Estimating the timing of early eukaryotic diversification with multigene molecular clocks

Laura Wegener Parfrey[a,b,2], Daniel J. G. Lahr[a,b], Andrew H. Knoll[c,1], and Laura A. Katz[a,b,1]

[a]Program in Organismic and Evolutionary Biology, University of Massachusetts, Amherst, MA 01003; [b]Department of Biological Sciences, Smith College, Northampton, MA 01063; and [c]Department of Organismic and Evolutionary Biology, Harvard University, Cambridge, MA 02138

Contributed by Andrew H. Knoll, July 1, 2011 (sent for review February 9, 2011)

Although macroscopic plants, animals, and fungi are the most familiar eukaryotes, the bulk of eukaryotic diversity is microbial. Elucidating the timing of diversification among the more than 70 lineages is key to understanding the evolution of eukaryotes. Here, we use taxon-rich multigene data combined with diverse fossils and a relaxed molecular clock framework to estimate the timing of the last common ancestor of extant eukaryotes and the divergence of major clades. Overall, these analyses suggest that the last common ancestor lived between 1866 and 1679 Ma, consistent with the earliest microfossils interpreted with confidence as eukaryotic. During this interval, the Earth's surface differed markedly from today; for example, the oceans were incompletely ventilated, with ferruginous and, after about 1800 Ma, sulfidic water masses commonly lying beneath moderately oxygenated surface waters. Our time estimates also indicate that the major clades of eukaryotes diverged before 1000 Ma, with most or all probably diverging before 1200 Ma. Fossils, however, suggest that diversity within major extant clades expanded later, beginning about 800 Ma, when the oceans began their transition to a more modern chemical state. In combination, paleontological and molecular approaches indicate that long stems preceded diversification in the major eukaryotic lineages.

microbial eukaryotes | Proterozoic oceans | taxon sampling | origin of eukaryotes

The antiquity of eukaryotes and the tempo of early eukaryotic diversification remain open questions in evolutionary biology. Proposed dates for the origin of the domain based on the fossil record and molecular clock analyses differ by up to 2 billion years (1). Microfossils attributed to eukaryotes occur at about 1800 Ma (2) and putative biomarkers of early eukaryotes have been found in 2700 Ma rocks (3). Such geological interpretations contrast with both molecular clock studies that place the origin of eukaryotes at 1250–850 Ma (4, 5), and a controversial hypothesis that rejects the eukaryotic interpretation of all older fossils and places eukaryogenesis at 850 Ma (6, 7).

Paleontologists generally agree that an unambiguous record of eukaryotic microfossils extends back to ∼1800 Ma (2, 8, 9). Microfossils of this age are assigned to eukaryotes because they combine informative characters that include complex morphology (e.g., the presence of processes and evidence for real-time modification of vegetative morphology), complex wall ultrastructure, and specific inferred behaviors (2, 9, 10). Despite being interpreted as eukaryotic, the taxonomic affinities of these fossils remain unclear (2). Eukaryotic fossils that can be assigned to extant taxonomic groups begin to appear ∼1200 Ma (11) and become more widespread, abundant, and diverse in rocks ∼800 Ma and younger (2, 12, 13).

Molecular estimation of divergence times has improved dramatically in recent years due the development of methods that incorporate uncertainty from sources that include phylogenetic reconstruction, fossil calibrations, and heterogeneous rates of molecular evolution (1, 14, 15). Relaxed clock approaches account for heterogeneity in evolutionary rates across branches and enable the use of complex models of sequence evolution (reviewed in refs. 16 and 17), although debate continues as to the best method for relaxing the clock (18–20). The process of calibrating molecular clocks has also been greatly improved with both the recognition that single calibration points are insufficient (21, 22), and the availability of methods incorporate uncertainty from the fossil record by specifying calibrations as time distributions rather than points (15, 16). Additional limitations in previous molecular clock studies of eukaryotes stem from the tradeoff between analyses of many taxa and calibration points but only a single gene (4), and analyses of many genes but a small number of taxa and calibrations (5, 23).

Molecular clock estimates rely on robust phylogenies. Reconstructions of relationships among eukaryotes have begun to stabilize in recent years with the increasing availability of multigene data from diverse lineages (24–26). The majority of the >70 lineages of eukaryotes fall within four major groups: Opisthokonta; Excavata; Amoebozoa; and Stramenopiles, Alveolates, and Rhizaria (SAR) (25, 26), while the placement of some photosynthetic lineages remains controversial (25, 27, 28). Greater data availability also yields more accurate estimates of divergence times because more nodes are available for calibration (29).

The availability of taxon- and gene-rich datasets coupled with flexible molecular clock methods make this an ideal time to revisit the timing of early eukaryotic evolution. Here, broadly sampled multigene trees are used to estimate dates, with rate heterogeneity across the tree and among genes incorporated into the model. We use 23 calibration points derived from diverse fossils of Proterozoic and Phanerozoic age specified as prior distributions (Table 1). The Proterozoic fossil record is sparse (2, 8, 9), and the taxonomic assignment of some Proterozoic fossils has been called into question by a minority of researchers (6). In the spirit of testing these ideas, we assess the impact of including calibration constraints derived from Phanerozoic fossils alone and Phanerozoic plus Proterozoic fossils. We also assess divergence dates across analyses that varied in the position of the root, and the number of taxa included, as well as across different software platforms and models.

Results

Taxon-rich analyses of multiple genes reveal a stability in divergence dates across the eukaryotic tree of life that is robust to changing taxon inclusion, position of the root, molecular clock model, and choice of calibration points (Phanerozoic only or both Phanerozoic and Proterozoic fossils). Collectively, these analyses provide a mean age for the root of extant eukaryotes to 1866–1679 Ma in analyses including both Proterozoic and Phanerozoic calibrations ("*All*" analyses; Fig. 1*A* and Table S1). Varying the position of the root had little impact on divergence

Author contributions: L.W.P. and L.A.K. designed research; L.W.P. and D.J.G.L. performed research; L.W.P., D.J.G.L., A.H.K., and L.A.K. analyzed data; and L.W.P., D.J.G.L., A.H.K., and L.A.K. wrote the paper.

