An uncorrelated relaxed-clock analysis suggests an earlier origin for flowering plants

Stephen A. Smith[a,1], Jeremy M. Beaulieu[b], and Michael J. Donoghue[b,1]

[a]National Evolutionary Synthesis Center, Durham, NC 27705; and [b]Department of Ecology and Evolutionary Biology, Yale University, New Haven, CT 06520

Contributed by Michael J. Donoghue, February 3, 2010 (sent for review November 15, 2009)

We present molecular dating analyses for land plants that incorporate 33 fossil calibrations, permit rates of molecular evolution to be uncorrelated across the tree, and take into account uncertainties in phylogenetic relationships and the fossil record. We attached a prior probability to each fossil-based minimum age, and explored the effects of relying on the first appearance of tricolpate pollen grains as a lower bound for the age of eudicots. Many of our divergence-time estimates for major clades coincide well with both the known fossil record and with previous estimates. However, our estimates for the origin of crown-clade angiosperms, which center on the Late Triassic, are considerably older than the unequivocal fossil record of flowering plants or than the molecular dates presented in recent studies. Nevertheless, we argue that our older estimates should be taken into account in studying the causes and consequences of the angiosperm radiation in relation to other major events, including the diversification of holometabolous insects. Although the methods used here do help to correct for lineage-specific heterogeneity in rates of molecular evolution (associated, for example, with evolutionary shifts in life history), we remain concerned that some such effects (e.g., the early radiation of herbaceous clades within angiosperms) may still be biasing our inferences.

divergence times | rates of molecular evolution | angiosperms | eudicots | land plants

Our understanding of the history of life depends critically on knowledge of the ages of major clades. The timing of land plant evolution is fundamental to the interpretation of earth history and macroevolution throughout the Phanerozoic. Age estimates bear directly on our interpretation of the tempo and mode of morphological and molecular evolution of plants themselves, but also on our interpretation of the evolution of many other groups. For example, the age of origin of the angiosperms has variously been related to the evolution of other plant lineages (e.g., ferns) (1) and biomes (e.g., tropical forests) (2), as well as to the major insect clades and their feeding habits (3–7), and even to the evolution of fungi (8, 9) and dinosaurs (10).

In plants, as in other major groups (e.g., mammals) (11), the ages of clades estimated from molecular phylogenetic analyses have not always corresponded well with the accepted fossil record. In particular, the application of molecular-clock methods has tended to yield older dates (12), in some cases much older than has seemed credible based on the stratigraphic record (13–15). The approaches used in molecular dating have been problematical for several reasons, and in some cases the results have been too easy to dismiss. Early attempts used a strict molecular clock (e.g., ref. 16). Recently, so-called "relaxed clock" methods have been used, which variously allow departures from clock-like behavior (17–20), but here too there are difficulties. First, molecular dating analyses have tended to treat the tree topology as complete and fixed, as opposed to taking into account phylogenetic uncertainties (21–25). For angiosperms this is problematical, considering the limited confidence we still have in the order of branching early in their diversification (24, 26).

Second, the information provided by fossils is often treated as fixed. Uncertainties associated with fossil calibrations are inherent, given the nature of fossilization (27), the difficulty of dating fossil localities, and the standard logic used in placing fossils into phylogenies to obtain minimum ages for lineages (28). One concern is the default practice of assigning fossils to the stem of the most inclusive crown clade to which they probably belong, thereby possibly biasing estimated ages (possibly throughout the tree) to be younger (29). In many cases, estimates have relied on lower bounds, based on what is assumed to be a tightly constrained fossil record. In angiosperms, the origin of eudicots, marked by the appearance of tricolpate pollen in the Late Barremian-Early Aptian (~125 Myr), has widely been used as a hard constraint (i.e., maximum age) to either calibrate or assess angiosperm molecular divergence times (21–25). However, using the first appearance of tricolpate pollen as a fixed calibration may underestimate the origin of eudicots and, by extension, other age estimates that have relied on this constraint. Tricoplate grains first appear in separated geographical areas and the grains themselves are not uniform in morphology (30–33), both observations implying that the tricoplate clade originated some time before its appearance in the fossil record. Although there may be value for some purposes in using the same fossil constraints in different studies, the reliance on the eudicot maximum-age constraint in multiple studies ties them all to the same underlying assumption, thereby compromising their independence.

