of phylogenetic analysis, that characters will be identically and independently distributed. Nonetheless, up to the present, the influx of multilocus data sets has been driven largely by studies using only a very limited number of loci, usually utilizing conserved PCR primers for orthologous nuclear loci that were previously developed in other taxonomic groups (e.g., Malcomber 2002; Townsend et al. 2004; Poux et al. 2005; Weisrock et al. 2005; Galewski et al. 2006; Harlin-Cognato and Honeycutt 2006; James et al. 2006; Noonan and Chippindale 2006; Zhang et al. 2006; Heckman et al. 2007; Li and Orti 2007).

Increasing the number of nuclear loci for phylogenetic analysis is more readily accomplished in some clades than others. For example, clades that contain model organisms with substantial genomic resources allow for the rapid assembly of large and broadly distributed phylogenomic loci (e.g., Rokas et al. 2003). This process is more problematic for nonmodel organisms. To date, the typical approach for increasing the number of amplifiable and informative nuclear sequence markers is to compare distantly related species with available genomic resources to develop conserved PCR primers for use in a more inclusive set of taxa (Palumbi and Baker 1994). This approach for nuclear marker generation has been successful for studies within diverse groups, such as placental mammals (Murphy et al. 2001) and fish (Li et al. 2007), and successful when subsequently applied to more exclusive terminal groups, for example, bats (Teeling et al. 2005) and cats (Johnson et al. 2006).

A major drawback to this approach is that increased evolu-

tionary divergence between genomes used in primer development will likely result in an increased PCR failure rate, leading to incomplete data matrices in phylogenetic analyses. For example, Murphy et al. (2001) developed a set of 15 nuclear DNA primer pairs based on human–mouse genomic comparisons, for use in placental mammalian phylogenetics. Of the 66 taxa used in Murphy et al. (2001), an average of six taxa were missing from each data set, with some matrices missing sequence data from as many as 14 taxa. Notably, the phylogenetically-mysterious primate genus *Tarsius* (Yoder 2003) is represented in only 10 of the Murphy et al. (2001) nuclear data sets and contains only 55% of the total aligned nuclear sequence data. Given the importance of taxon sampling in phylogenetic reconstruction (Zwickl and Hillis 2002; Hedtke et al. 2006), minimizing missing data should be a central objective of phylogenomic primer design methods. The analysis of data sets that are compromised by weak sampling, whether it be taxonomic or character insufficiencies, often results in phylogenetic hypotheses that either are conflicting or are weakly resolved.

Previous character-based phylogenetic analyses of lemurs (Primates, suborder Lemuriformes) have been subject to weak taxonomic sampling and small data sets that are based on only one or a few genetic loci (Crovella et al. 1993, 1995; Adkins and Honeycutt 1994; Yoder 1994, 2003; DelPero et al. 1995, 2001; Dutrillaux and Rumpler 1995; Porter et al. 1995, 1997b; Stanger 1996; Stanger-Hall 1997; Stanger-Hall and Cunningham 1998; Yoder and Irwin 1999; Murphy et al. 2001; Pastorini et al. 2003; Poux et al. 2005). Despite these weaknesses, virtually all phylogenetic studies of lemurs and related primates agree on three points: (1) lemurs are monophyletic; (2) the lemuriform clade is most closely related to the African and Asian lorisiform clade, together forming the tooth-combed primate clade strepsirrhini; and (3) the strepsirrhine clade is sister to all other living primates. Moreover, the majority of these studies have found that the aye-aye, genus *Daubentonia* (family Daubentoniidae), is sister to a clade composed of all other lemurs, including the recently extinct giant lemurs (Karanth et al. 2005). Beyond this result, however, there has been complete disagreement among studies as to the relationships among the other four major evolutionary lineages, also classified as families. Virtually every conceivable phylogenetic resolution has been presented among the Cheirogaleidae, Indriidae, Lemuridae, and Lepilemuridae (containing the single genus *Lepilemur*) in studies where taxon sampling has been adequate to investigate their relationships.

The accurate resolution of lemuriform phylogeny is critical for reconstructing primate evolutionary history and, by extension, for providing a comparative framework for understanding human biology. Lemurs constitute more than 15% of extant primate species diversity, and their position in the primate evolutionary tree offers unique power for hypothesizing the timing and phylogenetic placement of primate-specific traits. For example, the current project to sequence the genome of the gray mouse lemur, *Microcebus murinus* (http://www.genome.gov/10002154), will allow genomic researchers to identify genetic traits that are diagnostic of the primates, distinct from those of all other mammals. Given the independent evolution of lemurs, a fully resolved lemur phylogeny can be used as the foundation for quantitative comparative analyses, thus allowing for more powerful tests of evolutionary hypotheses across all primates (Nunn and Barton 2001). Moreover, our study is timely in addressing the biological community's current interest in both the methods (Telford 2007) and the organisms (Yoder 2007).

In order to develop this robust phylogeny, we have capitalized on available lemur genome sequence to create a molecular toolkit for use in phylogenetic comparisons across all strepsirrhine taxa targeted in our study. This phylogenomic toolkit includes PCR primer pairs designed across the genome and corresponding genomic sequences from all strepsirrhine primates in our study. Our toolkit resolves long-standing questions of lemur phylogeny and is used to calculate divergence time estimates for lemuriform evolutionary history. Importantly, this primer design approach should be useful for application to other diverse groups of species for which limited genomic sequence exists.

Results

Our study includes 29 species that span strepsirrhine evolutionary diversity (Table 1). The 11 new loci total 9 kb and include exons, introns, and noncoding regions and provide a largely balanced representation of loci across the genome (Supplemental Table 1). An additional 7.3 kb of nuclear DNA and mitochondrial DNA (mtDNA) sequence data was assembled from GenBank or generated de novo with published nuclear PCR primers (Supplemental Table 1). These published sequences represent seven nuclear loci (both intronic and exonic) and two mtDNA genes (MT-CO2 and MT-CYB). Combined, these loci yield 16,363 bp of aligned data, containing 5413 variable and 3950 parsimony informative characters. When mtDNA sequence data are excluded, the combined nuclear character matrix contains 14,539 bp of aligned sequence data, of which 4424 characters are variable and 3089 are parsimony informative. For a full description of character variation by locus, see Table 2 and Supplemental Table 2.

Phylogenetic analysis of the combined nuclear and nuclear plus mtDNA data sets yielded similar topologies across all optimality criteria. Branch support for several nodes diminished when the nuclear data were combined with the mtDNA data. Given the high degree of homoplasy seen in previous mtDNA-based reconstructions of lemur phylogeny (Yoder et al. 1996b), we focus on the topology favored by the combined nuclear data set as our best approximation of lemuriform phylogeny. The results of the Bayesian analysis are presented in Figure 1 and are consistent with maximum parsimony (MP) and maximum likelihood (ML) analyses. Branch support values for all combined analyses are provided in Table 3. The topology is well supported with MP and ML bootstrap values >90 and Bayesian posterior probabilities (PPs) equal to 1.0 for 22 of 25 numbered nodes. This result supports the hypothesis that lemurs are monophyletic, with the deepest split within the clade between the family Daubentoniidae and all remaining lemurs. The Cheirogaleidae and Lepilemuridae are supported as sister lineages. Moreover, the Indriidae (represented here by the three species contained within genus *Propithecus*) is resolved as the sister lineage to the Cheirogaleidae plus Lepilemuridae clade. As with most previous analyses, species of the family Lemuridae are resolved as a well-supported monophyletic group (Yoder and Irwin 1999; Pastorini et al. 2003; Roos et al. 2004).

Tree topologies for most individual loci were not substantially different from the total evidence trees generated from the concatenated nuclear data. Even so, branch support in the single-locus trees was often weak, leaving some uncertainty about how informative the individual loci were (data not shown; for discussion, see Methods). Separate analyses of exonic, intronic, or noncoding partitions yielded decreased support for some branches relative to the total evidence trees, but the overall topologies