The authors declare no conflict of interest.

Freely available online through the PNAS open access option.

[1]To whom correspondence may be addressed. E-mail: aknoll@oeb.harvard.edu or lkatz@smith.edu

[2]Present address: Department of Chemistry and Biochemistry, University of Colorado, Boulder, CO 80309.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1110633108/-/DCSupplemental.

Table 1. Calibration constraints for dating the eukaryotic tree of life

Taxon	Fossil	Eon*	Calibration†		Ref(s).
			Min	Dist	
Amniota	*Westlonthania*	Phan	328.3	4, 3	(54)
Angiosperms	Oldest angio pollen	Phan	133.9	2, 10	(55)
Ascomycetes	*Paleopyrenomycites*	Phan	400	4, 50	(56)
Coccolithophores	Earliest Heterococcolith	Phan	203.6	2, 8	(57)
Diatoms	Earliest diatoms	Phan	133.9	2, 100	(58)
Dinoflagellates	Earliest gonyaulacales	Phan	240	2, 10	(59)
Embryophytes	Land plant spores	Phan	471	2, 20	(60)
Endopterygota	Mecoptera	Phan	284.4	5, 5	(61)
Eudicots	Eudicot pollen	Phan	125	2, 1.5	(62, 63)
Euglenids	*Moyeria*	Phan	450	2, 40	(64)
Foraminifera	Oldest forams	Phan	542	2, 200	(65)
Gonyaulacales	Gonyaulacaceae split	Phan	196	2,10	(59)
Pennate diatoms	Oldest pennate	Phan	80	3, 5	(66)
Spirotrichs	Oldest tintinnids	Phan	444	2.5, 100	(67)
Trachaeophytes	Earliest trachaeophytes	Phan	425	4, 2.5	(68)
Vertebrates	*Haikouichthys*	Phan	520	3, 5	(69)
Animals	LOEMs, sponge biomarkers	Protero	632	2, 300	(70, 71)
Arcellinida	*Paleoarcella*	Protero	736	2, 300	(12)
Bilateria	*Kimberella*	Protero	555	2, 30	(72)
Chlorophytes	*Palaeastrum*	Protero	700	2.5, 300	(73)
Ciliates	Gammacerane	Protero	736	2.5, 300	(74)
Florideophyceae	Doushantuo red algae	Protero	550	2.5, 100	(75)
Red algae‡	*Bangiomorpha*	Protero	1174	3, 250	(11)

*Eon: Phan, Phanerozoic; Protero, Proterozoic. Proterozoic calibrations are excluded from *Phan* analyses.
†Calibration constraints are specified for BEAST using a gamma distribution with a minimum date in Ma based on the fossil record parameters as indicated: min, minimum divergence data; dist, gamma prior distribution (shape, scale). See Table S3 for details of PhyloBayes calibrations.
‡In the *All 720* analysis (*c*), the minimum age constraint for the red algae node is set to 720 Ma.

dates, especially for the estimated date of the root itself, which generally changed by <100 million years (myr; Fig. 1*A*). PhyloBayes estimates generally showed more uncertainty than those from BEAST analyses, but around similar means. Similarly, estimates were robust to changing models (uncorrelated or autocorrelated) and to the inclusion of only Phanerozoic (*Phan*) or all calibrations (*All*) with one exception: under the autocorrelated Cox–Ingersoll–Ross (CIR) model, estimates are much more recent in *Phan* analyses (1038 Ma and 1180 Ma; Fig. 1*A*).

Impact of Calibration Constraints on Estimates of the Origin of Extant Eukaryotes. We assessed the impact of including Proterozoic fossils, which are considered controversial by some (6, 7), by analyzing datasets without these seven calibration constraints (*Phan* analyses). In BEAST analyses, the exclusion of Proterozoic fossils shifted estimated divergence times toward the present, but not dramatically so: estimates for the mean age of root of extant eukaryotes fall between 1506–1471 Ma in *Phan* analyses [95% highest-probability density (HPD) range 1643–1347 Ma; Fig. 1*A*, Figs. S1, S5, and S7, analyses *b*, *f*, and *h*] compared with 1837–1717 Ma (95% HPD range 1954–1601 Ma; Figs. 1*A* and 2 and Figs. S4 and S6; analyses *a*, *e*, and *g*) when Proterozoic fossils were included (*All* analyses). Similar dates were recovered in *Phan* and *All* PhyloBayes analyses when the uncorrelated gamma model (UGAM) model (uncorrelated) of the molecular clock was assumed (Fig. 1*A*, analyses *i–l*).

Of the seven Proterozoic calibration points used in our analyses, only the *Bangiomorpha* point is controversial in terms of either systematic attribution or age. The *Bangiomorpha* calibration constraint is more than 400 myr older than our other Proterozoic constraints (Table 1). To determine whether this calibration point drives results in analyses with *All* calibrations, we assessed the age of the root with a much more conservative estimate for the age of this red alga (*All 720*; Fig. 1, analysis *c*). A number of factors place the age of *Bangiomorpha* ~1200 Ma (*SI Text*); however, given the importance of the fossil we assigned an age of 720 Ma to this constraint, representing the absolute younger bound of the Hunting Formation, Canada, in which it is found (*SI Text*) (11). In BEAST, placing the *Bangiomorpha* constraint at 720 Ma shifted the estimated age of the root by only 95 myr toward the present (Fig. 1*A* and Fig. S3, analysis *c*).

