



## Resolution of the Early Placental Mammal Radiation Using Bayesian Phylogenetics

William J. Murphy, *et al.*  
*Science* **294**, 2348 (2001);  
DOI: 10.1126/science.1067179

**The following resources related to this article are available online at  
[www.sciencemag.org](http://www.sciencemag.org) (this information is current as of June 23, 2008 ):**

**Updated information and services**, including high-resolution figures, can be found in the online version of this article at:

<http://www.sciencemag.org/cgi/content/full/294/5550/2348>

**Supporting Online Material** can be found at:

<http://www.sciencemag.org/cgi/content/full/294/5550/2348/DC1>

A list of selected additional articles on the Science Web sites **related to this article** can be found at:

<http://www.sciencemag.org/cgi/content/full/294/5550/2348#related-content>

This article **cites 19 articles**, 9 of which can be accessed for free:

<http://www.sciencemag.org/cgi/content/full/294/5550/2348#otherarticles>

This article has been **cited by** 447 article(s) on the ISI Web of Science.

This article has been **cited by** 78 articles hosted by HighWire Press; see:

<http://www.sciencemag.org/cgi/content/full/294/5550/2348#otherarticles>

This article appears in the following **subject collections**:

Evolution

<http://www.sciencemag.org/cgi/collection/evolution>

Information about obtaining **reprints** of this article or about obtaining **permission to reproduce this article** in whole or in part can be found at:

<http://www.sciencemag.org/about/permissions.dtl>

blue numbers (Fig. 1) appear to follow logically as we correlate peaks back from GIS 8. However, this correlation places the beginning of GIS 12 (the end of H5) at 48 ka. Both GISP2 (25) and GRIP (23) as well two other high-precision stalagmite records place the end of H5 at ~45 ka (26, 27). The differences among stalagmite records [Hulu, Sorel Cave, Israel (26), and Crevice Cave, Missouri (27)] highlight important regional differences in past climate at this time. If we take the end of H5 at 45 ka as a tie point, we obtain the correlation depicted by the brown numbers. At present, we cannot distinguish between the two. Beyond GIS 13, the correlation appears straightforward and is consistent with the only other high-resolution speleothem correlation that covers this whole time range (27). The oldest part of our record correlates to the end of GIS 21.

The Hulu record identifies a link between the East Asian Monsoon and North Atlantic climate and supports the idea that millennial-scale events first identified in Greenland are hemispheric or wider in extent (28–32). The Greenland events have been explained by changing rates of North Atlantic deep water formation, resulting in changing heat transport to the North Atlantic (28, 33). The millennial-scale changes that we observe may result similarly from massive and rapid changes in oceanic and atmospheric circulation patterns. The temporal relations between the Hulu and Greenland deglacial sequences are consistent with North Atlantic events that trigger large-scale circulation changes (28). Regardless of the trigger, our observations are consistent with the idea that Northern Hemisphere atmospheric circulation patterns are more meridional in character during Greenland interstadials and are more zonal during stadials. Our data support the idea that changes in the East Asian Monsoon are integral to millennial-scale changes in atmospheric/oceanic circulation patterns and are affected by orbitally induced insolation variations; however, our data do not show clear evidence that sea level itself has an observable effect on East Asian Monsoon intensity.

References and Notes

1. Z. S. An, *Quat. Sci. Rev.* **19**, 171 (2000).
2. J. E. Kutzbach, *Science* **214**, 59 (1981).
3. S. C. Clemens, D. W. Murray, W. L. Prell, *Science* **274**, 943 (1996).
4. P. Wang, *Marine Geology* **156**, 5 (1999).
5. S. C. Porter, Z. S. An, *Nature* **375**, 305 (1995).
6. T. S. Liu, Ed., *Loess and Environment* (China Ocean Press, Beijing, 1985), pp. 168–183.
7. G. Kukla et al., *Geology* **16**, 811 (1988).
8. W. X. Li et al., *Nature* **339**, 534 (1989).
9. K. R. Ludwig et al., *Science* **258**, 284 (1992).
10. R. L. Edwards, J. H. Chen, G. J. Wasserburg, *Earth Planet. Sci. Lett.* **81**, 175 (1987).
11. X. Luo, M. Rehkämper, D. Lee, A. N. Halliday, *Inter. J. Mass Spectrom. Ion Proc.* **171**, 105 (1997).
12. C. C. Shen et al., *Chem. Geol.*, in press.
13.  $\delta^{18}\text{O}$  values were determined on a Finnigan MAT 251 (Bremen, Germany) at the Nanjing Institute of Geol-

ogy and Paleontology, Chinese Academy of Sciences. The GBW4405 carbonate powder standard (a Chinese National Material Standard from the Chinese National Standards Bureau) was used as a reference standard and run every nine samples. The mean and standard deviation of the population of 338 standard runs was  $-8.56 \pm 0.08\%$  (relative to VPDB) for  $\delta^{18}\text{O}$ , as compared with the Standards Bureau value of  $-8.49\%$ .