Third, previous studies have relied on methods that have probably not accommodated sufficiently for heterogeneity in rates of molecular evolution (34). Rarely are datasets found to conform to a molecular clock, and broadly sampled plant phylogenies are no exception (22, 35, 36). Rates of evolution can vary among genes within a lineage, or among lineages. Incorporating data from multiple genes can help to compensate for rate heterogeneity across genes (23), but lineage-specific rate heterogeneity has not been adequately addressed and remains a potentially large source of error. Several analyses have demonstrated striking differences in molecular rates across large plant clades (36–39). These differences have been associated with traits such as growth habit, generation time, and population size, all of which are labile and might change multiple times along a branch, which will be especially difficult to detect when taxonomic sampling is sparse. When not accounted for, such rate heterogeneity can systematically bias slow-rate branches to appear younger and fast-rate branches to appear older (38, 40). Some methods for accommodating nonclock-like rates were designed to smooth differences across branches, under the assumption that rates are autocorrelated (17, 18): that is, they assume the inheritance of rate from parent node to child node. However, factors such as the evolution of life-history characteristics are expected to result in large differences between adjacent nodes, in which case autocorrelated methods are problematical (34). With the development of models of molecular evolution that are un-

Author contributions: S.A.S., J.M.B., and M.J.D. designed research; S.A.S. and J.M.B. performed research; S.A.S. and J.M.B. analyzed data; and S.A.S., J.M.B., and M.J.D. wrote the paper.

The authors declare no conflict of interest.

[1]To whom correspondence may be addressed: sasmith@nescent.org or michael.donoghue@yale.edu.

This article contains supporting information online at www.pnas.org/cgi/content/full/1001225107/DCSupplemental.