The autocorrelated CIR model combined with the low number of substitutions on deep branches of the eukaryotic tree appears more sensitive to the distribution of calibration dates included in these analyses. Under the CIR autocorrelated model, a consistent age was estimated with *All* calibrations included (1798–1691 Ma; Fig. 1*A*, analyses *m* and *o*), although confidence intervals are greater in PhyloBayes analyses in general (Fig. 1*A*, analyses *i–p*). However, excluding Proterozoic calibration points did cause estimated ages to shift more than 600 myr younger under the CIR model (1180–1038 Ma; Fig. 1*A*, analyses *n* and *p*), pushing the estimated age for the root of extant eukaryotes younger than the widely accepted date for the *Bangiomorpha* fossils. Similarly, the CIR analyses in PhyloBayes were sensitive to the age of the *Bangiomorpha* constraint, shifting more than 500 myr younger to 1296 Ma and 1167 Ma in analyses with *All* calibration points rooted with Opisthokonta and "Unikonta," respectively (Dataset S1). The necessity of using PhyloBayes to explore the differences between autocorrelated and uncorrelated models introduces confounding factors, as PhyloBayes requires both uniform distributions around calibration points and a fixed tree topology. Given that calibration points are likely best represented by more informative distributions, and that the topology of the tree is not fully known, we focus the rest of our discussions on the results from BEAST, although data from all PhyloBayes analyses are available in Fig. 1*A* and Dataset S1.

Origin of Major Clades. In most analyses, the major clades of extant eukaryotes diverged before 1200 Ma, with SAR, Excavata, and Amoebozoa arising within a similar time frame, as evidenced by overlapping 95% HPD ranges (Figs. 1 and 2, Figs. S1–S7, and Dataset S1). The 95% HPD intervals are wider for clades with few

EVOLUTION

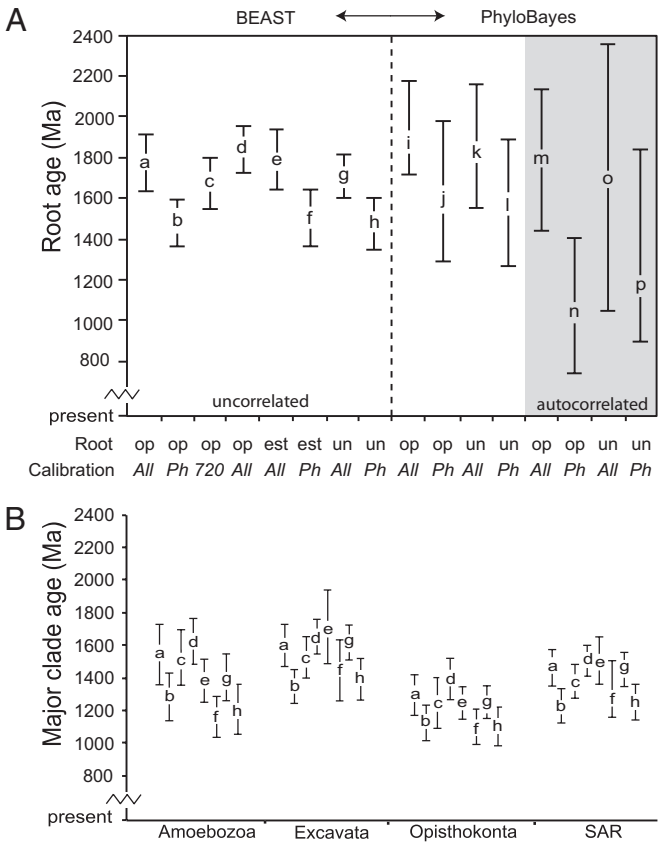

Fig. 1. Summary of mean divergence dates for the most recent common ancestor of major clades of extant eukaryotes. Letters are at the mean divergence time and denote analyses, as detailed in Table S1. Error bars represent 95% HPD for BEAST analyses (*a*–*h*) and the 95% confidence interval for PhyloBayes (analysis *i*–*p*). (*A*) Estimated age of the root of extant eukaryotes across analyses. Root position: Opis, root constrained to Opisthokonta; Uni, root constrained to "Unikonta"; Estim, root estimated by BEAST. Calibration: *All*, all Phanerozoic and Proterozoic CCs; *Phan*, Phanerozic CCs only; 720, *All* CCs with the minimum age of red algae set to 720 Ma. d = 91 taxa. (*B*) Estimated ages of major clades from BEAST analyses.

calibration points, such as Excavata and Amoebozoa (Fig. 1*B*). Estimates for the last common ancestor of extant Opisthokonta are younger than the other clades, at 1389–1240 Ma in analyses with *All* calibration constraints.

Exclusion of Proterozoic calibration constraints (*Phan* analyses) shifted age estimates for the origins of major extant eukaryotic clades younger by 200–300 myr (Fig. 1*B*). Differences in divergence times are relatively small for nested clades—e.g., the 95% HPD for Alveolata shifts from 1445 to 1236 Ma in analysis *a* (Fig. 2) to 1206–1020 Ma with only Phanerozoic calibration points (analysis *b*; Fig. S1). Not surprisingly, the differing calibration schemes had their most dramatic impact on the estimated age of the red algae, which changes from 1285 to 1180 Ma 95% HPD (Fig. 2) to 959–625 Ma 95% HPD when Proterozoic calibration points, including the constraint on red algae at 1174 Ma in accordance with the widely cited age for *Bangiomorpha*, are excluded (Fig. S1). Estimated ages of major clades were also much younger in analyses using the CIR model with *Phan* calibrations (analyses *n* and *p*; Dataset S1).

The topology of the eukaryotic tree produced through coestimation of phylogeny and divergence times in BEAST is broadly consistent with other analyses (*SI Text*) (25, 26). Hence, the BEAST topology was also used for the PhyloBayes analyses, which require a fixed topology. Though the relationships among the photosynthetic eukaryotes remain uncertain (25), our analyses suggest that many photosynthetic clades, such as red and green algae, diverged within a similar time frame (Fig. 2). These results imply an early acquisition of photosynthesis in eukaryotes, in accordance with both previous molecular clock estimates (30) and the ~1200 Ma age assigned to the red algal fossil *Bangiomorpha* (11).