Table 1. Taxon information

Strepsirrhine family	Latin name	Common name	Duke Lemur Center ID	Sex	DNA source
Cheirogaleidae	*Cheirogaleus medius*[a]	Fat-tailed dwarf lemur	687	Female	Buccal WGA
Cheirogaleidae	*Cheirogaleus medius*[a]	Fat-tailed dwarf lemur	1653	Female	Buccal WGA
Cheirogaleidae	*Microcebus berthae*	Berthe's mouse lemur	NA	Male	Tissue WGA
Cheirogaleidae	*Microcebus murinus*[a]	Gray mouse lemur	7006	Male	Buccal WGA
Cheirogaleidae	*Microcebus murinus*[a]	Gray mouse lemur	7021	Male	Buccal WGA
Cheirogaleidae	*Microcebus ravelobensis*	Golden brown mouse lemur	NA	Male	Tissue WGA
Cheirogaleidae	*Mirza coquereli*	Coquerel's mouse lemur	2307	Male	Buccal WGA
Daubentoniidae	*Daubentonia madagascariensis*	Aye-aye	6262	Female	Tissue WGA
Galagidae	*Galago moholi*	Southern lesser bushbaby	3188	Male	Buccal WGA
Galagidae	*Otolemur garnettii garnettii*	Small-eared galago	8030	Male	Tissue WGA
Indriidae	*Propithecus d. diadema*	Diademed sifaka	6564	Male	Blood DNA
Indriidae	*Propithecus tattersalli*	Golden-crowned sifaka	6196	Male	Buccal WGA
Indriidae	*Propithecus verreauxi coquereli*	Coquerel's sifaka	6583	Male	Buccal WGA
Lemuridae	*Eulemur coronatus*	Crowned lemur	6478	Male	Buccal WGA
Lemuridae	*Eulemur fulvus albifrons*	White-fronted lemur	5547	Female	Blood DNA
Lemuridae	*Eulemur fulvus collaris*	Collared lemur	5776	Female	Buccal WGA
Lemuridae	*Eulemur fulvus fulvus*	Brown lemur	6123	Female	Buccal WGA
Lemuridae	*Eulemur fulvus rufus*	Red-fronted lemur	6459	Female	Buccal WGA
Lemuridae	*Eulemur fulvus sanfordi*	Sanford's lemur	6386	Female	Buccal WGA
Lemuridae	*Eulemur macaco flavifrons*	Blue-eyed black lemur	6288	Male	Buccal WGA
Lemuridae	*Eulemur m. macaco*	Black lemur	5859	Female	Coriell/IBPIR cell line PR00254
Lemuridae	*Eulemur mongoz*	Mongoose lemur	6077	Male	Buccal WGA
Lemuridae	*Eulemur rubriventer*	Red-bellied lemur	6038	Female	Blood DNA
Lemuridae	*Hapalemur g. griseus*	Eastern lesser bamboo lemur	1370	Female	Buccal WGA
Lemuridae	*Lemur catta*	Ring-tailed lemur	NA	Female	Coriell cell line AG07100
Lemuridae	*Varecia variegata rubra*	Red ruffed lemur	6633	Female	Buccal WGA
Lemuridae	*Varecia v. variegata*[a]	Black and white ruffed lemur	6178	Male	Tissue WGA
Lemuridae	*Varecia v. variegata*[a]	Black and white ruffed lemur	4598	Male	Buccal WGA
Lepilemuridae	*Lepilemur ruficaudatus*	Red-tailed sportive lemur	NA	Female	Tissue WGA
Lorisidae	*Loris tardigradus*	Slender loris	2914	Female	Buccal WGA
Lorisidae	*Nycticebus coucang*	Slow loris	1960	Male	Buccal WGA
Lorisidae	*Nycticebus pygmaeus*	Lesser slow loris	1925	Male	Tissue WGA

[a]More than one animal of this species was used in analyses.
WGA, whole-genome amplification; NA, not available.

were not substantially altered (see Supplemental Fig. 1A–C; Supplemental material). Similarly, analyses excluding single loci ("leave-one-out" analysis) from the concatenated set did not yield substantially different branch support from the overall total evidence topology (Supplemental material; Supplemental Table 3).

Concordance across data sets

A Bayesian analysis of concordance across 13 of our nuclear loci produced a primary concordance tree with a topology completely congruent with our total evidence trees (Fig. 2; Supplemental Table 4). Identical results were obtained when using α priors of 0.1, 1, and 10. A key result of these analyses was the high level of concordance across loci for all branches. All but two branches in the primary concordance tree have clade concordance factors (CCFs) >0.89, with 95% confidence intervals including a CCF of 1 (i.e., these nodes are supported by all 13 assessed loci). The two nodes with slightly lower CCFs are the placement of *Propithecus* (family Indriidae) in a clade with the Cheirogaleidae and Lepilemuridae (Fig. 1, node 8), and the placement of *Eulemur* as sister to the *Lemur* plus *Hapalemur* clade (Fig. 1, node 18). Our analyses still yield relatively high CCFs for these nodes (>0.761), in favor of the primary concordance tree. However, these analyses do identify three loci (*ABCA1*, *CFTR*-pair A, and *RBP3*) that appear to strongly contradict these latter two relationships, with joint posterior densities that favor the placement of the Indriidae as the sister lineage to the Lemuridae and favor the sister relationship of the genera *Eulemur* and *Varecia* (Supplemental Table 4). Further

details regarding concordance among loci can be found in the Supplemental materials.

Lemur divergence estimates

Divergence times were estimated using the combined nuclear data and two calibrations: the Human–Chimpanzee divergence at 5–7 million years ago (Mya) (Kumar et al. 2005), and the Galagidae–Lorisidae divergence at 38–42 Mya (Fig. 3; Table 4) (Seiffert et al. 2003). Divergence between the Lemuriformes and Lorisiformes is estimated at 75 Mya (with a 95% credibility interval [CI] of 66.9–84.4 Mya). The mean age estimate for the earliest lemuriform divergence between *Daubentonia* and all other lemurs is 66 Mya, with a CI (54.9–74.7 Mya) that overlaps the 65 Mya Cretaceous/Tertiary (K/T) boundary. These age estimates are congruent with previous estimates of strepsirrhine and lemuriform antiquity (Roos et al. 2004; Yoder and Yang 2004) (Supplemental Fig. 2). Similarly, these analyses agree with previous findings suggesting an Eocene/Oligocene timing for diversification of the majority of family-level lemur lineages (excluding Daubentoniidae).

Discussion

The power of our data set for resolving the evolutionary history of the lemuriformes demonstrates the utility of phylogenetic marker development utilizing genomic resources from a focal set of taxa. In our study, the use of BAC library sequence data from two relatively divergent species (*Lemur* and *Microcebus*) led to the

Table 2. Data partition information

Locus	Size (bp)	Variable sites[a]	Parsimony informative sites[a]	Model[b]
ABCA1	684	300/248	196/149	GTR+Γ
AXIN1	949	234/203	178/145	GTR+I+Γ
CFTR-Pair A	865	326/266	224/167	GTR+Γ
CFTR-Pair B	849	268/213	178/134	GTR+Γ
ERC2	767	261/211	179/138	GTR+Γ
LRPPRC-Pair A	924	255/206	175/129	GTR+Γ
LRPPRC-Pair B	813	221/180	139/102	GTR+Γ
LUC7L	697	207/161	145/98	GTR+Γ
Noncoding	871	143/104	93/60	HKY+Γ
SREBF2	764	244/200	187/148	HKY+Γ
ZNF202	828	184/135	134/92	GTR+Γ
ADORA	383	111/93	88/64	GTR+Γ
FGA	716	298/246	196/145	HKY+Γ
RBP3	1238	359/285	267/193	HKY+Γ
SLC11A1	531	173/133	110/72	GTR+Γ
RAG1	825	185/161	144/117	GTR+I+Γ
TTR	965	288/179	185/86	GTR+Γ
VWF	906	367/289	271/196	GTR+Γ
MT-CO2	684	369/322	324/270	HKY+I+Γ
MT-CYB	1140	620/570	537/483	GTR+I+Γ
Combined nuclear	14539	4424/3513	3089/2235	#
Combined all loci	16363	5413/4405	3950/2988	#
Nuclear exonic	4702	1182/961	896/677	#
Noncoding	9575	3194/2518	2156/1530	#
Nuclear Intronic	8602	3036/2400	2058/1465	#

[a]Ingroup + Outgroup/Ingroup.
[b]Best-fit model as determined using Akaike Information Criterion (AIC) using MrModeltest2.2. GTR, general time-reversible model; HKY, the Hasegawa, Kishino, and Yano model; Γ, rate heterogeneity; I, the proportion of sites assumed to be invariant. #, For combined Bayesian analyses, data were partitioned using the above models assessed for individual loci. The "combined nuclear," "combined all," and exonic maximum likelihood analyses were set up using the GTR+I+Γ model, while intronic and noncoding maximum likelihood analyses were analyzed using a GTR+Γ model.

development of PCR primer pairs with exceptionally high efficacy across strepsirrhine taxa. In contrast, the use of previously published mammalian primers yielded data sets featuring incomplete taxon sampling (see Supplemental Table 5), some of which are unable to contribute to the phylogenetic resolution of key lemuriform family-level lineages. Overall, the combination of 11 newly developed loci combined with seven previously published nuclear loci, provides the most comprehensive strepsirrhine phylogenetic analysis to date, with more than 400 kb compared among 29 strepsirrhine and three anthropoid primate taxa. The unequivocal support of lemur monophyly and the position of Daubentoniidae as the most basal lineage is not surprising given past studies (Yoder et al. 1996a, 2003; Porter et al. 1997a; Stanger-Hall and Cunningham 1998; Pastorini et al. 2003; Poux and Douzery 2004; Roos et al. 2004; Yoder and Yang 2004; Poux et al. 2005; DelPero et al. 2006). However, the fact that many of these past phylogenies did not sufficiently resolve particular family relationships (or did so in the absence of convincing branch support) (Yoder and Yang 2004; Poux et al. 2005) raised the issue of whether these data sets were merely too small or whether these phylogenetic relationships represented a hard polytomy (i.e., rapid speciation event into three or more lineages essentially simultaneously). Our phylogenetic resolution of the major lemur lineages with high levels of concordance across loci is strong evidence against a polytomous history among the major lemur lineages and instead argues for a deep bifurcating history. It also

accentuates the point that difficult phylogenetic problems are best addressed via the application of large data sets drawn from across the breadth of the genome.