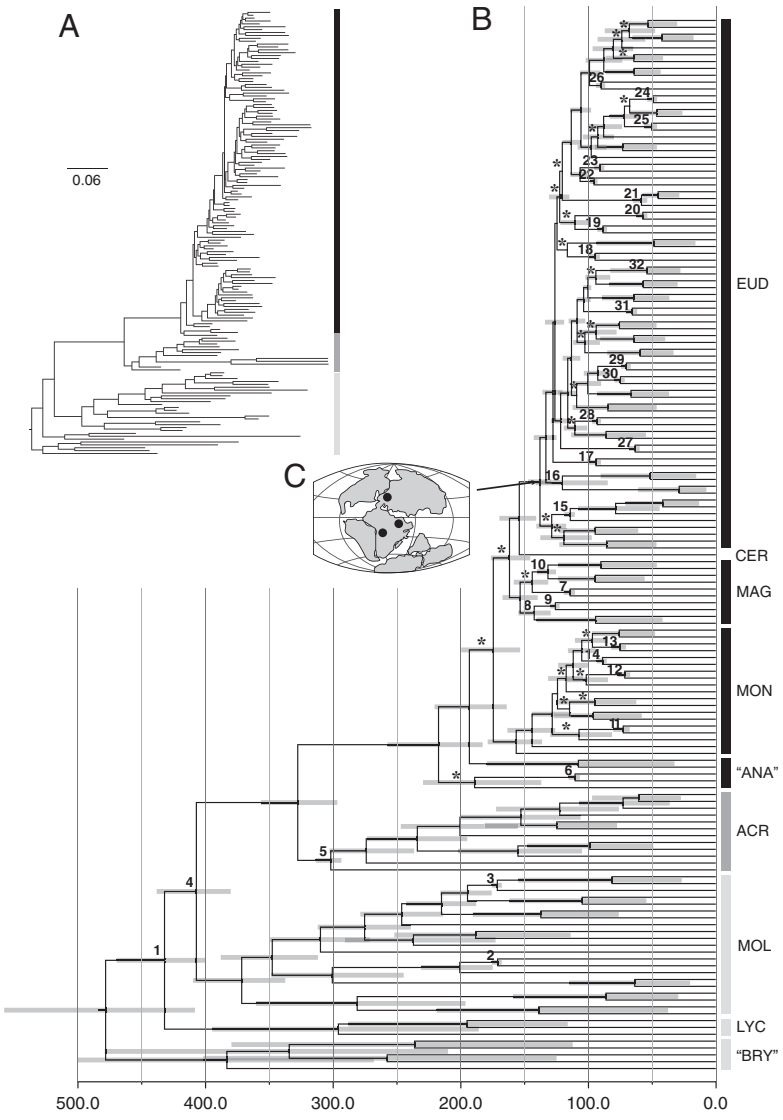

Fig. 1. Phylogenetic tree and divergence time estimates for land plants. (*A*) MrBayes consensus tree from a three-gene (*atp*B, *rbc*L, and 18S) analysis (see Fig S1 for taxon names). Branch lengths represent average substitutions per site. Shaded bars represent *Angiospermae* (black), *Acrogymnospermae* (dark gray), and the rest of the land plants (light gray). (*B*) The maximum clade credibility tree from the divergence time analysis of the same three-gene dataset as in *A*. Studies focused on the root of the land plants (e.g., ref. 81), including outgroups, place the root along the liverwort branch ("bryophytes" paraphyletic). Nodes marked by an asterisk (*) are supported by <0.95 posterior probability. The 95% highest posterior density (HPD) estimates for each well-supported clade are represented by bars. Numbers at nodes correspond to the fossil calibrations in Table S2. (*C*) Map with localities for the first tricolpate pollen records. Clade names follow Cantino et al. (47): ACR, *Acrogymnospermae*; ANA, ANITA grade; BRY, bryophytes; CER, *Ceratophyllum*; EUD, *Eudicotyledonae*; LYC, *Lycopodiophyta*; MAG, *Magnoliidae*; MOL, *Monilophyta*; MON, *Monocotyledonae*.

correlated over a phylogeny (41, 42), we can at least begin to accommodate such rate shifts.

Here, we present a hypothesis of divergence times for land plants, with special emphasis on the timing of the origin of crown-clade angiosperms. To address the issues noted above, we have used the uncorrelated lognormal (UCLN) relaxed-clock model of Drummond et al. (41), which permits the rate of molecular substitution to be uncorrelated across the tree while incorporating uncertainty in both tree topology and multiple fossil calibrations. We also allow fossil calibrations to act as probabilistic priors rather than as point estimates, and explore the possibility of dramatic rate differences associated with life-history evolution. Our results bear on the ages now often assumed in broad comparative analyses in plants (43, 44) and on the possible link between the radiation of plants and other organisms. However, as we discuss below, we remain concerned about the possible effects of rate heterogeneity (especially related to shifts in life history) and highlight the need to develop methods that more explicitly take this into account.