Discussion

The molecular clock analyses presented here suggest that the last common ancestor of extant eukaryotes lived between 1866 and 1679 Ma when both Phanerozoic and Proterozoic fossils are considered. We favor these more-inclusive analyses as they should reveal a more accurate picture of eukaryotic diversification, especially because the chosen fossils are widely accepted by paleontologists, and calibration constraints were assigned in a conservative manner that accounts for age uncertainties. Estimated ages are younger when we remove Proterozoic calibration constraints, though not dramatically so, with the notable exception of the autocorrelated model CIR as implemented in PhyloBayes with only Phanerozoic calibrations. Thus, our results tend to place the last common ancestor of extant eukaryotes deep within the Proterozoic Eon.

Our estimates for the timing of the origin of extant eukaryotes are in line with fossil evidence (2, 13), but reject the hypothesis that eukaryotes originated only 850 Ma (6, 7). Fossils provide minimum dates, leaving open the possibility that clades evolved much earlier than their first fossil appearance (2, 31). Thus, it is not surprising that divergence times for many eukaryotic clades are older than their first unambiguous fossil occurrence (Table 2). The paleontological literature contains some references to eukaryotic fossils older than our estimate of the last common ancestor. In some cases, these paleontological reports are incorrect or ambiguous. For example, large carbonaceous fossils assigned to the genus *Grypania* were originally reported to be older than our molecular clock estimate (32), but more recent radiometric dates indicate an age of 1874 ± 9 Ma (33), consistent with the clock analyses presented here. Older still are the 50- to 300-μm spheroidal microfossils described from ~3200 Ma rocks by Javaux et al. (34), and proposed as possible eukaryotes by Buick (35), and sterane biomarkers from 2700 Ma shales (3). Whether these materials record Archean eukaryotes remains a subject of debate (34, 36). Our molecular clock estimates suggest that if these fossils do represent eukaryotes, they record stem lineages—early representatives of eukaryotic groups that went extinct—that were present before the emergence of extant eukaryotic clades.

The major lineages of extant eukaryotes (Opisthokonta, SAR, Excavata, and Amoebozoa) are projected to have diverged from one another by the Mesoproterozoic era (1600–1000 Ma), relatively early in the history of the domain (Fig. 1 and Table 2). This, in turn, suggests that these lineages were present for hundreds of millions of years before the observed increase in the abundance and diversity of eukaryotic microfossils beginning ~800 Ma (2, 37–40). Our molecular clock estimates indicate that stem groups were present well before recognizable members of crown lineages—monophyletic groups consisting of living representatives and their ancestors—diversified. A similar pattern of long stems preceding diversification is seen in animal and plants and may be a consistent pattern in evolution (38).

Fossils and our molecular clock analyses agree that eukaryotes originated and diversified during a time when oceans differed substantially from the modern seas. Increasingly, geochemical data indicate that for much of the Proterozoic eon, mildly oxic surface waters lay above an oxygen-minimum zone that was persistently anoxic and commonly sulfidic (41, 42). Such conditions are compatible with scenarios for eukaryogenesis that rely on anaerobic methanogens in symbiotic partnership with facultatively aerobic proteobacteria or sulfate reducers (see references in ref. 43), because facultatively anaerobic mitochondria may have enabled early eukaryotes to live in the sulfidic Proterozoic oceans (44). Because sulfide interferes with the function of mitochondria in aerobically respiring eukaryotes, the radiation of diverse species within eukaryotic clades may have become pos-