14. Supplemental data are available at Science Online at [www.sciencemag.org/cgi/content/full/294/5550/2345/DC1](http://www.sciencemag.org/cgi/content/full/294/5550/2345/DC1).
15. C. H. Hendy, *Geochim. Cosmochim. Acta* **35**, 801 (1971).
16. J. A. Dorale, R. L. Edwards, E. Ito, L. A. González, *Science* **282**, 1871 (1998).
17. S. Zheng et al., *Chinese Sci. Bull.* **13**, 801 (1983).
18. I. Friedman, J. R. O'Neil, *U.S. Geol. Surv. Prof. Paper* 440-KK (1977), p. 3.
19. B. Li, D. Yuan, J. Qin, Y. Lin, M. Zhang, *Sci. China Ser. D* **43**, 227 (2000).
20. A. Berger, *J. Atmos. Sci.* **35**, 2362 (1978).
21. D. Paillard, L. Labeyrie, P. Yiou, *EOS* **77**, 379 (1996).
22. The Greenland Summit Ice Cores [CD-ROM], 1997. Available from the National Snow and Ice Data Center, University of Colorado at Boulder, and

the World Data Center-A for Paleoclimatology, National Geophysical Data Center, Boulder, CO. Also available online at: [www.ngdc.noaa.gov/paleo/icecore/greenland/summit/index.html](http://www.ngdc.noaa.gov/paleo/icecore/greenland/summit/index.html).

23. W. Dansgaard et al., *Nature* **364**, 218 (1993).
24. G. C. Bond et al., *Nature* **365**, 143 (1993).
25. D. A. Meese et al., *J. Geophys. Res.* **102**, 26411 (1997).
26. M. Bar-Matthews, A. Ayalon, A. Kaufman, G. J. Wasserburg, *Earth Planet. Sci. Lett.* **166**, 85 (1999).
27. J. A. Dorale, thesis, University of Minnesota (2000).
28. W. S. Broecker, *Nature* **372**, 421 (1994).
29. M. Kienast et al., *Science* **291**, 2132 (2001).
30. I. Hendy, J. Kennett, *Geology* **27**, 291 (1999).
31. K. A. Hughen, J. R. Southon, S. J. Lehman, J. T. Overpeck, *Science* **290**, 1951 (2000).
32. H. Schulz et al., *Nature* **393**, 54 (1998).
33. D. Roemmich, *J. Phys. Oceanogr.* **10**, 1972 (1981).
34. We thank two anonymous reviewers for constructive criticisms that improved this contribution considerably. Supported by National Natural Science Foundation of China grant 49972055, U.S. NSF grants EAR-9712037 and ESH-9809459, and Chinese Academy of Sciences grants KZCX2-108 and KZCX1-Y-05.

20 July 2001; accepted 7 November 2001

# Resolution of the Early Placental Mammal Radiation Using Bayesian Phylogenetics

William J. Murphy,<sup>1\*</sup> Eduardo Eizirik,<sup>1,2\*</sup> Stephen J. O'Brien,<sup>1,†</sup> Ole Madsen,<sup>3</sup> Mark Scally,<sup>4,5</sup> Christophe J. Douady,<sup>4,5</sup> Emma Teeling,<sup>4,5</sup> Oliver A. Ryder,<sup>6</sup> Michael J. Stanhope,<sup>5,7</sup> Wilfried W. de Jong,<sup>3,8</sup> Mark S. Springer<sup>4,†</sup>

Molecular phylogenetic studies have resolved placental mammals into four major groups, but have not established the full hierarchy of interordinal relationships, including the position of the root. The latter is critical for understanding the early biogeographic history of placentals. We investigated placental phylogeny using Bayesian and maximum-likelihood methods and a 16.4-kilobase molecular data set. Interordinal relationships are almost entirely resolved. The basal split is between Afrotheria and other placentals, at about 103 million years, and may be accounted for by the separation of South America and Africa in the Cretaceous. Crown-group Eutheria may have their most recent common ancestry in the Southern Hemisphere (Gondwana).