Results and Discussion

Phylogenetic Results. We conducted Bayesian and maximum-likelihood (ML) phylogenetic analyses on 154 species of land plants using previously published sequences of 18S, *atp*B, and *rbc*L (Table S1). These genes were chosen because of their utility in identifying major plant clades (45, 46). Our taxonomic sampling within *Angiospermae* (flowering plants) represents most large clades of *Mesangiospermae* (mesangiosperms, or "core angiosperms"), including *Magnoliidae* (magnoliids), *Monocotyledoneae* (monocots), and *Eudicotyledoneae* (eudicots), all in the sense of Cantino et al. (47), as well as species sampled in previous large-scale tracheophyte analyses (36, 46). To obtain preliminary estimates for the divergence times of crown clades within angiosperms, at least two species were sampled to represent very large clades. Both our ML and Bayesian analyses confirmed previous phylogenetic inferences within angiosperms, with the few exceptions likely because of our full partitioning of the data into gene regions and our somewhat different sample of taxa (see *Materials and Methods* for details). Likewise, relationships among the major clades of land plants largely confirmed previous analyses based on ML nonpartitioned analyses (36, 46). Monophyletic *Spermatophyta* (seed plants), *Acrogymnospermae* (containing the four major lineages of extant "gymnosperms"), and *Angiospermae* are well supported in both our ML and Bayesian analyses. Monocots and eudicots are well supported as monophyletic. In agreement with recent analyses, *Ceratophyllum* is placed sister to eudicots (24), and we see an accelerated rate of molecular evolution in Gnetales (Fig. 1*A* and Figs. S1, S2, S3 and S4) (48).

Table 1. Divergence-time estimates (in Myr) for major clades of land plants as estimated using 33 fossil calibrations (including the eudicot pollen calibration) and 32 fossil calibrations (excluding the eudicot pollen calibration)

Clade	With eudicot calibration	Without eudicot calibration
Land plants	477 (407–557)	474 (419–577)
Tracheophyta	432 (399–469)	434 (404–490)
Spermatophyta	327 (296–356)	330 (301–366)
Acrogymnospermae	301 (293–313)	302 (294–316)
Gnetophyta	155 (104–202)	158 (110–208)
Angiospermae	217 (182–257)	228 (193–270)
Mesangiospermae	174 (153–200)	184 (160–210)
Magnoliidae	155 (136–181)	163 (138–198)
Monocotyledoneae	156 (139–167)	164 (141–191)
Eudicotyledoneae	137 (128–147)	153 (138–172)

Dates within parentheses denote the 95% HPD.

Previous phylogenetic analyses have differed with respect to the relationships of *Amborella* and Nymphaeales to the rest of angiosperms (49) and among the magnoliids, monocots, and eudicots (24, 26). The most comprehensive analysis to date, based on 61 chloroplast genes, supported *Amborella* as sister to the rest of angiosperms but was uncertain regarding magnoliids, eudicots, and monocots (24). We conducted ML analyses on individual genes to help identify differential support for problematical relationships (Figs. S2, S3, and S4). The 18S sequences support Nymphaeales as sister to the rest of angiosperms, and weakly (51% BS) support the nonmonophyly of *Acrogymnospermae*, with cycads sister to angiosperms. *atp*B sequences support the monophyly of *Acrogymnospermae* and place magnoliids as sister to eudicots, but uncertainty remains as to the placement of *Amborella* and Nymphaeales in relation to the rest of the angiosperms. *rbc*L sequences do not clearly resolve the placement of magnoliids or monocots, and weakly support *Amborella* as sister to Nymphaeales (79% BS) and a monophyletic *Acrogymnospermae*. Our divergence time analyses take into account this uncertainty in the underlying tree topology.

Dating Analyses and Results. Our age estimates for divergences within land plants were based on 33 fossil calibrations (Table S2). The origin of land plants centered on 477 Myr (95% HPD: 407–557 Myr) (Table 1), which corresponds well with the earliest known occurrence of microfossils assigned to land plants from the middle Ordovician (~470 Myr) (50). The origin of tracheophytes was estimated at 432 Myr (95% HPD: 399–469 Mya), which corresponds to the middle Silurian. The first fossil fragments widely assigned to early *Tracheophyta* are also of Silurian age (~419) (51). We estimated the origin of *Spermatophyta* (crown seed plants) in the Middle Carboniferous (327 Myr; 95% HPD: 296–356 Myr), which broadly corresponds with the fossil record (52). A Middle Carboniferous age for crown seed plants follows the Devonian (>360 Myr) evidence of the "progymnosperm" lineages, Archaeopteridales and Aneurophytales (53, 54), and Early Carboniferous evidence of Paleozoic seed ferns (55). *Acrogymnospermae* (301 Myr, 95% HPD: 293–313) are estimated to have originated some 30 million years after crown seed plants.