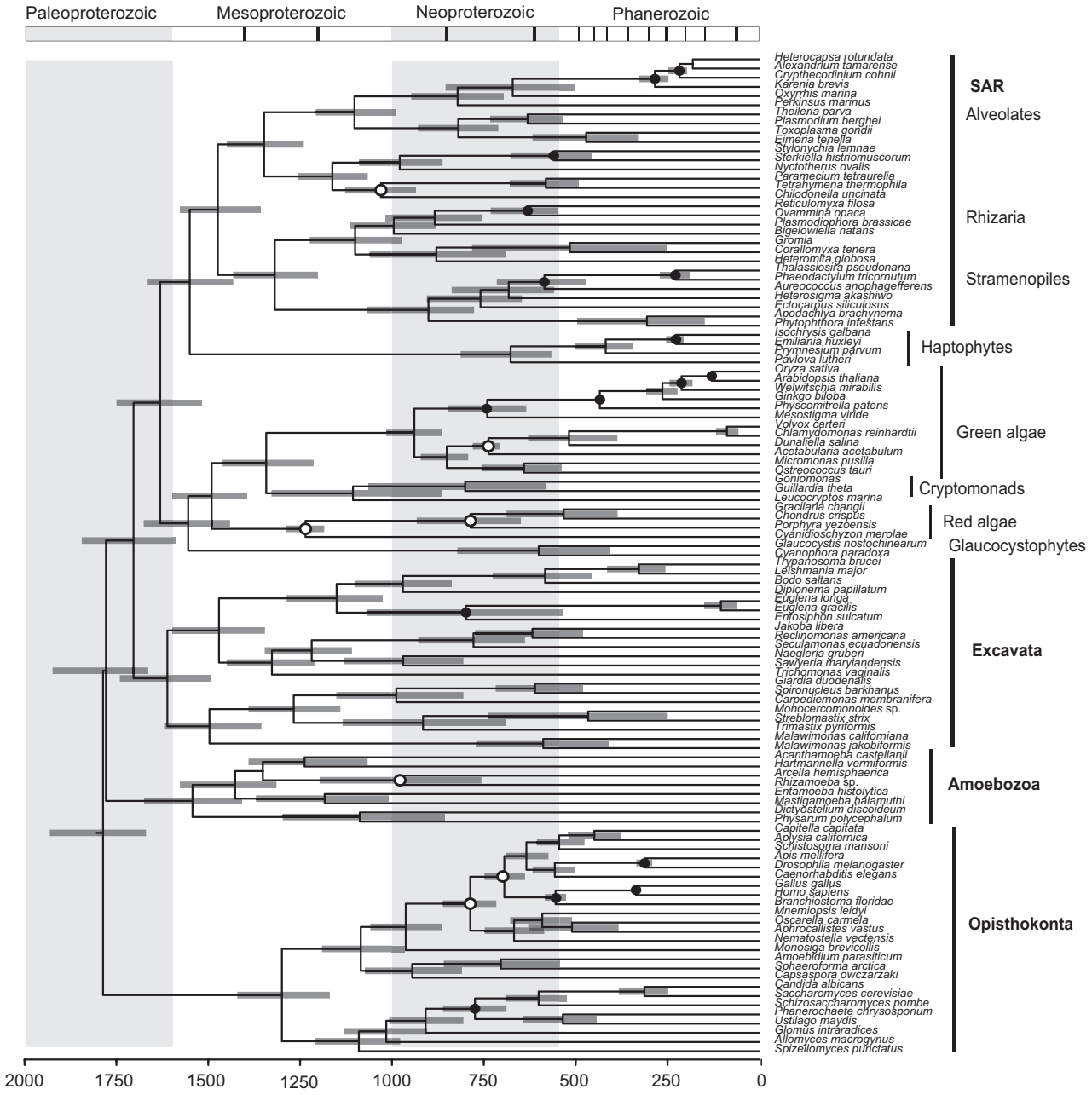

Fig. 2. Time-calibrated tree of extant eukaryotes using *All* calibration points, 109 taxa, and root constrained to Opisthokonta. Nodes are at mean divergence times and gray bars represent 95% HPD of node age. (*Upper*) Geological time scale; (*Lower*) Absolute time scale in Ma. Thick vertical bars demarcate eras and thin vertical lines denote periods, with dates derived from the 2009 International Stratigraphic Chart. Node calibrated with Phanerozoic fossils (●); node calibrated with Proterozoic fossils (○). Estimated ages of calibrated nodes differ from calibration constraints (Table 1) because they have been modified by relaxed clock analysis of sequence data.

sible only when sulfidic subsurface waters began to wane about 800 Ma (45). Alternatively, early eukaryotic evolution may have occurred in coastal environments sheltered from the impact of sulfidic waters or in freshwater systems, which are both poorly sampled by the geologic record and not impacted by sulfidic oceanic water masses (46). Consistent with this view, moderately diverse assemblages of fossil eukaryotes occur in well-ventilated lake deposits of the 1200 to 900 Ma Torridonian succession, Scotland (47, 48), and in coastal marine deposits of the ~1500 to 1400-Ma Roper Group, Australia (49).

Within Proterozoic oceans, low concentrations of biologically available nitrogen may also have inhibited the diversification of photosynthetic eukaryotes (50). Many cyanobacteria and other photosynthetic bacteria are capable of nitrogen fixation, ameliorating the impact of nitrate and ammonia limitation on primary production. Eukaryotes, however, have no such capacity; thus, it may not be a coincidence that biomarkers indicating an expanding importance of algae in marine primary production occur in conjunction with geochemical data recording the spread of oxygen through later Neoproterozoic oceans (51). In our analyses, the clade that contains extant photosynthetic taxa, including green algae plus land plant and red algae, arose between 1670 and 1428 Ma, but diversification within these lineages occurred later in the Neoproterozoic and may correspond to a changing redox profile in the oceans (Fig. 2).

Table 2. Comparison of major node ages to fossil dates

Major clade	Estimated age, Ma	Oldest fossil, Ma	Ref.
Eukaryotes	*	1800	(2)
Extant eukaryotes	1679–1866	1200	(11)
Amoebozoa	1384–1624	800	(12)
Excavata	1510–1699	450	(64)
Opisthokonta	1240–1481	632	(71)
Rhizaria	1017–1256	550	(65)
SAR	1365–1577	736	(74)

Estimated age is range of mean dates from *All* analyses.
*The age of the root of all eukaryotes is not estimated because molecular clock studies can only inform the timing of extant clades.

Discrepancy Between These and Previous Molecular Clock Studies. Previous molecular clock studies yielded vastly different dates for the root of extant eukaryotes, ranging from 3970 to 1100 Ma (1). In a recent analysis of small subunit ribosomal DNA (SSU-rDNA) from 83 broadly sampled eukaryotes, Berney and Pawlowski (4) placed the origin of eukaryotes at 1100 Ma, a conclusion that was robust to changing the position of the root. They had numerous Phanerozoic calibration constraints specified as either minimum or maximum divergence dates (4), but they found that including Proterozoic calibration points, such as *Bangiomorpha* at 1200 Ma, shifted their estimates of the origin and diversification of eukaryotes by 1000–2500 Ma. The age discrepancy observed by Berney and Pawlowski (4), when Proterozoic calibration constraints are included, contrasts sharply with the relative stability of dates seen in our analyses (Fig. 1*A*). We hypothesize that the increased gene and taxon sampling, as well as the use of flexible prior distributions of calibration points as implemented in BEAST, are major factors contributing to the stability of molecular clock estimation in our analyses.

Conclusion

Our molecular clock analyses yield a timeline of eukaryotic evolution that is congruent with the paleontological record and robust to varying analytical conditions. According to our analyses, crown (extant) groups of eukaryotes arose in the Paleoproterozoic era (2500–1600 Ma) and began to diversify soon thereafter, suggesting that early eukaryotic evolution was influenced by anoxic and sulfidic water masses in contemporaneous oceans. The stability in our analysis across a range of variables is a welcome departure from the large age discrepancies reported in earlier molecular analyses, reflecting improved paleontological interpretation, advancements in molecular methods, and the rapidly growing body of molecular data from diverse eukaryotes.