Implications of loci concordance

Both the Bayesian concordance and "leave-one-out" analyses indicate that most loci contribute phylogenetic signal to the overall topology and that most loci are in agreement (Fig. 2; Supplemental Tables 3, 4; Methods). The Bayesian concordance analyses did, however, uncover an interesting pattern in which three loci were consistent in yielding joint posterior densities that favor an alternative hypothesis that shifts the Indriidae to a sister position with the Lemuridae and places *Eulemur* and *Varecia* as sister lineages (Supplemental Table 4). The cause for this discordance may be related to the short evolutionary durations preceding nodes 8 and 18. As the durations between branching events decrease, there is an increased chance that gene trees will disagree due to an incomplete lineage sorting process (Hudson 1992). While this incongruence is more often thought to affect genealogical discordance at the population-species boundary, it has also been shown to have substantial effects at deeper levels of evolutionary divergence (Takahashi et al. 2001).

Short evolutionary durations between deep branching events are also expected to challenge phylogenetic resolution due to a limited set of synapomorphic substitutions and the loss of signal due to multiple substitutions. This challenge is well recognized in phylogenetics (Felsenstein 1978; Huelsenbeck 1998; Fishbein et al. 2001; Weisrock et al. 2005) and has been highlighted in previous attempts at resolving lemur phylogeny with mtDNA sequence data (Adkins and Honeycutt 1994; Yoder et al. 1996b) in which lemur monophyly is only recovered when using third position transversions in protein-coding genes. These previous results are the likely explanation for the decreased branch support in our analyses of the combined nuclear and mtDNA data. Furthermore, it should not be too surprising to find alternative phylogenetic hypotheses for some taxa in a small number of individual nuclear gene trees. Overall, the vast majority of loci used in this study resolve concordant relationships for nearly all branches in our tree, including the placement of the Indriidae, Cheirogaleidae, and Lepilemuridae into a clade. These consistent phylogenetic patterns across numerous independent loci are indicative of an evolutionary history that is recoverable, despite relatively short durations between deep branching events.

A biogeographical context for lemur divergence time estimates

This resolved lemur phylogeny provides an opportunity to explore basal primate origins, the historical biogeography of Madagascar (Yoder et al. 1996a), and primate speciation events (Horvath and Willard 2007) and should contribute to conservation efforts (Ryder 2005; Lehman 2006). Divergence age analyses calculated here and elsewhere (Yoder and Yang 2004; Poux et al. 2005) estimated that lemurs did not arrive in Madagascar until 50–80 Mya. Given that the time of lemuriform divergence clearly postdates the separation of Madagascar from other landmasses by many millions of years, it seems clear that lemurs arrived via oceanic dispersal (Kappeler 2000; Yoder et al. 1996a, 2003) though other biogeographic mechanisms have been proposed (McCall 1997; Arnason et al. 2000). Once the ancestral lemurs
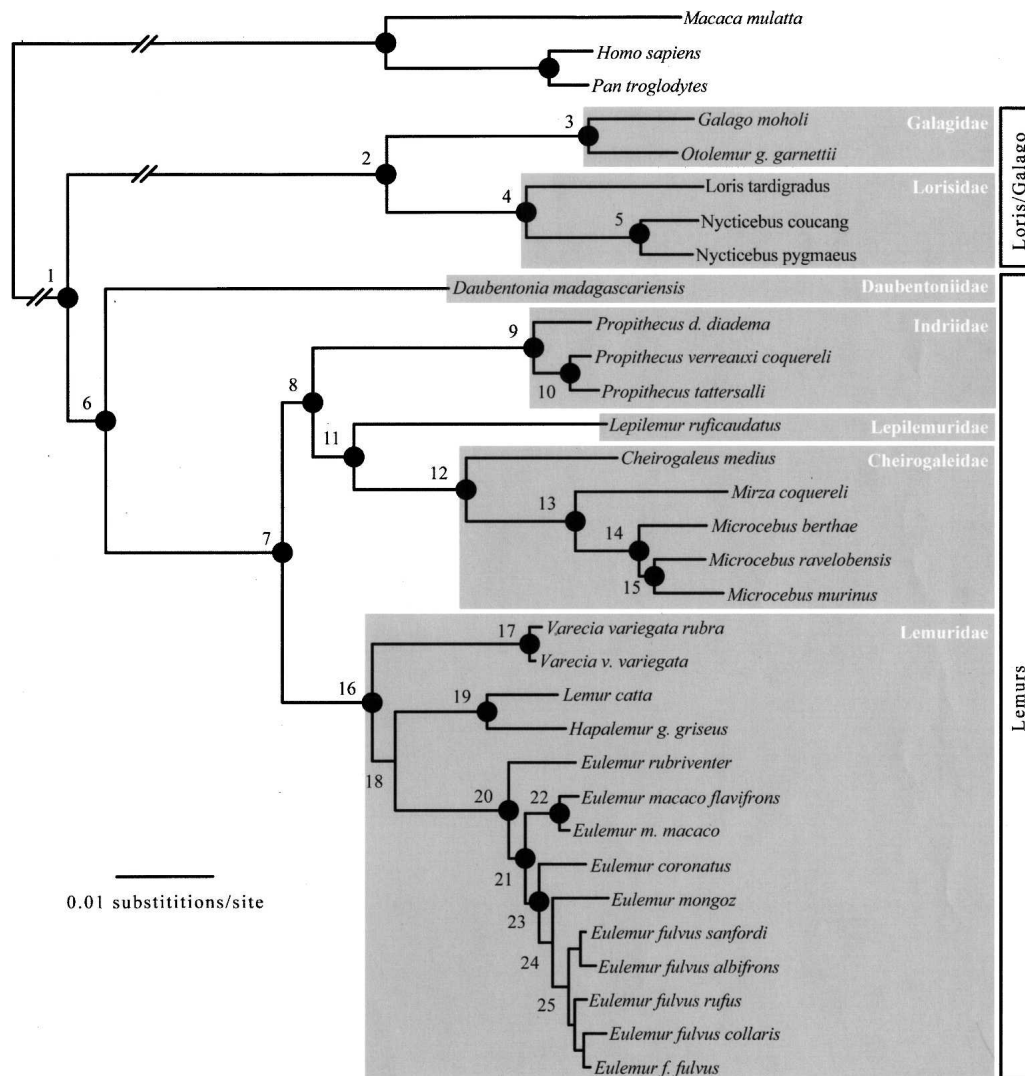
Figure 1. Lemur Bayesian consensus phylogram using 18 nuclear loci combined. This phylogram is based on the combined posterior distribution of four independent analyses of the combined and partitioned nuclear data sets. The topology is identical to those resolved using maximum parsimony (MP) and maximum likelihood (ML). Nodes with strong measures of branch support (posterior probabilities = 1.0, MP bootstrap >90%, ML >90%) are marked with filled black circles. Branch support values are indicated in Table 3. Open vertical boxes span the two infraorders of strepsirrhine primates. Shaded boxes encompass family-level taxonomic groups.

arrived on Madagascar, there appear to have been at least two prominent episodes of lineage diversification, at the family and species levels. Following the deepest divergence between *Daubentonia* and the common ancestor of all remaining lemurs, the four remaining family-level lineages diverged within a window of 16 million years (Myr), between the end of the Eocene and through the Oligocene (23.8–40 Mya) (see Fig. 3). Coincident with this timeframe, rainforests began emerging (Wells 2003; Yoder and Yang 2004) and many other mammals such as carnivores, tenrecs, and rodents may have colonized Madagascar (Poux et al. 2005). As a working hypothesis, one can imagine that the combination of the changing climate combined with the arrival of new predators and competitors would have altered the ecological landscape and thus served as an impetus for lineage diversification within the lemuriform clade.

The divergence time estimates for the origin of diversification in the genera *Eulemur* and *Microcebus*, two of the most diverse lemur genera, are both hypothesized to have emerged in the late Miocene (~8 Mya). The past 8 Myr on Madagascar initially began with a wetter climate from monsoons (Wells 2003) and, if congruent with global patterns, would have experienced plant biome changes (Cerling et al. 1997), some of which may have been in response to Pleistocene glaciation cycles (Wilme et al. 2006). Again, though cause and effect are hypothetical, such environmental changes may have stimulated speciation within the *Eulemur* and *Microcebus* lineages.