We estimated the origin of crown angiosperms to be 217 Myr (95% HPD: 182–257 Myr), in the Late Triassic. This result was robust to the inclusion or exclusion of a 125 Myr minimum-age calibration on the node corresponding to crown eudicots (Table 1). A Late Triassic origin for crown angiosperms is typically not estimated with molecular methods. Of the several analyses carried out by Sanderson and Doyle (35), a few estimated a Late Triassic origin of crown angiosperms. However, this was sensitive to the underlying tree topology, codon position, and taxon sampling, with a majority of estimates falling instead between 140 and 190 Myr. Most molecular divergence-time analyses for crown angiosperms have reported dates within this range of 140 to 190 Myr (reviewed by 34). For example, Bell et al. (22) and Magallón and Sanderson (23) estimated crown angiosperms to be 140 to 180 Myr (with Bayesian relaxed-clock methods) and 163 to 189 Myr (using penalized likelihood), respectively. More recently, Moore et al. (24) estimated the origin of crown angiosperms to be ~170 Myr using a data set of 61 plastid genes. Magallón and Castillo (25) reported an age of 130 Myr; however, this depended on a maximum age constraint applied to the crown corresponding to the oldest putative fossil angiosperm pollen from the Valanginian to Hauterivian (~130–140 Myr) (33).

Our age estimates for nodes within the *Mesangiospermae* suggest that crown magnoliids, monocots, and eudicots had all originated by the Late Jurassic (Table 1). Although the relationships among these three lineages remains uncertain, the short time interval between the origin of the corresponding crown groups suggests a rapid succession in the origin of the major angiosperms lineages. Although a Late Jurassic origin for magnoliids, monocots, and eudicots is much older than previous analyses have reported, it is important to note that we cannot reject an Early Cretaceous origin (the lower 95% HPD <144 Myr) (Table 1). When we removed the minimum age calibration for crown eudicots at 125 Myr, the estimates for the origin of crown magnoliids and monocots remained centered on the Late Jurassic, again with the lower 95% credibility interval encompassing the Early Cretaceous (Table 1). Taken together, our age estimates for the origin of major crown groups within mesangiosperms clearly predate the first putative angiosperm fossils (~130–140 Myr) (33).

Rate Heterogeneity. One possible explanation for the lack of correspondence between our molecular divergence estimate and the accepted fossil record is that we have failed to properly account for rate heterogeneity. Two metrics were used to evaluate the appropriateness of assuming a model of uncorrelated rates of molecular variation and to assess overall rate heterogeneity across the tree. First, the degree of autocorrelation of molecular rates from parent to child throughout the phylogeny was estimated through the covariance parameter. Across land plants, we estimated a rate covariance of $\rho = 0.074$, which was not significantly different from zero (95% HPD: −0.038 to 0.176). Although this result does not support the autocorrelation of rates, we note that the UCLN may be inadequate in detecting significant rate autocorrelation even when it exists (56). Second, we examined the coefficient of variation to assess the overall degree of rate heterogeneity across our tree. Assuming uncorrelated substitution rates, the coefficient of variation is the variance of rates scaled by the associated mean. A significant portion of the posterior density centered near zero is evidence that the data are "clock-like," whereas a posterior density that does not encompass zero provides evidence for significant rate heterogeneity. The estimated

coefficient of variation indicated that the rate of molecular substitution varied by 69.7% (95% HPD: 63.2–76.3%) across the entire tree. This high degree of rate variation suggests an influence of rate heterogeneity.