Deciphering higher level relationships among mammalian orders is a difficult problem in systematics (1–6) and has important ramifi-

cations for evolutionary biology, genomics, and biomedical sciences (7, 8). Studies based on different, multikilobase molecular data sets (5, 6) independently resolved placental mammals into four superordinal groups: Afrotheria, Xenarthra, Laurasiatheria, and Euarchontoglires. However, hierarchical relationships within these groups and at deeper levels in the placental tree remain unclear. A precise resolution of the relationships among the major groups and elucidation of the root of the placental tree are critical for interpreting biogeographic patterns and evolutionary processes involved in the early diversification of placental mammals.

We combined and expanded the large data sets of Madsen et al. (5) and Murphy et al. (6) to yield a 16,397–base pair molecular data set that includes 19 nuclear and 3 mitochon-

<sup>1</sup>Laboratory of Genomic Diversity, National Cancer Institute, Frederick, MD 21702, USA. <sup>2</sup>Department of Biology, University of Maryland, College Park, MD 20742, USA. <sup>3</sup>Department of Biochemistry, University of Nijmegen, Netherlands. <sup>4</sup>Department of Biology, University of California, Riverside, California 92521, USA. <sup>5</sup>Queen's University of Belfast, Biology and Biochemistry, Belfast, UK. <sup>6</sup>Center for Reproduction of Endangered Species, Zoological Society of San Diego, San Diego, CA 92112, USA. <sup>7</sup>Bioinformatics, Glaxo-SmithKline, Collegeville, PA 19426, USA. <sup>8</sup>Institute for Systematics and Population Biology, Amsterdam, Netherlands.

\*These authors contributed equally to this work. †To whom correspondence should be addressed. E-mail: mark.springer@ucr.edu (M.S.S.) or obrien@ncifcrf.gov (S.J.O.)

REPORTS

drial gene sequences for 42 placentals (representing all major lineages) and 2 marsupial outgroups (9). The data set is dominated by nuclear exons. Among genes that have been evaluated for resolving deep level mammalian relationships, nuclear exons have more power than mitochondrial genes on a per-residue basis (10). Large molecular data sets have the potential to resolve longstanding controversies in systematics (5, 6, 11, 12), especially when they are analyzed with appropriate models of DNA sequence evolution and with statistically robust estimation procedures that extract the maximum amount of information from molecular sequence data (13).

We used a general-time-reversible + gamma + invariants (GTR +  $\Gamma$  + I) model of sequence evolution (14) and likelihood-based inferential techniques (15), including parametric bootstrap tests (13, 16, 17) and Bayesian methods (18, 19) with Markov chain Monte Carlo (MCMC) sampling to assess phylogenetic relationships and examine alternative positions for the root of the placental tree. Given enough data and a correct model of sequence evolution, likelihood methods are statistically consistent and have been shown to be powerful tools for resolving complex phylogenetic problems (13).