The value of phylogenomic toolkits

Despite the large investment required to develop phylogenomic toolkits such as the one presented here, the analytical payoff can be enormous. Although our taxonomic focus was strepsirrhine phylogeny, the primer design approach used in this study can be applied to any organismal group, even those with limited

Table 3. Branch support values

Branch	Bayesian posterior probability					Maximum likelihood bootstrap					Parsimony bootstrap/decay index				
	nDNA	mtDNA +nDNA	nDNA Exon	nDNA Intron	nDNA Non-coding	nDNA	mtDNA +nDNA	nDNA Exon	nDNA Intron	nDNA Non-coding	nDNA	mtDNA +nDNA	nDNA Exon	nDNA Intron	nDNA Non-coding
1	1.0	1.0	1.0	1.0	1.0	100	100	100	100	100	100/670	100/773	100/197	100/438	100/466
2	1.0	1.0	1.0	1.0	1.0	100	100	100	100	100	100/356	100/369	100/100	100/242	100/252
3	1.0	1.0	1.0	1.0	1.0	100	100	100	100	100	100/155	100/168	100/45	100/106	100/108
4	1.0	1.0	1.0	1.0	1.0	100	100	100	100	100	100/94	100/94	100/29	100/60	100/65
5	1.0	1.0	1.0	1.0	1.0	100	100	100	100	100	100/88	100/114	100/28	100/50	100/58
6	1.0	1.0	1.0	1.0	1.0	99	90	91	96	97	100/35	94/20	98/15	99/19	99/20
7	1.0	1.0	0.98	1.0	1.0	100	100	100	100	100	100/149	100/36	100/95	100/100
8	1.0	1.0	1.0	1.0	1.0	100	99	76	90	97	94/10	<50/2	77/3	70/2	75/3
9	1.0	1.0	1.0	1.0	1.0	100	100	100	100	100	100/147	100/189	100/45	100/87	100/99
10	1.0	1.0	0.89	1.0	1.0	100	100	64	97	96	100/22	100/56	73/2	87/6	93/8
11	1.0	1.0	1.0	1.0	1.0	100	100	93	98	100	100/19	94/16	95/7	98/11	98/11
12	1.0	1.0	1.0	1.0	1.0	100	100	100	100	100	100/63	100/74	100/19	100/30	100/31
13	1.0	1.0	1.0	1.0	1.0	100	100	100	100	100	100/73	100/85	100/24	100/64	100/66
14	1.0	1.0	1.0	1.0	1.0	100	100	100	100	100	100/57	100/91	99/10	100/45	100/46
15	1.0	1.0	NR	1.0	1.0	97	60	NR	95	97	99/10	55/1	NR/NR	98/7	98/7
16	1.0	1.0	1.0	1.0	1.0	100	100	100	100	100	100/82	100/86	100/24	100/48	100/52
17	1.0	1.0	1.0	1.0	1.0	100	100	100	100	100	100/153	100/235	100/51	100/100	100/102
18	1.0	1.0	1.0	1.0	1.0	91	100	81	81	90	85/8	96/16	88/4	66/2	74/3
19	1.0	1.0	1.0	1.0	1.0	100	100	100	100	100	100/84	100/100	100/15	100/67	100/69
20	1.0	1.0	1.0	1.0	1.0	100	100	100	100	100	100/86	100/127	100/24	100/58	100/60
21	1.0	1.0	1.0	0.98	1.0	90	60	100	67	68	99/10	70/4	92/4	88/4	90/5
22	1.0	1.0	1.0	1.0	1.0	100	100	91	100	100	100/28	100/35	83/3	100/22	100/23
23	1.0	1.0	NR	1.0	1.0	97	77	77	96	99	93/6	<50/1	NR	95/6	93/5
24	1.0	0.99	0.87	0.99	1.0	76	74	NR	71	82	81/5	<50/3	NR	63/2	83/3
25	1.0	1.0	1.0	NR	NR	89	100	79	NR	NR	89/5	100/9	76/3	NR	53/NR

NR, not resolved; mtDNA, mitochondrial DNA; nDNA, nuclear DNA.

amounts of available genomic sequence data. Our resulting lemur-based PCR primers proved to be applicable to divergent strepsirrhine taxa separated by as much as 84 Myr of evolutionary divergence (Fig. 3; Table 4). Consequently, toolkit design can span great divergences but will be more easily applied to closely related species. The development of conserved primer pairs that work in a diverse set of species is obviously geared toward phylogenetic analysis at and above the species level, but may also prove useful in analyses at the intraspecific level, either in genealogical-based analyses (Rosenberg and Nordborg 2002), or as the basis for single nucleotide polymorphism development (Brumfield et al. 2003). Consequently, intraspecific application of our toolkit may provide useful estimates of genetic diversity within populations and be useful for directing conservation strategies (Ryder 2005; Lehman 2006). Our hope is that these PCR primers will become a standard molecular resource utilized by lemur researchers with access to species not represented here, thus offering a means for filling in missing taxa and enhancing the overall lemur phylogeny. While we have extensive species sampling within some genera (e.g., *Eulemur*), other genera are missing entirely (*Phaner, Allocebus, Indri, Avahi*) and will be important to include in future studies to verify their placement with respect to those currently assessed. As these missing taxa are well established as being included within the major family-level lineages of focus in this study, we do not expect the tree topology to change substantially.

The ultimate goal of our approach was to design a method applicable to any clade-specific phylogenetic problem wherein at least one species offers genomic sequence. This strategy was tested in the evolutionarily diverse lemurs and was successful for resolving their evolutionary history. While we anticipate that our phylogenomic toolkit (the 11 newly designed primer pairs and corresponding strepsirrhine genomic sequence) will be of use to other primatologists for genomic comparisons and for resolution of the evolutionary history of their study taxa, its application extends much further. This primer design approach can be applied to any organism with available genomic sequence for phylogenetic analyses, diversity assessments, or other molecular evolutionary comparisons across diverse taxa. Primer design in other nonmodel organisms will facilitate the exploration of more complete comparative genomics, and will lead to a better understanding of the tree of life.

Methods

Taxon sampling

Sequence data were collected from 29 strepsirrhine taxa representing all five lemuriform families and both lorisiform families (Galagidae and Lorisidae) (Table 1). Sequence data from three outgroup taxa with complete genome sequence were also used (*Homo sapiens, Pan troglodytes, Macaca mulatta*). The majority of ingroup samples were obtained from the Duke Lemur Center (http://lemur.duke.edu/). Samples of *Microcebus ravelobensis, Microcebus berthae*, and *Lepilemur ruficaudatus* are from the personal collection of the authors (A.D. Yoder and P. Kappeler). The sample for *Lemur catta* (cell line AG07100) was obtained from Coriell Cell Repositories (http://locus.umdnj.edu/). The sample for *Eulemur m. macaco* (cell line PR00254) was obtained from Integrated Primate Biomaterials and Information Resource (http://ccr.coriell.org/Sections/Collections/IPBIR/?SsId=18).
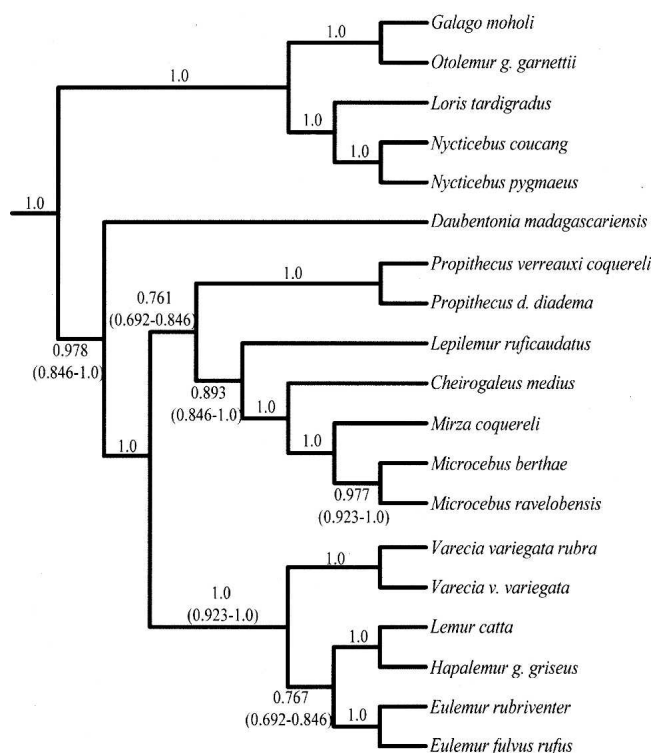
Figure 2. Bayesian primary concordance tree. The Bayesian primary concordance tree is shown, resulting from analyses of 13 nuclear loci, using the program BUCKy (Larget 2006). Numbers on branches represent the average concordance factor. The 95% confidence intervals that are not exclusively 1.0 are included in parentheses. Results are identical across analyses using α priors of 0.1, 1, and 10. Loci not used in this analysis are listed in the Supplementary material.

Genomic DNA sample collection

Genomic DNA was isolated from buccal cells collected on brushes (PUREGENE kit, Qiagen), from blood, and from cell line material. Genomic DNA was isolated using the PUREGENE kit and the manufacturer's recommended protocol (Qiagen). Approximately 10–100 ng genomic DNA (1/20 volume buccal or 1/80 volume blood DNA isolated) was used in whole-genome amplification (WGA) reactions (REPLI-g [Qiagen] or GenomiPhi [GE Healthcare]) according to the manufacturer's recommendations. Each set of WGA reactions was set up with a negative control to ensure no cross-contamination of genomic DNA samples occurred during the WGA procedure.