It has been noted that nonsynonymous sites may show less rate heterogeneity than synonymous sites (37, 39, 57, 58). Different rates of evolution associated with different life histories may be more evident in synonymous sites. To examine whether the UCLN model had sufficiently accommodated for rate heterogeneity, we estimated divergence times using only the first and second codon positions for *atp*B and *rbc*L, which focused rate and date estimates on nonsynonymous sites. In this analysis, only the divergence time for land plants changed significantly, with the coding data supporting a slightly younger date of 442 Myr (95% HPD: 401–514 Myr) (Table S3). Importantly, the coefficient of variation indicated that significant rate heterogeneity remained, as the substitution rate varied 67.6% from the mean (95% HPD: 59.5–75.8%). Thus, although using only first and second sites may remove some bias caused by rate heterogeneity, it certainly does not eliminate it (39), and this is presumably true not just for the UCLN method but for methods that assume autocorrelated rates.

To explore whether the evolution of different life histories might help to explain shifts in the rate of molecular evolution (38, 39, 58), and specifically to test whether a shift to the herbaceous habit at the base of the angiosperms (59) could explain the long branch subtending crown angiosperms, we reconstructed growth habit over the posterior distribution of dated trees (Fig. 1*B*). We found an 88.6% and a 97.4% posterior probability of "woody" being the ancestral state for crown angiosperms and crown seed plants, respectively. This finding supports the results of Feild et al. (60, 61), who also argued on the basis of such reconstructions coupled with physiological data that the first angiosperms were woody plants living in "dark and disturbed" environments. Here it is relevant that fossil lineages inferred to be along the stem subtending the angiosperms, such as the glossopterids, Caytoniales, and Bennettitales (62, 63; but see ref. 64), are considered to be woody. These observations work against a simple argument that a shift to herbaceous habit along the line leading to angiosperms resulted in a faster rate of molecular evolution and, hence, to the inference that crown angiosperms are much older than they actually are. However, it is important to appreciate that the length of the branch leading to crown angiosperms spans ~100 Myr, which would allow many potentially confounding and undetected changes in life history and population size. Of course, the long branch subtending the angiosperms may not entirely reflect a faster rate of molecular evolution; it may simply indicate high extinction along that branch. We return to these concerns below.

Use of First Occurrence of Tricolpate Pollen. Tricolpate pollen grains (and derivative conditions) characterize modern eudicots, and the appearance of such grains in the fossil record is taken as evidence of the existence of the eudicot lineage, if not the crown clade. The first appearance of tricolpate pollen grains at the Barremian-Aptian boundary (~125 Myr) has commonly been used as a fossil calibration in divergence-time analyses. Whether placed along the stem or at the crown of the eudicots, this tricolpate pollen calibration has primarily been treated as a maximum-age constraint (22–25, 36, 65). This treatment could be justified based on the observation that the pollen record is substantial and that tricolpate grains, which are easy to identify, have not been recovered from any earlier sediments. It may be possible in this case to infer a likely maximum age using the statistical methods proposed by Marshall (66). In the meantime, several aspects of the appearance of tricolpate grains in the fossil record suggest that 125 Myr may not be an appropriate maximum age for eudicots. The first tricolpate grains have the same aperture configuration (tricolpate, not tricolporoidate, or triporate), but show "considerable structural variety" (33) in the sculpturing of the exine. In addition, the Barremian-Aptian tricolpate pollen localities are geographically widespread, first at several Gondwanan sites (present-day northern and equatorial Africa), with specimens becoming more common in Laurasia during the Aptian-Albian (Fig. 1*C*) (30, 31). Based on these observations, it is possible that the appearance of tricolpate grains reflects the rise to dominance of the eudicot lineage as opposed to the origin of the group. Finally, it is not clear whether these pollen grains represent the emergence of the tricolpate apomorphy along the branch leading to crown eudicots or whether they represent the appearance of modern lineages of eudicots (i.e., within the crown). There are too few characters to place them with any certainty within the eudicot phylogeny (33). In view of these caveats, we favor the use of 125 Myr as a minimum age for the origin of the eudicot crown clade, with an associated probabilistic prior (67).