Materials and Methods

Alignments. Alignments are derived from the 15 protein-coding genes analyzed in Parfrey et al. (dataset 15:10 of ref. 25). Using this 88-taxon dataset as a starting point, taxa were added to capture additional lineages, particularly those with fossil data available (Table S2). Rapidly evolving taxa (e.g., *Encephalitozoon cuniculi*) and orphans (e.g., *Breviata anathema*) were removed to minimize rate heterogeneity for the clock analysis. The resulting 109-taxon data matrix includes 5,696 characters, with each taxon having between three and 15 of the target genes (36% missing character data; Table S2; analyses *a*–*c* and *e*–*p*). A 91-taxon alignment was created by removing additional taxa with either long branches or high levels of missing data to ensure that our results were not driven by these potential sources of artifact (analysis *d*).

Molecular Dating Analyses. Dating analyses were predominantly performed in BEAST v1.5.4 (52), and we also assessed results obtained in PhyloBayes 3.2f (53) (see *SI Text* for analysis details). BEAST offers a number of desirable features, including flexible specification of prior distributions that enable the uncertainty of the fossil record to be realistically modeled, as well as the ability to coestimate divergence times with topology (15). We compared divergence dates for eukaryotes obtained from different models to assess whether our conclusions were driven by the choice of a particular model (*SI Text*, Fig. 1 and Table S1).

Calibration Constraints. Calibration constraints were specified with prior distributions to incorporate errors arising from age dating, stratigraphy, and clade assignment (Table 1). The impact of Proterozoic fossils was assessed by analyzing the data with only the 16 Phanerozoic calibration constraints (*Phan* analyses *b*, *f*, *h*, *j*, *l*, *n*, and *p*) or with Phanerozoic and Proterozoic calibration constraints (*All* analyses *a*, *c*–*e*, *g*, *i*, *k*, *m*, and *o*). Calibration constraints were specified with prior distributions in BEAST using BEAUTi v1.5.4 (52) and were derived from a conservative reading of the fossil record (i.e., we err toward younger rather than older ages; *SI Text*). Distributions were specified with long tails unless the fossil record provided minimum-divergence information. Calibration constraints used for PhyloBayes had to be specified as a uniform distribution (Table S3).

Assessing Impact of the Root on the Inferred Age of Eukaryotes. Molecular clock analyses require a rooted tree. However, the position of the eukaryotic root remains an open question; therefore, we compared age estimates from molecular clock analyses with multiple positions for the root of extant eukaryotes. First, the root was constrained to the branch leading to the Opisthokonta or to Opisthokonta + Amoebozoa ("Unikonta") in accordance with current hypotheses (see *SI Text* for discussion of the position of the eukaryotic root). In BEAST, the root was specified by constraining a monophyletic ingroup. PhyloBayes requires the tree topology to be fixed, and we used the tree in Fig. 2 rooted on either Opisthokonta or "Unikonta". Finally, for the third condition, the root was estimated by the molecular clock criterion, as implemented in BEAST (*SI Text*), which yielded variable estimates of the location of the root.

ACKNOWLEDGMENTS. We thank Ben Normark, Rob Dorit, and Sam Bowser for useful discussions, and Jeff Thorne and Bengt Sennblad for helpful discussions about molecular clock models. This manuscript has been improved following the comments of Emmanuelle Javaux, Andrew Roger, and Heroen Verbruggen. We thank Jessica Grant and Tony Caldanaro for technical help. This research was supported by the National Aeronautics and Space Administration Astrobiology Institute (A.H.K.) and by National Science Foundation Assembling the Tree of Life Grant 043115 and National Science Foundation Systematics Grant 0919152 (to L.A.K). D.J.G.L. is supported by Conselho Nacional de Desenvolvimento Científico e Tecnológico-Brazil Doutorado no Exterior Fellowship 200853/2007-4.

1. Roger AJ, Hug LA (2006) The origin and diversification of eukaryotes: Problems with molecular phylogenetics and molecular clock estimation. *Philos Trans R Soc Lond B Biol Sci* 361:1039–1054.
2. Knoll AH, Javaux EJ, Hewitt D, Cohen P (2006) Eukaryotic organisms in Proterozoic oceans. *Philos Trans R Soc Lond B Biol Sci* 361:1023–1038.
3. Brocks JJ, Logan GA, Buick R, Summons RE (1999) Archean molecular fossils and the early rise of eukaryotes. *Science* 285:1033–1036.
4. Berney C, Pawlowski J (2006) A molecular time-scale for eukaryote evolution recalibrated with the continuous microfossil record. *Proc Roy Soc Lond B* 273:18671872.
5. Douzery EJP, Snell EA, Bapteste E, Delsuc F, Philippe H (2004) The timing of eukaryotic evolution: Does a relaxed molecular clock reconcile proteins and fossils? *Proc Natl Acad Sci USA* 101:15386–15391.
6. Cavalier-Smith T (2002) The phagotrophic origin of eukaryotes and phylogenetic classification of Protozoa. *Int J Syst Evol Microbiol* 52:297–354.
7. Cavalier-Smith T (2010) Deep phylogeny, ancestral groups and the four ages of life. *Philos Trans R Soc Lond B Biol Sci* 365:111–132.
8. Porter SM (2004) The fossil record of early eukaryotic diversification. *Paleontol Soc Papers* 10:35–50.
9. Javaux EJ, Knoll AH, Walter M (2003) Recognizing and interpreting the fossils of early eukaryotes. *Orig Life Evol Biosph* 33:75–94.
10. Javaux EJ, Knoll AH, Walter MR (2004) TEM evidence for eukaryotic diversity in mid-Proterozoic oceans. *Geobiology* 2:121–132.
11. Butterfield NJ (2000) *Bangiomorpha pubescens* n. gen., n. sp.: Implications for the evolution of sex, multicellularity, and the Mesoproterozoic/Neoproterozoic radiation of eukaryotes. *Paleobiol* 26:386–404.
12. Porter SM, Meisterfeld R, Knoll AH (2003) Vase-shaped microfossils from the Neoproterozoic Chuar Group, Grand Canyon: A classification guided by modern testate amoebae. *J Paleontol* 77:409–429.
13. Javaux EJ (2007) The early eukaryotic fossil record. *Adv Exp Med Biol* 607:1–19.
14. Welch JJ, Bromham L (2005) Molecular dating when rates vary. *Trends Ecol Evol* 20: 320–327.
15. Drummond AJ, Ho SYW, Phillips MJ, Rambaut A (2006) Relaxed phylogenetics and dating with confidence. *PLoS Biol* 4:e88.
16. Ho SYW, Phillips MJ (2009) Accounting for calibration uncertainty in phylogenetic estimation of evolutionary divergence times. *Syst Biol* 58:367–380.