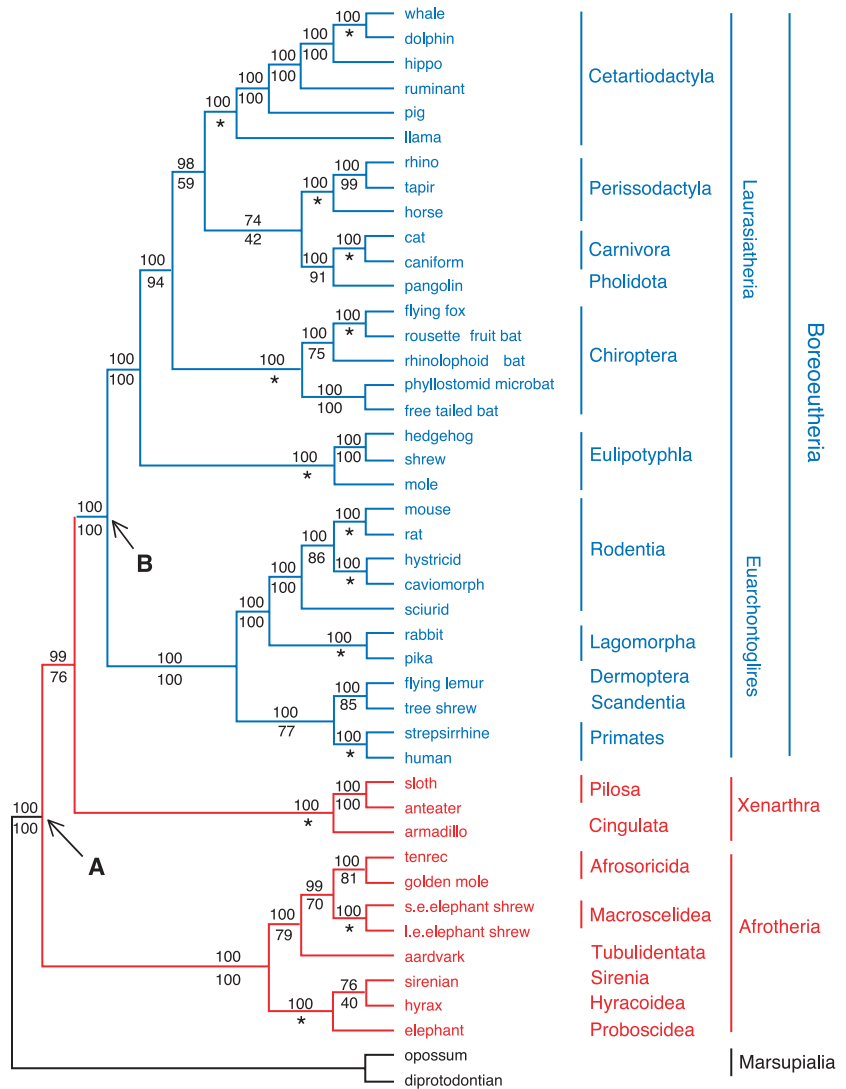
Figure 1 shows an evolutionary tree with Bayesian posterior probabilities for individual branches. Three independent MCMC runs, each starting with random trees for each of four simultaneous chains, resulted in concordant joint posterior probability distributions for the topology and the estimated parameters of the model of sequence evolution (15). This result suggests that the chains were run for a sufficient number of generations and sampled the same posterior probability landscape. With the exception of two nodes, the entire placental superordinal tree is resolved with posterior probabilities greater than 0.95. All additional Bayesian analyses that varied taxon and gene sampling (15) resulted in well-resolved trees with high posterior probabilities (15). Furthermore, an identical topology was also obtained with maximum likelihood (ML), strengthening our confidence in the Bayesian results. For some nodes, nonparametric ML bootstrap values were lower than Bayesian posterior probabilities (Fig. 1), consistent with the suggestion of Hillis and Bull (20) that nonparametric bootstrap support may be too conservative.

Our results firmly place Laurasiatheria and Euarchontoglires as sister taxa that together constitute a clade named Boreoeutheria, with a Northern Hemisphere origin according to the available fossil record (15, 21). Deeper in the placental tree, Xenarthra and Boreoeutheria are sister taxa. The basal

split among crown-group placentals is between Afrotheria versus Xenarthra + Boreoeutheria (Fig. 1). Previous molecular studies (5, 6) had suggested three most likely positions for the root of the placental tree: (i) the base of Afrotheria, (ii) the base of Xenarthra, or (iii) the branch that separates Xenarthra and Afrotheria from Boreoeutheria. Bayesian results and parametric bootstrap tests reject both the Xenarthra root (posterior probabilities < 0.01; para-

metric bootstrap  $P \leq 0.01$ ) and the root between Xenarthra + Afrotheria and Boreoeutheria (posterior probabilities < 0.04; parametric bootstrap  $P < 0.01$ ) (Fig. 1) (15).