Reexamining Biological Patterns. Of particular interest for botanists and entomologists is the possible correlation of the early evolution of angiosperms with the rise of the major lineages of holometabolous insects (Coleoptera, Hymenoptera, Diptera, and Lepidoptera). Labandeira and Sepkoski (3) noted that a number of major insect radiations date to the late Permian (~254 Myr), with trophic diversity proliferating dramatically during this period. Based on the apparent incongruity with the angiosperm fossil record, they concluded that angiosperms had little impact on the early evolution of holometabolous insects.

Our results, taken at face value, push the origin of the angiosperm crown clade much closer in time to the diversification of the major lineages of holometabolous insects. Molecular estimates for the origin of Coleoptera (285 Myr) (5) predate crown angiosperms, but the origin of the most diverse herbivorous lineages of Coleoptera (i.e., Chrysomeloidea, Curculionoidea) are estimated to be ~230 Myr, which is in the range of our age estimate for crown angiosperms (68). Also congruent are Triassic fossils of Diptera and Hymenoptera (68, 69). Molecular age estimates for the origin of Lepidoptera, as well as for the ant and bee clades nested within Hymnoptera, correspond well with our age estimates for the major crown clades within *Mesangiospermae* (7, 70, 71), as does the fossil record of long-proboscid Mecoptera (72). However, it is important to note that even if our inferred dates were correct, the absence of clear-cut angiosperm fossils during the Triassic and Jurassic may signify that the first angiosperms were not abundant, widespread, or ecologically very significant, in which case it would be difficult to argue that the appearance of the angiosperms dramatically increased insect diversity during that time period.

Conclusions

Regarding the tempo of plant evolution, our results show generally good correspondence with the fossil record (e.g., for crown tracheophytes and seed plants). However, they also imply that crown angiosperms originated in the Triassic (or possibly in the Jurassic), well before the Cretaceous radiations that were responsible for the dramatic rise of the angiosperms. That is, they suggest that crown angiosperms were in existence for some 50 Myr (or more) before the radiation of the mesangiosperms, and some 60 Myr (or more) before the diversification of monocots and eudicots. The only living remnants of the lineages that existed in this inferred interval are *Amborella*, Nymphaeales/Hydatellales, and Austrobaileyales. Today, these lineages are species-poor, but they exhibit tremendous morphological and ecological disparity. One possibility is that these groups were once much more diverse, and that we are left today with only a few survivors. In this case, it may be that most of these plants lived in environments that were not conducive to fossilization. However, it is also possible that angiosperms were simply not diverse or ecologically dominant plants during the Late Triassic and Jurassic. For example, as Feild et al. (60, 61) have argued, the physiology and ecological prefer-

ences of the early angiosperms (living in dark, wet, and disturbed understory habitats, probably with low population sizes) may have restricted their abundance, geographical spread, and diversification. These same factors might also account for the lack of fossils during this interval.

Regarding eudicots, our results suggest that the first appearance of tricolpate grains at ca. 125 Myr underestimates the origin of the tricolpate clade by perhaps 3 to 22 Myr. This finding is problematic because the record of fossil pollen is judged to be very good through this time period (33). If our inferences are correct, the appearance of tricolpate grains may not signal the origin of the crown group, but rather the rise in abundance and geographic expansion of the tricolpate lineage.

Dismissing our angiosperm date as an artifact will be tempting. However, as the date reflects the current state of knowledge of fossils and phylogeny, as well as the current state of development of relevant analytical tools, we believe that these dates should not be set aside lightly. Yet, we hasten to acknowledge that our analysis is unlikely to be the final word on the subject and, moreover, there are several reasons to proceed cautiously. Perhaps most importantly, we remain concerned about the impacts of lineage-specific rate heterogeneity on molecular age estimates, despite having tried to accommodate this. It is increasingly clear that there may be extreme differences in molecular rate depending upon life history and other factors (39), and current methods may be unable to cope. It is possible that the effects of lineage-specific rate heterogeneity can "trickle-down" to nodes at some distance from the inferred shift in life history and molecular rate. For this reason we are concerned that, although a shift to herbaceousness may not have marked the origin on the angiosperms, multiple shifts to the herbaceous habit not far within angiosperm, followed by several rapid radiations, might result in an older age estimate for angiosperms as a whole. This possibility needs to be explored further using simulations and also suggests the need to develop methods (akin, perhaps, to so-called local-clock methods) (20) that allow shifts to different rate categories as a function of evolutionary shifts in an underlying parameter that might drive rate changes, such as life history or population size.