17. Rutschmann F (2006) Molecular dating of phylogenetic trees: A brief review of current methods that estimate divergence times. *Divers Distrib* 12:35–48.
18. Linder M, Britton T, Sennblad B (2011) Evaluation of Bayesian models of substitution rate evolution—parental guidance versus mutual independence. *Syst Biol* 60:329–342.
19. Ho SYW (2009) An examination of phylogenetic models of substitution rate variation among lineages. *Biol Lett* 5:421–424.
20. Lepage T, Bryant D, Philippe H, Lartillot N (2007) A general comparison of relaxed molecular clock models. *Mol Biol Evol* 24:2669–2680.
21. Graur D, Martin W (2004) Reading the entrails of chickens: Molecular timescales of evolution and the illusion of precision. *Trends Genet* 20:80–86.
22. Hug LA, Roger AJ (2007) The impact of fossils and taxon sampling on ancient molecular dating analyses. *Mol Biol Evol* 24:1889–1897.
23. Hedges SB, Blair JE, Venturi ML, Shoe JL (2004) A molecular timescale of eukaryote evolution and the rise of complex multicellular life. *BMC Evol Biol* 4:2.
24. Adl SM, et al. (2005) The new higher level classification of eukaryotes with emphasis on the taxonomy of protists. *J Eukaryot Microbiol* 52:399–451.
25. Parfrey LW, et al. (2010) Broadly sampled multigene analyses yield a well-resolved eukaryotic tree of life. *Syst Biol* 59:518–533.
26. Hampl V, et al. (2009) Phylogenomic analyses support the monophyly of Excavata and resolve relationships among eukaryotic "supergroups". *Proc Natl Acad Sci USA* 106: 3859–3864.
27. Lane CE, Archibald JM (2008) The eukaryotic tree of life: Endosymbiosis takes its TOL. *Trends Ecol Evol* 23:268–275.
28. Baurain D, et al. (2010) Phylogenomic evidence for separate acquisition of plastids in cryptophytes, haptophytes, and stramenopiles. *Mol Biol Evol* 27:1698–1709.
29. Heath TA, Hedtke SM, Hillis DM (2008) Taxon sampling and the accuracy of phylogenetic analyses. *J Syst Evol* 46:239–257.
30. Yoon HS, Hackett JD, Ciniglia C, Pinto G, Bhattacharya D (2004) A molecular timeline for the origin of photosynthetic eukaryotes. *Mol Biol Evol* 21:809–818.
31. Donoghue PCJ, Benton MJ (2007) Rocks and clocks: Calibrating the Tree of Life using fossils and molecules. *Trends Ecol Evol* 22:424–431.
32. Han TM, Runnegar B (1992) Megascopic eukaryotic algae from the 2.1-billion-year-old Negaunee-Iron-Formation, Michigan. *Science* 257:232–235.
33. Schneider DA, Bickford ME, Cannon WF, Schulz KJ, Hamilton MA (2002) Age of volcanic rocks and syndepositional iron formations, Marquette Range Supergroup: Implications for the tectonic setting of Paleoproterozoic iron formations of the Lake Superior. *Can J Earth Sci* 39:999–1012.
34. Javaux EJ, Marshall CP, Bekker A (2010) Organic-walled microfossils in 3.2-billion-year-old shallow-marine siliciclastic deposits. *Nature* 463:934–938.
35. Buick R (2010) Early life: Ancient acritarchs. *Nature* 463:885–886.
36. Rasmussen B, Fletcher IR, Brocks JJ, Kilburn MR (2008) Reassessing the first appearance of eukaryotes and cyanobacteria. *Nature* 455:1101–1104.
37. Knoll AH (1994) Proterozoic and early Cambrian protists: Evidence for accelerating evolutionary tempo. *Proc Natl Acad Sci USA* 91:6743–6750.
38. Knoll AH (2011) The multiple origins of complex multicellularity. *Annu Rev Earth Planet Sci* 39:217–239.
39. Yin L, Yuan X (2007) Radiation of Meso-Neoproterozoic and early Cambrian protists inferred from the microfossil record of China. *Palaeogeogr Palaeocl* 254:350–361.
40. Porter SM (2006) Heterotrophic Eukaryotes. *Neoproterozoic Geobiology and Paleobiology*, eds Xiao S, Kaufman AJ (Springer, Dordrecht, The Netherlands), pp 1–21.
41. Canfield DE (1998) A new model for Proterozoic ocean chemistry. *Nature* 396: 450–453.
42. Johnston DT, Wolfe-Simon F, Pearson A, Knoll AH (2009) Anoxygenic photosynthesis modulated Proterozoic oxygen and sustained Earth's middle age. *Proc Natl Acad Sci USA* 106:16925–16929.
43. Embley TM, Martin W (2006) Eukaryotic evolution, changes and challenges. *Nature* 440:623–630.
44. Mentel M, Martin W (2008) Energy metabolism among eukaryotic anaerobes in light of Proterozoic ocean chemistry. *Philos Trans R Soc Lond B Biol Sci* 363:2717–2729.
45. Johnston DT, et al. (2010) An emerging picture of Neoproterozoic ocean chemistry: Insights from the Chuar Group, Grand Canyon, USA. *Earth Planet Sci Lett* 290:64–73.