Primer design

We chose two approaches to design PCR primer pairs to obtain loci conserved in a diverse set of strepsirrhine taxa. First, we reasoned that loci conserved between human and a nonprimate outgroup, such as the dog, would likely also be conserved in most primate taxa and could be used to develop loci for deep phylogenetic comparisons (i.e., presently the standard approach to primer design). Initial primer design using human and dog sequence targeted six genomic regions that were assessed for conservation in both species using the conservation track in the UCSC genome browser (http://genome.ucsc.edu/). By use of the conservation track alignments, 17 primer pairs were designed by hand to be 100% identical between human and dog, 20–24 bp in length and 45%–55% GC rich, and to terminate with a G or C nucleotide at the 3' end. Two primers were designed with degen-

erate bases at two sites not identical between human and dog. No primers were designed within, or spanning, known repetitive elements.

Our second approach used lemur and human sequence comparisons since sequence conserved between one lemur species and human would likely be conserved among other strepsirrhine taxa. At the onset of this project, as is typically true for nonmodel organisms, no partially or fully sequenced genome existed for any strepsirrhine primate. Therefore, we relied upon publicly available bacterial artificial chromosome (BAC) sequences from two lemur species (the ringtailed lemur, *L. catta,* and the gray mouse lemur, *M. murinus*). We downloaded three ringtailed lemur BAC sequences (accession nos. AC123971, AC145463, and AC145484) from GenBank (http://www. ncbi.nlm.nih.gov/) and masked them for known primate repeat elements (RepeatMasker 2004, A. Smit; http://www. repeatmasker.org). The repeat masked BAC sequences were compared with the human genome (hg18, March 2006) by BLAST (blastall v2.2.10, http://www.ncbi.nlm.nih.gov/BLAST/), and 16 primer pairs were designed by hand using the above parameters.

The human and lemur sequence comparisons were much more successful than human and dog and therefore were used for further primer design with a semi-automated program, primer_conservation.py. Using primer_conservation.py, 48 *L. catta* and 42 *M. murinus*, repeat masked lemur BACs were compared with the human genome by BLAST. Coordinates of 100% identity were submitted to Primer3 for primer design (v 0.3.0) (Rozen and Skaletsky 2000; source code available at http://fokker.wi.mit.edu/ primer3/). Parameters for primer design were as follows: 600–1000 bp product size, 22–26 bp primer length, 45%–65% GC, 1-bp GC clamp at the 3' end, and primer pairs that did not span, or contain, a known repeat element. The output primers were assessed in the human genome to ensure no duplicated regions were utilized for further analyses. By use of this program with human and lemur sequence comparisons, an additional 31 primer pairs were designed in 16 of the 90 newly downloaded BAC sequences. The remaining BAC sequences did not return primer pairs with our parameters and were not utilized further.

Combined, these approaches yielded a total of 64 primer pairs that were assessed for success in PCR assays on a panel of four taxa chosen for diverse placement in the phylogenetic tree (*H. sapiens, Daubentonia madagascariensis, L. catta,* and *Otolemur g. garnettii*). A successful primer pair was one that amplified a single band in PCR assays in all four taxa. Of the 17 human–dog primer pairs, only one (6%) was successful, while 14 of 16 (88%) human–lemur hand-designed primer pairs and 15 of 31 (48%) human–lemur program-designed primer pairs were successful in the four-taxon panel. The human–dog genomic regions chosen for primer design were different from the human–lemur regions and could have contributed to the lower level of primer pair success.

Next, the 29 successful human–lemur–based primers were assessed in PCR assays using an expanded panel of 10 primates (*H. sapiens, Cheirogaleus medius, D. madagascariensis, Eulemur f. fulvus, L. catta, M. murinus, Nycticebus coucang, O. g. garnettii, Propithecus tattersalli,* and *Varecia v. variegata*). This resulted in 16 of 29 (55%) primer pairs that successfully amplified a product in all 10 primate taxa. The primer pairs with the largest, most robust amplicons (using *L. catta* accession nos. AC145533, AC146284, and AC139880 in addition to *M. murinus* accession nos. AC145757 and AC186699) were selected, and 11 primer pairs successfully amplified all 32 primates in our panel (see Supplemental Table 5). Primer pairs that were not successful in the
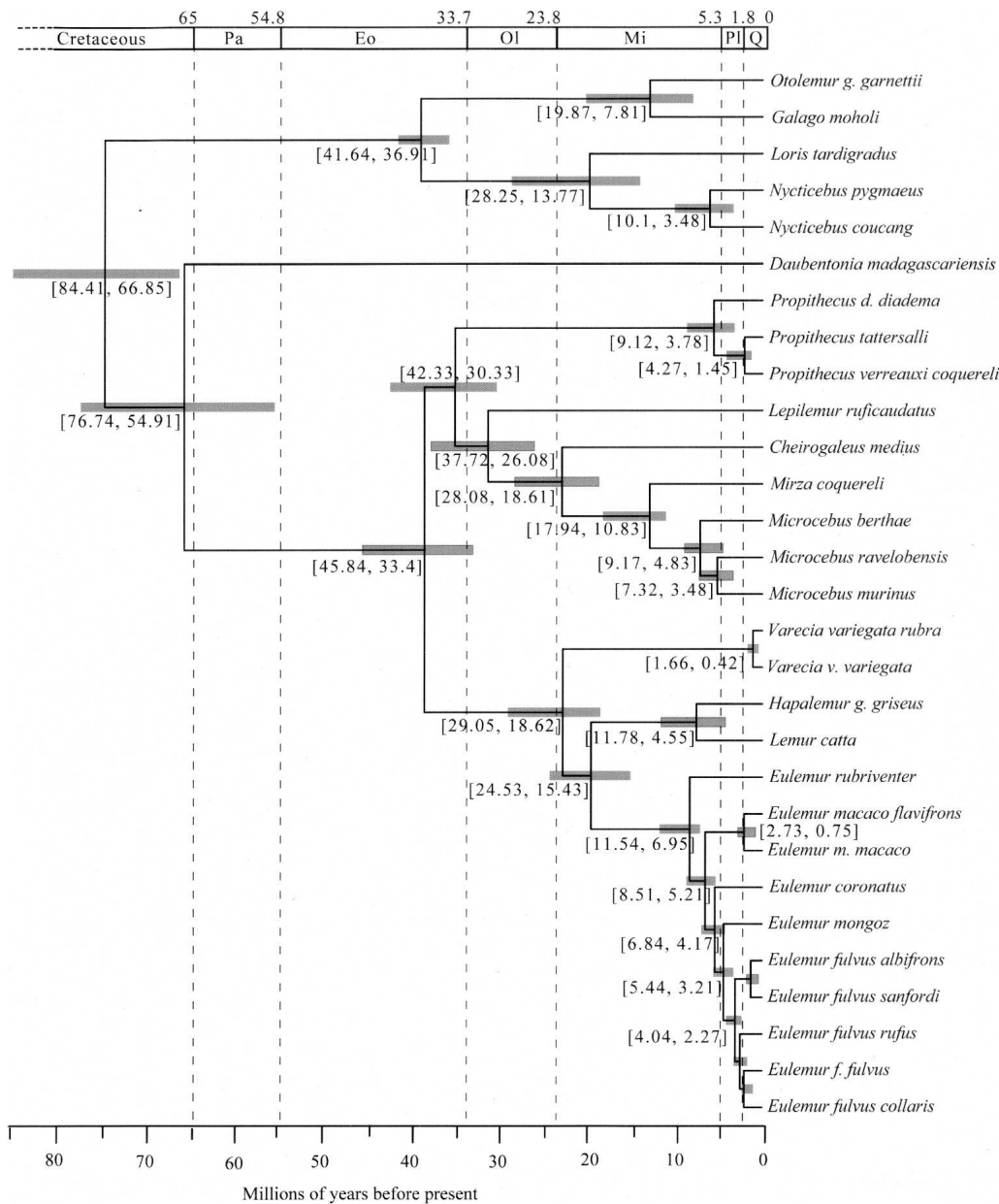
Figure 3. Divergence time estimates for the strepsirrhine lineages. An ultrametric tree with divergence time estimates resulting from the combined posterior distribution of four independent BEAST analyses (see Methods) of the combined nuclear data set. The results are based on prior date estimates of (1) the 6 Mya (5–7 Mya) split between *Homo sapiens* and *Pan troglodytes* (Kumar et al. 2005) and (2) the 40 Mya (38–42 Mya) split between the Lorisidae and Galagidae (Seiffert et al. 2003). Shaded gray boxes and numbers within brackets span the 95% highest posterior density of divergence time estimates. The scale bar is divided up by time (in millions of years before present) according to the Cretaceous, Tertiary, and Quaternary periods. The Tertiary period is shown according to epochs (Pa indicates Paleocene; Eo, Eocene; Ol, Oligocene; Mi, Miocene; Pl, Pliocene). Full details of time estimates from all BEAST analyses are presented in Table 4.

panel of four or 10 taxa often failed to amplify in multiple taxa, though no consistent pattern emerged.