Within each of the four major clades, relationships that were previously unresolved or controversial are also well resolved. There is now strong support for a basal split between paenungulates versus other afrotherians (armadillo, elephant shrews, and afro-



**Fig. 1.** Phylogeny of living placental mammals reconstructed using a Bayesian phylogenetic approach. An identical topology was obtained with maximum likelihood [ $-\ln L = 211110.54$ ; see (15) for methodological details]. The number above each branch refers to the Bayesian posterior probability (shown as percentages; i.e., 95 represents a posterior probability of 0.95) of the node derived from 26,250 MCMC sampled trees on the basis of the complete 16.4-kb data. Additional analyses with the full data set and with data sets that varied taxon sampling (i.e., jackknifing single outgroup taxa) and character sampling (nuclear only and nuclear coding loci only) produced similarly high posterior probabilities (15). Values below branches represent percent support in maximum likelihood (GTR +  $\Gamma$  + I) nonparametric bootstrap. An asterisk indicates nodes constrained in the ML nonparametric bootstrap analysis. (A) Bifurcation between Afrotheria and Xenarthra + Boreoeutheria at approximately 103 million years, which corresponds to the vicariant event that separated Africa and South America (Fig. 2B). (B) Branch where dispersal from South America to Laurasia is hypothesized to have occurred (15). Blue, monophyletic Northern Hemisphere group (i.e., Boreoeutheria); red, paraphyletic Southern Hemisphere group (i.e., Xenarthra + Afrotheria); black, outgroups.



cidans). Among the latter, afrosoricidans and elephant shrews were well supported as sister taxa. Within Euarchontoglires, our molecular results are the first to render robust support for the monophyly and internal structure of Euarchonta (3). Euarchonta is similar to the morphology-based Archonta hypothesis, but bats are excluded. Inside Laurasiatheria there is now strong support for the basal position of Eulipotyphla and a carnivore + pangolin clade.

The resolution of the placental root and the pattern of basal divergences lead to a plausible biogeographic inference regarding the origin and diversification of this group. Afrotheria and Xenarthra have Gondwanan origins in Africa and South America, respectively. Given their basal positions in the placental tree, the hypothesis that crown-group eutherians have their most recent common ancestry in Gondwana demands consideration. This view is at odds with a prevalent and long-held view that crown-group eutherians have their most recent common ancestry in the Northern Hemisphere (22–24). We tested this hypothesis by estimating molecular divergence dates using the quartet dating (25) and linearized tree methods (26) for specific nodes in the placental tree (Fig. 2A) (15). Our point estimates for the basal split among living placentals range from 101 to 108 million years ago (Mya) (Fig. 2A) (15), in agreement with independent molecular estimates of the split between Afrotheria-Boreoeutheria representatives at 105 Mya (27). It is striking that this date coincides with the vicariant event that separated South America and Africa approximately 100 to 120 Mya (Fig. 2B) (28, 29). We suggest a causal relationship between the sundering of Africa and South America and basal cladogenesis among crown-group eu-

therians, placing their origin in Gondwana. Subsequently, a trans-hemispheric dispersal event from Gondwana to Laurasia was of fundamental importance in the early history of crown-group eutherians (15). Given the sister-group relationship between Xenarthra and Boreoeutheria, the directionality of this event was likely from South America to Laurasia. Molecular dates for the Xenarthra-Boreoeutheria split (88 to 100 Mya) and for basal divergences within Boreoeutheria (79 to 88 Mya) suggest a window during the Late Cretaceous during which dispersal occurred (Fig. 2A) (15).

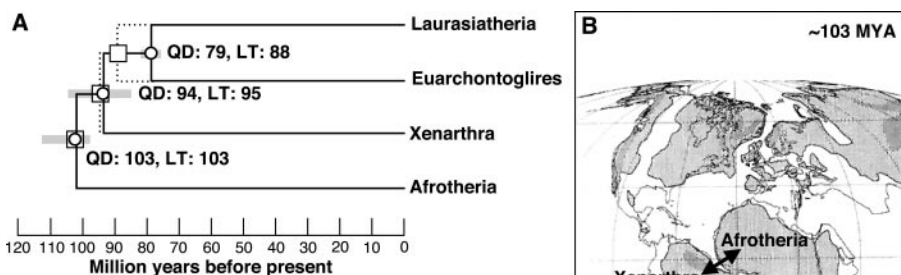
The suggestion that crown-group placentals have their most recent common ancestry in Gondwana does not imply that stem eutherians also have their origins in Gondwana. Kumar and Hedges (27) and Penny *et al.* (30) suggested that marsupials and placentals split 173 to 176 Mya on the basis of molecular data. At this time, Gondwana and Laurasia remained connected and stem eutherians may have established a Pangean distribution before the vicariant separation of Gondwana and Laurasia 160 to 170 Mya (28, 29). Early Cretaceous eutherians in the Northern Hemisphere, such as *Prokennalestes* (31), may be representatives of an extinct evolutionary radiation in Laurasia that predates Boreoeutheria.