46. Cavalier-Smith T (2009) Megaphylogeny, cell body plans, adaptive zones: Causes and timing of eukaryote basal radiations. *J Eukaryot Microbiol* 56:26–33.
47. Strother PK, Battison L, Brasier MD, Wellman CH (2011) Earth's earliest non-marine eukaryotes. *Nature* 473:505–509.
48. Parnell J, Boyce AJ, Mark D, Bowden S, Spinks S (2010) Early oxygenation of the terrestrial environment during the Mesoproterozoic. *Nature* 468:290–293.
49. Javaux EJ, Knoll AH, Walter MR (2001) Morphological and ecological complexity in early eukaryotic ecosystems. *Nature* 412:66–69.
50. Anbar AD, Knoll AH (2002) Proterozoic ocean chemistry and evolution: A bioinorganic bridge? *Science* 297:1137–1142.
51. Knoll AH, Summons RE, Waldbauer JR, Zumberge J (2007) The geological succession of primary producers in the oceans. *The Evolution of Primary Producers in the Sea*, eds Falkowski PG, Knoll AH (Elsevier, Burlington, MA), pp 133–163.
52. Drummond AJ, Rambaut A (2007) BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evol Biol* 7:214.
53. Lartillot N, Lepage T, Blanquart S (2009) PhyloBayes 3: A Bayesian software package for phylogenetic reconstruction and molecular dating. *Bioinformatics* 25:2286–2288.
54. Smithson TR, Rolfe WDI (1990) *Westlothiana* Gen. nov.—naming the earliest known reptile. *Scott J Geol* 26:137–138.
55. Crane PR, Friis EM, Pedersen KR (1995) The origin and early diversification of angiosperms. *Nature* 374:27–33.
56. Taylor TN, Hass H, Kerp H (1999) The oldest fossil ascomycetes. *Nature* 399:648.
57. Bown PR (1998) *Calcareous Nannofossil Biostratigraphy* (Kluwer Academic, London).
58. Harwood DM, Nikolaev VA, Winter DM (2007) Cretaceous records of diatom evolution, radiation, and expansion. *Paleontol Soc Papers* 13:33–59.
59. Fensome RA, Saldarriaga JF, Taylor F (1999) Dinoflagellate phylogeny revisited: Reconciling morphological and molecular based phylogenies. *Grana* 38:66–80.
60. Rubinstein CV, Gerrienne P, de la Puente GS, Astini RA, Steemans P (2010) Early Middle Ordovician evidence for land plants in Argentina (eastern Gondwana). *New Phytol* 188:365–369.
61. Dostál O, Prokop J (2009) New fossil insects (Diaphanopterodea: Martynoviidae) from the Lower Permian of the Boskovice Basin, southern Moravia. *Geobios* 42:495–502.
62. Friis EM, Pedersen KR, Crane PR (2010) Diversity in obscurity: Fossil flowers and the early history of angiosperms. *Philos Trans R Soc Lond B Biol Sci* 365:369–382.
63. Sun G, Dilcher DL, Wang H, Chen Z (2011) A eudicot from the Early Cretaceous of China. *Nature* 471:625–628.
64. Gray J, Boucot AJ (1989) Is *Moyeria* a euglenoid? *Lethaia* 22:447–456.
65. McIlroy D, Green OR, Brasier MD (2001) Palaeobiology and evolution of the earliest agglutinated Foraminifera: *Platysolenites*, *Spirosolenites* and related forms. *Lethaia* 34:13–29.
66. Kooistra W, Gersonde R, Medlin L, Mann DG (2007) The origin and evolution of the diatoms: Their adaptation to a planktonic existence. *The Evolution of Primary Producers in the Sea*, eds Falkowski PG, Knoll AH (Elsevier, Burlington, MA), pp 201–249.
67. Lipps HJ (1993) *Fossil Prokaryotes and Protists* (Blackwell Scientific, Boston).
68. Kenrick P, Crane PR (1997) The origin and early evolution of plants on land. *Nature* 389:33–39.
69. Shu DG, et al. (1999) Lower Cambrian vertebrates from South China. *Nature* 402: 42–46.
70. Love GD, et al. (2009) Fossil steroids record the appearance of Demospongiae during the Cryogenian period. *Nature* 457:718–721.
71. Cohen PA, Knoll AH, Kodner RB (2009) Large spinose microfossils in Ediacaran rocks as resting stages of early animals. *Proc Natl Acad Sci USA* 106:6519–6524.
72. Martin MW, et al. (2000) Age of Neoproterozoic bilatarian body and trace fossils, White Sea, Russia: Implications for metazoan evolution. *Science* 288:841–845.
73. Butterfield NJ, Knoll AH, Swett K (1994) Paleobiology of the Neoproterozoic Svanbergfjellet Formation, Spitsbergen. *Fossils Strata* 34:1–84.
74. Summons RE, Walter MR (1990) Molecular fossils and microfossils of prokaryotes and protists from Proterozoic sediments. *Am J Sci* 290-A:212–244.
75. Xiao SH, Knoll AH, Yuan XL, Pueschel CM (2004) Phosphatized multicellular algae in the Neoproterozoic Doushantuo Formation, China, and the early evolution of florideophyte red algae. *Am J Bot* 91:214–227.