PCR and sequencing

Standard PCR assays were conducted using ~50–150 ng DNA as template and Platinum Taq High Fidelity (Invitrogen) enzyme (for details, see Supplemental material). PCR products were directly sequenced using both forward and reverse PCR primers and BigDye Terminator v3.1 (Applied Biosystems) (for details, see Supplemental material). PCR products failing to produce high-

quality sequence by direct sequencing methods were ligated using the pGEM-T Easy Vector System II (Promega) with ~20–40 ng PCR product as insert (for details, see Supplemental Material). Sequence quality was assessed using either *phred/phrap/consed* (Gordon et al. 1998) or Sequencher (Gene Codes Corporation). Base calls with a *phred* quality score of less than 30 were scored as Ns. Accession numbers corresponding to each sequence are deposited in GenBank (Supplemental Table 6). Outgroup sequences (human, chimpanzee, macaque) were obtained using PSL map (http://hgdownload.cse.ucsc.edu/downloads.html) based on the human coordinates.

Table 4. Bayesian divergence estimates in millions of years

Node	Combined divergence estimates	Human–chimpanzee (5–7 Myr)	Galago-Loris (38–42 Myr)
	Divergence priors		
1	75.04 (66.85–84.41)	70.64 (60.84–81.67)	75.04 (66.86–84.87)
2	39.38 (36.91–41.64)	32.53 (26.08–39.58)	39.43 (37.20–41.84)
3	13.23 (7.81–19.87)	12.21 (7.67–17.22)	13.34 (7.74–19.65)
4	20.93 (13.77–28.25)	18.41 (13.18–24.43)	20.67 (13.84–27.85)
5	6.44 (3.48–10.1)	6.22 (3.58–9.37)	6.45 (3.50–10.0)
6	66.22 (54.91–74.74)	62.69 (52.76–74.36)	66.3 (54.68–77.54)
7	39.33 (33.4–45.84)	39.36 (33.81–45.63)	39.64 (33.58–46.22)
8	35.94 (30.33–42.33)	36.05 (30.54–42.01)	36.15 (29.96–42.56)
9	6.2 (3.78–9.12)	6.26 (3.97–9.04)	6.18 (3.72–9.00)
10	2.74 (1.45–4.27)	2.78 (1.55–4.23)	2.74 (1.39–4.28)
11	31.78 (26.08–37.72)	31.95 (26.62–37.58)	31.85 (26.12–38.21)
12	23.05 (18.61–28.08)	23.29 (18.85–28.26)	23.14 (18.24–28.26)
13	14.11 (10.83–17.94)	14.34 (10.94–18.00)	14.21 (10.78–17.93)
14	6.96 (4.83–9.17)	7.08 (5.00–9.29)	7 (4.74–9.24)
15	5.35 (3.48–7.32)	5.49 (3.59–7.54)	5.39 (3.53–7.52)
16	23.42 (18.62–29.05)	24.04 (19.10–29.64)	23.92 (18.69–29.83)
17	0.97 (0.42–1.66)	0.98 (0.42–1.64)	0.97 (0.40–1.66)
18	19.73 (15.43–24.53)	20.46 (15.71–25.62)	20.27 (15.43–25.51)
19	7.78 (4.55–11.78)	8.5 (4.66–12.09)	7.96 (4.36–12.46)
20	9.16 (6.95–11.54)	9.33 (7.11–11.74)	9.28 (6.94–11.83)
21	6.77 (5.21–8.51)	6.92 (5.33–8.61)	6.84 (5.14–8.66)
22	1.65 (0.75–2.73)	1.68 (0.80–2.73)	1.68 (0.74–2.76)
23	5.44 (4.17–6.84)	5.6 (4.32–7.04)	5.52 (4.13–7.04)
24	4.29 (3.21–5.44)	4.4 (3.31–5.56)	4.33 (3.18–5.58)
25	3.1 (2.27–4.04)	3.16 (2.33–4.10)	3.13 (2.26–4.09)

Phylogenetic analyses

All sequences were globally aligned using MUSCLE (http://www.drive5.com/muscle/; Edgar 2004a,b), with comparisons to ClustalW (Wilbur and Lipman 1983; Myers and Miller 1988) for manual editing. Regions of ambiguous alignment were excluded from analyses. Parsimony analyses were performed on each individual locus data set as well as on a combined data set with all 18 nuclear loci, including and excluding previously published mtDNA sequence data. Parsimony analyses were performed using the program PAUP* v4.0 (Swofford 2002). Heuristic searches were performed with 100 random-addition replicates and TBR branch swapping. All parsimony analyses gave equal weight to all character changes. To assess support for branches in parsimony trees, bootstrap percentages (BPs) were calculated using 1000 pseudo-replicates with 10 random additions per replicate. Decay indices were calculated using constraint trees generated in TreeRot v2 (Sorenson 1999) and analyzed in PAUP*.

Maximum likelihood and Bayesian analyses were performed using the best-fit model selected using the Akaike Information Criterion (AIC) in the program MrModeltest v2.2 (Nylander 2004). ML analyses were performed in the program GARLI v0.9.5.1 (www.bio.utexas.edu/faculty/antisense/garli/Garli.html; Zwickl 2006). Three replicate analyses were performed using random starting trees. Default conditions were used for the genetic algorithm search. Branch support for ML trees was generated using bootstrapping with 100 pseudoreplicates. Bootstrap analyses were performed using default conditions except that runs were set to terminate after 5000 generations without an improvement in topology. Combined GARLI analyses were unpartitioned using the general time-reversible (GTR) model with rate heterogeneity (Γ) and a proportion of sites assumed to be invariant (I). All Bayesian analyses (Huelsenbeck et al. 2001; Ronquist and Huelsenbeck 2003) were run with MrBayes v3.1.2 and used four Markov chains with the temperature profile at the default setting of 0.2. Flat dirichlet priors were used for GTR substitution-rate

parameters and for all base-frequency parameters. A flat Beta prior was used in estimating the transition/transversion substitution-rate parameter. Uniform priors were used for the Γ and I parameters. Unconstrained, uniform priors were used for topology and branch-length estimation. A molecular clock was not enforced. Combined analyses were performed in a partitioned framework, allowing for independent parameter estimation at each locus. Partitioned analyses involving mtDNA data treated MT-CO2 and MT-CYB as separate partitions. All Bayesian analyses were run for 1 million generations and replicated four times. Individual runs were compared with each other in the program Tracer v1.3 (Drummond et al. 2003) to assess convergence. Tree and parameter samplings from the latter 500,000 generations of each run were pooled and used as an estimate of the posterior distribution. A Bayesian concordance analysis was performed to assess concordance in gene trees across loci following the method of Ane et al. (2007). Full details of this method are given in the Supplementary materials.

Divergence time estimation

A rate-smoothed tree with divergence times was estimated from our 18 locus nuclear data set using a Bayesian relaxed clock method in BEAST v1.4 (Drummond et al. 2006). BEAST analyses used a relaxed lognormal model of lineage rate variation, a GTR+I+Γ model of nucleotide substitution, and a Yule prior for branching rates. No fossil-based time calibrations are currently available for the Lemuriformes (Simons et al. 1995). However, two calibrations are available that are relevant to our data set: (1) The split between slow lorises (*Loris* and *Nycticebus*) and galagos (*Galago* and *Otolemur*) has been dated at 38–42 Mya (Seiffert et al. 2003), and we used this date as a normal-distribution prior on this node with a mean of 40 Mya and a SD of 1.2 Myr. (2) The split between *H. sapiens* and *P. troglodytes* has been estimated through fossil and molecular evidence to have occurred at 5–7 Mya (Kumar et al. 2005), and we use this date as a normal-distribution prior with a mean of 6 Mya and a SD of 0.6 Myr. The result of each of these divergence time priors was assessed individually and in combination. In all BEAST analyses, we performed four replicate analyses of 10,000,000 generations each. The sampling distributions of each run were visualized using the program Tracer v1.3 (Drummond et al. 2003) to ensure the achievement of a stable posterior distribution in all parameter estimates, and then combined using the program LogCombiner v1.4. A consensus chronogram with node height distributions was then generated using TreeAnnotater v1.4. Because we only have two priors to calibrate our divergence calculations, our results may suffer from underestimates of nodal divergence times (Yoder and Yang 2004). Until basal terrestrial Tertiary fossil lemurs are discovered (if any even exist), these molecular calculations will serve as our best estimate of divergence dates for the Malagasy lemurs.

Acknowledgments

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.

Ane, C., Larget, B., Baum, D.A., Smith, S.D., and Rokas, A. 2007. Bayesian estimation of concordance among gene trees. *Mol. Biol. Evol.* **24:** 412–426.