Foote *et al.* (32) argued that Cretaceous molecular dates for the early evolutionary history of crown-group placentals are incongruent with the fossil record. One explanation for the discrepancy between molecular dates and paleontological data is the “Garden of Eden” hypothesis (32), which postulates early placental diversification in regions with a poorly known fossil record. Notably, the Cretaceous fossil record for most of Gondwana fits this description (32). A second explanation is the phylogenetic placement of fossil taxa (32). Foote *et*

*al.* (32) assumed that Cretaceous eutherians were stem taxa rather than crown-group eutherians. In contrast, the cladistic analysis of Archibald *et al.* (33) suggests that 85- to 90-million-year-old zalambdalestids and zhelestids, both from Laurasia, are members of crown-group Eutheria and have affinities with Glires and archaic ungulates, respectively. In the context of our molecular results, zalambdalestids may be early representatives of Euarchontoglires, and zhelestids may be an early branch within Laurasiatheria. Reliable phylogenetic placements for these and other Cretaceous taxa are critical for evaluating whether the available fossil record is compatible with Cretaceous molecular dates. The well-resolved phylogeny that we have established for living orders of placental mammals provides a molecular scaffold (34) that may be used in cladistic analyses of morphological characters to examine inter-relationships among living and fossil forms. Understanding the phylogenetic and biogeographic patterns connecting extinct and living mammalian lineages remains one of the major challenges ahead.

References and Notes

1. G. G. Simpson, *Bull. Am. Mus. Nat. Hist.* **85**, 1 (1945).
2. M. J. Novacek, *Nature* **356**, 121 (1992).
3. P. J. Waddell, N. Okada, M. Hasegawa, *Syst. Biol.* **48**, 1 (1999).
4. F. G. Liu *et al.*, *Science* **291**, 1786 (2001).
5. O. Madsen *et al.*, *Nature* **409**, 610 (2001).
6. W. J. Murphy *et al.*, *Nature* **409**, 614 (2001).
7. S. J. O'Brien, E. Eizirik, W. J. Murphy, *Science* **292**, 2264 (2001).
8. K. R. Wollenberg, W. R. Atchley, *Proc. Natl. Acad. Sci. U.S.A.* **97**, 3288 (2000).
9. The data set (available upon request from M.S.S. or S.J.O.) is a concatenation of 19 nuclear gene segments (*ADORA3*, *ADRA2B*, *ADRB2*, *APP*, *ATP7A*, *BDNF*, *BMI1*, *BRCA1*, *CNR1*, *CREM*, *EDG1*, *IRBP*, *PLCB4*, *PNOG*, *RAG1*, *RAG2*, *TYR*, *VWF*, and *ZFX*) and the two complete mitochondrial ribosomal RNA (rRNA) subunit genes (12S and 16S) plus the intervening valine tRNA. Seventy-six new nucleotide sequences were obtained using novel and previously described PCR primers (5, 6, 15). Sequences were collected on ABI 377 or 3700 automated sequencers (ABI, Inc., Foster City, CA) using BigDye terminator chemistry (ABI, Inc.). Alignments are as previously published (5, 6), modified by eye to include new taxa. All regions for which reliable homology could not be established (e.g., mitochondrial rRNA and tRNA hypervariable loops, polynucleotide expansions in untranslated regions) were removed from the analyses, resulting in a 16,397 base pair data set. Because the two original data sets sampled different species, we have performed phylogenetic analyses on a data set where taxa from the same (monophyletic) taxonomic group were combined to produce chimeric sequences (15). Methodological details for Bayesian and maximum-likelihood phylogenetic analyses are given in supplemental material (15).
10. M. S. Springer *et al.*, *Mol. Biol. Evol.* **18**, 132 (2001).
11. D. M. Hillis, *Nature* **383**, 130 (1996).
12. Y. L. Qui *et al.*, *Nature* **402**, 404 (1999).
13. S. Whelan, P. Lio, N. Goldman, *Trends Genet.* **17**, 262 (2001).
14. Estimated using the program Modeltest [described by D. Posada, K. A. Crandall, *Bioinformatics* **14**, 817 (1998)].
15. Supplemental material is available on Science Online



**Fig. 2.** Biogeographic scenario for the basal divergence among crown-group placental mammals. **(A)** Maximum likelihood molecular divergence estimates for the early radiation of placental mammals, estimated with the quartet-dating (QD) and linearized tree (LT) methods (25, 26). Open squares, point estimates based on LT; open circles, median point estimates based on QD; gray bars, range of 95% confidence intervals based on QD. A summary of QD and LT methods and results can be found in supplemental material (15). **(B)** Final vicariant separation of Africa and South America, approximately 100 to 120 Mya (28, 29), isolates Afrotheria in Africa and the common ancestor of Xenarthra and Boreoeutheria in South America. Reprinted with permission from Cambridge University Press (28).

at [www.sciencemag.org/cgi/content/full/294/5550/2348/DC1](http://www.sciencemag.org/cgi/content/full/294/5550/2348/DC1).

16. D. M. Hillis, B. K. Mable, C. Moritz, Eds., *Molecular Systematics* (Sinauer, Sunderland, MA, 1996), pp. 515–543.

17. N. Goldman, J. P. Anderson, A. G. Rodrigo. *Syst. Biol.* **49**, 652 (2000).

18. B. Larget, D. Simon, *Mol. Biol. Evol.* **16**, 750 (1999).

19. J. P. Huelsenbeck, F. Ronquist. *Bioinformatics* **17**, 754 (2001).

20. D. M. Hillis, J. J. Bull, *Syst. Biol.* **42**, 182 (1993).

21. E. Eizirik, W. J. Murphy, S. J. O'Brien. *J. Hered.* **92**, 212 (2001).

22. A. R. Wallace, *The Geographical Distribution of Animals* (Hafner, New York, 1962).

23. W. D. Matthew, *Ann. N.Y. Acad. Sci.* **24**, 171 (1915).

24. R. Rainger, *Agenda for Antiquity: Henry Fairfield Osborn and Vertebrate Paleontology at the American Museum of Natural History, 1890–1935* (Univ. of Alabama Press, Tuscaloosa, 1991).

25. A. Rambaut, L. Bromham, *Mol. Biol. Evol.* **15**, 442 (1998).

26. N. Takezaki, A. Rzhetsky, M. Nei, *Mol. Biol. Evol.* **12**, 823 (1995).

27. S. Kumar, S. B. Hedges, *Nature* **392**, 917 (1998).

28. A. G. Smith, D. G. Smith, B. M. Funnell, *Atlas of Cenozoic and Mesozoic Coastlines*. (Cambridge Univ. Press, Cambridge, 1994).

29. W. W. Hay *et al.*, *Geol. Soc. Am. Spec. Pap.* **332** (1999), pp. 1–48.

30. D. Penny, M. Hasegawa, P. J. Waddell, M. D. Hendy, *Syst. Biol.* **48**, 76 (1999).

31. M. C. McKenna, S. K. Bell, *Classification of Mammals Above the Species Level* (Columbia Univ. Press, New York, 1997).

32. M. Foote, J. P. Hunter, C. M. Janis, J. J. Sepkoski Jr., *Science* **283**, 1310 (1999).

33. J. D. Archibald, A. O. Averlanov, D. G. Ekdale, *Nature* **414**, 62 (2001).

34. M. S. Springer *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **98**, 6241 (2001).

35. Supported by the NSF (M.S.S.), the Training and Mobility of Researchers (TMR) program of the European Commission (W.W.d.J. and M.J.S.), and the Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), Brazil (E.E.).

17 October 2001; accepted 12 November 2001

# The Closest Living Relatives of Land Plants

Kenneth G. Karol,<sup>1\*</sup> Richard M. McCourt,<sup>2</sup> Matthew T. Cimino,<sup>1</sup> Charles F. Delwiche<sup>1</sup>

The embryophytes (land plants) have long been thought to be related to the green algal group Charophyta, though the nature of this relationship and the origin of the land plants have remained unresolved. A four-gene phylogenetic analysis was conducted to investigate these relationships. This analysis supports the hypothesis that the land plants are placed phylogenetically within the Charophyta, identifies the Charales (stoneworts) as the closest living relatives of plants, and shows the Coleochaetales as sister to this Charales/land plant assemblage. The results also support the unicellular flagellate *Mesostigma* as the earliest branch of the charophyte lineage. These findings provide insight into the nature of the ancestor of plants, and have broad implications for understanding the transition from aquatic green algae to terrestrial plants.