Arnason, U., Gullberg, A., Burguete, A.S., and Janke, A. 2000. Molecular estimates of primate divergences and new hypotheses for primate dispersal and the origin of modern humans. *Hereditas* **133:** 217–228.

Bapteste, E., Brinkmann, H., Lee, J.A., Moore, D.V., Sensen, C.W., Gordon, P., Durufle, L., Gaasterland, T., Lopez, P., Muller, M., et al. 2002. The analysis of 100 genes supports the grouping of three highly divergent amoebae: Dictyostelium, Entamoeba, and Mastigamoeba. *Proc. Natl. Acad. Sci.* **99:** 1414–1419.

Brumfield, R.T., Beerli, P., Nickerson, D.A., and Edwards, S.V. 2003. The utility of single nucleotide polymorphisms in inferences of population history. *Trends Ecol. Evol.* **18:** 249–256.

Cannarozzi, G., Schneider, A., and Gonnet, G. 2007. A phylogenomic study of human, dog, and mouse. *PLoS Comput. Biol.* **3:** e2. doi: 10.1371/journal.pcbi.0030002.

Cerling, T.E., Harris, J.M., MacFadden, B.J., Leakey, M.G., Quade, J., Eisenmann, V., and Ehleringer, J.R. 1997. Global vegetation change through the Miocene/Pliocene boundary. *Nature* **389:** 153–158.

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. *Hum. Evol.* **10:** 35–44.

DelPero, M., Crovella, S., Cervella, P., Ardito, G., and Rumpler, Y. 1995. Phylogenetic relationships among Malagasy lemurs as revealed by mitochondrial DNA sequence analysis. *Primates* **36:** 431–440.

DelPero, M., Masters, J.C., Cervella, P., Crovella, S., Ardito, G., and Rumpler, Y. 2001. Phylogenetic relationships among the Malagasy lemuriforms (Primates: Strepsirrhini) as indicated by mitochondrial sequence data from the 12S rRNA gene. *Zool. J. Linn. Soc.* **133:** 83–103.

DelPero, M., Pozzi, L., and Masters, J.C. 2006. A composite molecular phylogeny of living lemuroid primates. *Folia Primatol.* **77:** 434–445.

Drummond, A., Pybus, O.G., and Rambaut, A. 2003. Inference of viral evolutionary rates from molecular sequences. *Adv. Parasitol.* **54:** 331–358.

Drummond, A.J., Ho, S.Y., Phillips, M.J., and Rambaut, A. 2006. Relaxed phylogenetics and dating with confidence. *PLoS Biol.* **4:** e88. doi: 10.1371/journal.pbio.0040088.

Dutrillaux, B. and Rumpler, Y. 1995. Phylogenetic relations among Prosimii with special reference to Lemuriformes and Malagasy nocturnals. In *Creatures of the dark. The nocturnal prosimians.* (eds. L. Alterman et al.), pp. 141–150. Plenum Press, New York.

Edgar, R.C. 2004a. MUSCLE: A multiple sequence alignment method with reduced time and space complexity. *BMC Bioinformatics* **5:** 113. doi: 10.1186/1471-2105-5-113.

Edgar, R.C. 2004b. MUSCLE: Multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* **32:** 1792–1797.

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

Fishbein, M., Hibsch-Jetter, C., Soltis, D.E., and Hufford, L. 2001. Phylogeny of saxifragales (angiosperms, eudicots): Analysis of a rapid, ancient radiation. *Syst. Biol.* **50:** 817–847.

Galewski, T., Tilak, M.K., Sanchez, S., Chevret, P., Paradis, E., and Douzery, E.J. 2006. The evolutionary radiation of Arvicolinae rodents (voles and lemmings): Relative contribution of nuclear and mitochondrial DNA phylogenies. *BMC Evol. Biol.* **6:** 80. doi: 10.1186/1471-2148-6-80.

Gordon, D., Abajian, C., and Green, P. 1998. *Consed*: A graphical tool for sequence finishing. *Genome Res.* **8:** 195–202.

Harlin-Cognato, A.D. and Honeycutt, R.L. 2006. Multi-locus phylogeny of dolphins in the subfamily Lissodelphininae: Character synergy improves phylogenetic resolution. *BMC Evol. Biol.* **6:** 87. doi:

10.1186/1471-2148-6-87.

Heckman, K.L., Mariani, C.L., Rasoloarison, R., and Yoder, A.D. 2007. Multiple nuclear loci reveal patterns of incomplete lineage sorting and complex species history within Western mouse lemurs (Microcebus). *Mol. Phylogenet. Evol.* **43:** 353–367.

Hedtke, S.M., Townsend, T.M., and Hillis, D.M. 2006. Resolution of phylogenetic conflict in large data sets by increased taxon sampling. *Syst. Biol.* **55:** 522–529.

Horvath, J.E. and Willard, H.F. 2007. Primate comparative genomics: Lemur biology and evolution. *Trends Genet.* **23:** 173–182.

Hudson, R.R. 1992. Gene trees, species trees and the segregation of ancestral alleles. *Genetics* **131:** 509–513.

Huelsenbeck, J.P. 1998. Systematic bias in phylogenetic analysis: Is the Strepsiptera problem solved? *Syst. Biol.* **47:** 519–537.

Huelsenbeck, J.P., Ronquist, F., Nielsen, R., and Bollback, J.P. 2001. Bayesian inference of phylogeny and its impact on evolutionary biology. *Science* **294:** 2310–2314.

James, T.Y., Kauff, F., Schoch, C.L., Matheny, P.B., Hofstetter, V., Cox, C.J., Celio, G., Gueidan, C., Fraker, E., Miadlikowska, J., et al. 2006. Reconstructing the early evolution of Fungi using a six-gene phylogeny. *Nature* **443:** 818–822.

Johnson, W.E., Eizirik, E., Pecon-Slattery, J., Murphy, W.J., Antunes, A., Teeling, E., and O'Brien, S.J. 2006. The late Miocene radiation of modern Felidae: A genetic assessment. *Science* **311:** 73–77.

Kappeler, P.M. 2000. Lemur origins: Rafting by groups of hibernators? *Folia Primatol.* **71:** 422–425.

Karanth, K.P., Delefosse, T., Rakotosamimanana, B., Parsons, T.J., and Yoder, A.D. 2005. Ancient DNA from giant extinct lemurs confirms single origin of Malagasy primates. *Proc. Natl. Acad. Sci.* **102:** 5090–5095.

Kriegs, J.O., Churakov, G., Kiefmann, M., Jordan, U., Brosius, J., and Schmitz, J. 2006. Retroposed elements as archives for the evolutionary history of placental mammals. *PLoS Biol.* **4:** e91. doi: 10.1371/journal.pbio.0040091.

Kumar, S., Filipski, A., Swarna, V., Walker, A., and Hedges, S.B. 2005. Placing confidence limits on the molecular age of the human-chimpanzee divergence. *Proc. Natl. Acad. Sci.* **102:** 18842–18847.

Larget, B. 2006. Bayesian untangling of Concordance Knots (BUCKy), version 1.1. Department of Statistics, University of Wisconsin. http://www.stat.wisc.edu/larget/bucky.html.

Lehman, S.M. 2006. Conservation biology of Malagasy strepsirhines: A phylogenetic approach. *Am. J. Phys. Anthropol.* **130:** 238–253.

Li, C. and Orti, G. 2007. Molecular phylogeny of Clupeiformes (Actinopterygii) inferred from nuclear and mitochondrial DNA sequences. *Mol. Phylogenet. Evol.* **44:** 386–398.

Li, C., Orti, G., Zhang, G., and Lu, G. 2007. A practical approach to phylogenomics: The phylogeny of ray-finned fish (Actinopterygii) as a case study. *BMC Evol. Biol.* **7:** 44. doi: 10.1186/1471-2148-7-44.

Malcomber, S.T. 2002. Phylogeny of Gaertnera Lam. (Rubiaceae) based on multiple DNA markers: Evidence of a rapid radiation in a widespread, morphologically diverse genus. *Evolution Int. J. Org. Evolution* **56:** 42–57.

McCall, R.A. 1997. Implications of recent geological investigations of the Mozambique Channel for the mammalian colonization of Madagascar. *Proc. Biol. Sci.* **264:** 663–665.

Murphy, W.J., Eizirik, E., Johnson, W.E., Zhang, Y.P., Ryder, O.A., and O'Brien, S.J. 2001. Molecular phylogenetics and the origins of placental mammals. *Nature* **409:** 614–618.

Murphy, W.J., Pevzner, P.A., and O'Brien, S.J. 2004. Mammalian phylogenomics comes of age. *Trends Genet.* **20:** 631–639.

Murphy, W.J., Pringle, T.H., Crider, T.A., Springer, M.S., and Miller, W. 2007. Using genomic data to unravel the root of the placental mammal phylogeny. *Genome Res.* **17:** 413–421.

Myers, E.W. and Miller, W. 1988. Optimal alignments in linear space. *Comput. Appl. Biosci.* **4:** 11–17.