The evolutionary origin of the embryophytes (or land plants) from their green algal ancestor was a pivotal event in the history of life. This monophyletic group has altered the biosphere and now dominates the terrestrial environment, but uncertainty as to the identity of their closest living relatives has persisted in the literature after more than a century of scrutiny (1–3). Morphological and molecular studies have identified two distinct lineages within the green plants *sensu lato*, termed Charophyta and Chlorophyta. The Charophyta comprise the land plants and at least five lineages (orders) of fresh water green algae, and are sister to the Chlorophyta, which consist of essentially all other green algae. Previous molecular analyses have verified monophyly of most of the charophyte orders (4–6), but branching patterns among these lineages have been only weakly supported, with results that were sensitive to taxon selection and method of phylogenetic reconstruction. Similarly, analyses of mor-

phological and genome structural data have clarified some relationships (7–10), but have been limited by the number of characters available, uncertain homology assessment, and a lack of character independence.

Identifying the closest living relatives of land plants has been difficult. Roughly 470 million years of evolution since the colonization of the land, coupled with rapid radiation and numerous extinction events (2, 3, 11), has resulted in an inherently difficult phylogenetic problem, with much information from the early, common history of evolution obscured by subsequent evolution in the now independent lineages (12).

To investigate the evolutionary origin of land plants and identify the closest living relatives of this group, we analyzed DNA sequence data from four genes representing three plant genomes: *atpB* and *rbcL* (plastid), *nad5* (mitochondrial), and the small subunit (SSU) rRNA gene (nuclear). The data set used for phylogenetic analyses excludes introns and unalignable regions for a total length of 5147 base pairs [Appendix 1 (13)] (14). We sampled 34 representative charophytes, including eight land plants, and six outgroup taxa [Appendix 2 (13)]. The data were analyzed with Bayesian inference (BI), maximum likelihood (ML), maximum parsimony (MP), and minimum evolution with

two distance measures [LogDet (ME-ld) and maximum likelihood (GTR+I+ $\Gamma$ ; ME-ml) distances] [Appendix 3 (13)]. Both BI and ML are probabilistic methods that utilize explicit models of sequence evolution to test phylogenetic hypotheses. Advantages of BI are that it is relatively fast and provides probabilistic measures of tree strength that are more directly comparable with traditional statistical measures than those more commonly used in phylogenetic analyses (15, 16). To measure phylogenetic stability, posterior probabilities (PP) as inferred by BI were calculated and bootstrapping was performed for the ML, MP, and ME analyses.

Using BI and ML on the combined four-gene data set (Fig. 1), we found the order Charales sister to the land plants with strong statistical support (PP = 1.0, ML = 94) and a monophyletic Coleochaetales sister to the Charales/land plant clade (PP = 1.0, ML = 59). The MP and ME analyses [Appendix 4 (13)] also support the result that Charales have a closer relationship to land plants than do Coleochaetales (MP = 80, ME-ld = 97, ME-ml = 92). The overall structure of the best tree is consistent with previous work in that the classically recognized orders were also recovered (land plants, PP = 1.0, ML = 100, MP = 100, ME-ld = 100, ME-ml = 100; Charales, PP = 1.0, ML = 100, MP = 100, ME-ld = 100, ME-ml = 100; Coleochaetales, PP = 1.0, ML = 62, MP = <50, ME-ld = 75, ME-ml = <50; Zygnematales, PP = 1.0, ML = 99, MP = 93, ME-ld = 68, ME-ml = <50; and Klebsormidiales PP = 1.0, ML = 100, MP = 100, ME-ld = 100, ME-ml = 100). There was also support for placement of the enigmatic filamentous alga *Entransia* (6) with the Klebsormidiales (PP = 1.0, ML = 77, MP = 77, ME-ld = <50, ME-ml = 64). The rare, monotypic genus *Chlorokybus* was found sister to the remainder of the unambiguous charophytes, while all analyses strongly support the inclusion of *Mesostigma* within the Charophyta (PP = 1.0, ML = 97, MP = 100, ME-ld = 100, ME-ml = 100).

The phylogenetic placement of *Mesostigma*, a unicellular, scaly green flagellate

<sup>1</sup>Department of Cell Biology and Molecular Genetics, University of Maryland, College Park, MD 20742, USA.

<sup>2</sup>Department of Botany, Academy of Natural Sciences, 1900 Benjamin Franklin Parkway, Philadelphia, PA 19103, USA.

\*To whom correspondence should be addressed. E-mail: karol@umail.umd.edu