Nikolaev, S., Montoya-Burgos, J.I., Margulies, E.H., Rougemont, J., Nyffeler, B., and Antonarakis, S.E. 2007. Early history of mammals is elucidated with the ENCODE multiple species sequencing data. *PLoS Genet.* **3:** e2. doi: 10.1371/journal.pgen.0030002.

Noonan, B.P. and Chippindale, P.T. 2006. Vicariant origin of malagasy reptiles supports late cretaceous antarctic land bridge. *Am. Nat.* **168:** 730–741.

Nunn, C.L. and Barton, R.A. 2001. Comparative methods for studying primate adaptation and allometry. *Evol. Anthropol.* **10:** 81–98.

Nylander, J.A.A. 2004. *MrModeltest v2.* Program distributed by the author. Evolutionary Biology Centre, Uppsala University, Uppsala, Sweden.

Palumbi, S.R. and Baker, C.S. 1994. Contrasting population structure from nuclear intron sequences and mtDNA of humpback whales. *Mol. Biol. Evol.* **11:** 426–435.

Pastorini, J., Thalmann, U., and Martin, R.D. 2003. A molecular approach to comparative phylogeography of extant Malagasy lemurs. *Proc. Natl. Acad. Sci.* **100:** 5879–5884.

Pollard, D., Iyer, V.N., Moses, A.M., and Eisen, M.B. 2006. Widespread discordance of gene trees with species tree in *Drosophila*: Evidence for incomplete lineage sorting. *PLoS Genet.* **2:** e173. doi: 10.1371/journal.pgen.0020173.

Porter, C.A., Sampaio, I., Schneider, H., Schneider, M.P., Czelusniak, J., and Goodman, M. 1995. Evidence on primate phylogeny from epsilon-globin gene sequences and flanking regions. *J. Mol. Evol.* **40:** 30–55.

Porter, C.A., Czelusniak, J., Schneider, H., Schneider, M.P., Sampaio, I., and Goodman, M. 1997a. Sequences of the primate epsilon-globin gene: Implications for systematics of the marmosets and other New World primates. *Gene* **205:** 59–71.

Porter, C.A., Page, S.L., Czelusniak, J., Schneider, H., Schneider, M.P.C., Sampio, I., and Goodman, M. 1997b. Phylogeny and evolution of selected primates as determined by sequences of the ε-globin locus and 5′ flanking regions. *Int. J. Primatol.* **18:** 261–295.

Poux, C. and Douzery, E.J. 2004. Primate phylogeny, evolutionary rate variations, and divergence times: A contribution from the nuclear gene IRBP. *Am. J. Phys. Anthropol.* **124:** 1–16.

Poux, C., Madsen, O., Marquard, E., Vieites, D.R., de Jong, W.W., and Vences, M. 2005. Asynchronous colonization of Madagascar by the four endemic clades of primates, tenrecs, carnivores, and rodents as inferred from nuclear genes. *Syst. Biol.* **54:** 719–730.

Rokas, A., Williams, B.L., King, N., and Carroll, S.B. 2003. Genome-scale approaches to resolving incongruence in molecular phylogenies. *Nature* **425:** 798–804.

Ronquist, F. and Huelsenbeck, J.P. 2003. MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* **19:** 1572–1574.

Roos, C., Schmitz, J., and Zischler, H. 2004. Primate jumping genes elucidate strepsirrhine phylogeny. *Proc. Natl. Acad. Sci.* **101:** 10650–10654.

Rosenberg, N.A. and Nordborg, M. 2002. Genealogical trees, coalescent theory and the analysis of genetic polymorphisms. *Nat. Rev. Genet.* **3:** 380–390.

Rozen, S. and Skaletsky, H. 2000. Primer3 on the WWW for general users and for biologist programmers. In *Bioinformatics methods and protocols: Methods in molecular biology* (eds. S. Krawetz and S. Misener), pp. 365–386. Humana Press, Totowa, NJ.

Ryder, O.A. 2005. Conservation genomics: Applying whole genome studies to species conservation efforts. *Cytogenet. Genome Res.* **108:** 6–15.

Seiffert, E.R., Simons, E.L., and Attia, Y. 2003. Fossil evidence for an ancient divergence of lorises and galagos. *Nature* **422:** 421–424.

Simons, E.L., Burney, D.A., Chatrath, P.S., Godfrey, L.R., Jungers, W.L., and Rakotosamimanana, B. 1995. AMS ^{14}C dates for extinct lemurs from caves in the Ankarana Massif, Northern Madagascar. *Quaternary Res.* **43:** 249–254.

Sorenson, M.D. 1999. *TreeRot, version 2*. Boston University, Boston.

Stanger, K.F. 1996. The phylogenetic relationships of the extant malagasy lemurs (Lemuriformes) based on DNA sequence data. *Am. Zool.* **36:** 134A.

Stanger-Hall, K.F. 1997. Phylogenetic affinities among the extant Malagasy lemurs (Lemuriformes) based on morphology and behavior. *J. Mamm. Evol.* **4:** 163–194.

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

Swofford, D.L. 2002. *PAUP*. Phylogenetic analysis using parsimony (and other methods)*. Version 4. Sinauer Associates, Sunderland, MA.

Takahashi, K., Terai, Y., Nishida, M., and Okada, N. 2001. Phylogenetic relationships and ancient incomplete lineage sorting among cichlid fishes in Lake Tanganyika as revealed by analysis of the insertion of retroposons. *Mol. Biol. Evol.* **18:** 2057–2066.

Teeling, E.C., Springer, M.S., Madsen, O., Bates, P., O'Brien, S.J., and Murphy, W.J. 2005. A molecular phylogeny for bats illuminates biogeography and the fossil record. *Science* **307:** 580–584.

Telford, M.J. 2007. Phylogenomics. *Curr. Biol.* **17:** R945–R946.

Townsend, T., Larson, A., Louis, E., and Macey, J.R. 2004. Molecular phylogenetics of squamata: The position of snakes, amphisbaenians, and dibamids, and the root of the squamate tree. *Syst. Biol.* **53:** 735–757.

Weisrock, D.W., Harmon, L.J., and Larson, A. 2005. Resolving deep phylogenetic relationships in salamanders: Analyses of mitochondrial and nuclear genomic data. *Syst. Biol.* **54:** 758–777.

Weisrock, D.W., Shaffer, H.B., Storz, B.L., Storz, S.R., and Voss, S.R. 2006. Multiple nuclear gene sequences identify phylogenetic species boundaries in the rapidly radiating clade of Mexican ambystomatid salamanders. *Mol. Ecol.* **15:** 2489–2503.

Wells, N.A. 2003. Some hypotheses on the Mesozoic and Cenozoic paleoenvironmental history of Madagascar. In *The natural history of Madagascar* (eds. S.M. Goodman and J.P. Benstead), pp. 16–74. University of Chicago Press, Chicago.

Wilbur, W.J. and Lipman, D.J. 1983. Rapid similarity searches of nucleic acid and protein data banks. *Proc. Natl. Acad. Sci.* **80:** 726–730.

Wilme, L., Goodman, S.M., and Ganzhorn, J.U. 2006. Biogeographic evolution of Madagascar's microendemic biota. *Science* **312:** 1063–1065.

Yoder, A.D. 1994. Relative position of the Cheirogaleidae in strepsirrhine phylogeny: A comparison of morphological and molecular methods and results. *Am. J. Phys. Anthropol.* **94:** 25–46.

Yoder, A.D. 2003. Phylogeny of the lemurs. In *The natural history of Madagascar* (eds. S.M. Goodman and J. Benstead), pp. 1242–1247. University of Chicago Press, Chicago.

Yoder, A.D. 2007. Lemurs. *Curr. Biol.* **17:** R866–R868.

Yoder, A.D. and Irwin, J.A. 1999. Phylogeny of the Lemuridae: Effects of character and taxon sampling on resolution of species relationships within *Eulemur*. *Cladistics* **15:** 351–361.

Yoder, A.D. and Yang, Z. 2004. Divergence dates for Malagasy lemurs estimated from multiple gene loci: Geological and evolutionary context. *Mol. Ecol.* **13:** 757–773.

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

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

Yoder, A.D., Burns, M.M., Zehr, S., Delefosse, T., Veron, G., Goodman, S.M., and Flynn, J.J. 2003. Single origin of Malagasy Carnivora from an African ancestor. *Nature* **421:** 734–737.

Zhang, N., O'Donnell, K., Sutton, D.A., Nalim, F.A., Summerbell, R.C., Padhye, A.A., and Geiser, D.M. 2006. Members of the *Fusarium solani* species complex that cause infections in both humans and plants are common in the environment. *J. Clin. Microbiol.* **44:** 2186–2190.

Zwickl, D.J. 2006. "Genetic algorithm approaches for the phylogenetic analysis of large biological sequence datasets under the maximum likelihood criterion." Ph.D. dissertation, The University of Texas at Austin, Austin.

Zwickl, D.J. and Hillis, D.M. 2002. Increased taxon sampling greatly reduces phylogenetic error. *Syst. Biol.* **51:** 588–598.