It is interesting to reflect, however, that as older fossils are discovered and incorporated into various lineages, this will tend to shift the angiosperm date back further in time. Furthermore, as our taxonomic sampling improves, and as knowledge of fossils increases to the point of allowing us to place them within clades with greater precision, there may be a general tendency for these to be placed further up within the clades with which they are associated. This process will also tend to push the age of crown angiosperms further back in time. It is possible that a closer match between molecular inferences and the stratigraphic record will eventually be obtained, as dating methods are improved to cope with extreme rate heterogeneity and as older fossils are discovered. However, it is also possible that a significant gap will remain and, if so, this might tell us something important about the rise of flowering plants.

Materials and Methods

Phylogenetic Analyses. Sequences for each gene region were aligned separately with MUSCLE (73) within three partitions: angiosperms, acrogymnosperms, and the remaining green plants. These separate alignments were then combined using profile alignment techniques (73; see also ref. 74), and the aligned gene regions were concatenated using Phyutility (75).

Maximum-likelihood analyses were conducted with RAxML (Ver. 7.0.1) (76). Runs were partitioned into gene regions with parameters unlinked. We used the GTRMIX substitution and rate heterogeneity model. ML analyses were conducted by first running 100 rapid bootstrap analyses; every tenth bootstrap tree was used as a start tree for a full ML search. The best tree from those searches was considered to be the ML tree. Bootstraps were summarized with Phyutility (75). We conducted these analyses on the concatenated dataset as well as on the individual gene regions (Figs S1, S2, S3, and S4).

Bayesian phylogenetic analyses were conducted with MrBayes (Ver. 3.1.2) (77, 78) using the Metropolis coupled Markov Chain Monte Carlo (MCMC) algorithm. Two analyses, each consisting of four incrementally heated chains, were run for 10 million generations, sampling every 1,000th tree. The posterior distribution of trees was summarized after removing 1 million generations as burn-in. A GTR+Γ model was applied to each gene region and the associated parameters were unlinked. Posterior distributions for parameter estimates and likelihood scores were visualized in Tracer (Ver. 1.4) to approximate convergence.

Divergence-Time Estimation. Simultaneous divergence-time and phylogenetic analyses were conducted using MCMC methods implemented in BEAST (Ver. 1.4.7; 42). BEAST employs an uncorrelated relaxed-clock (UCLN) model to estimate divergence times and allows topologies to be considered "fixed" or estimated to accommodate for phylogenetic uncertainty (41, 42). Here, we allowed BEAST to estimate the topology. For each branch, the UCLN independently draws substitution rates from a lognormal distribution, allowing substitution rates to be uncorrelated across the phylogeny. The absolute estimates of divergence times are then calculated from fossil calibrations, each with an associated probabilistic prior. We attached a lognormal prior probability (67) to the minimum-age estimates obtained from 33 fossil calibrations (using the International Commission on Stratigraphy 2007), 27 of these for crown clades within angiosperms (Table S2). The mean and standard distributions of these calibrations were chosen to acknowledge that, although the fossil age represents the minimum age of the lineage, there remains a probability that the true age extends (in most cases ~10–15 Myr) further back in time.

Our divergence-time analyses were carried out using two partitioning strategies. The first partitioned the data by gene region (*atp*B, *rbc*L, and 18S), with the rate parameters unlinked and assuming a GTR+Γ substitution model. The second partitioned the first and second codon positions of *atp*B and *rbc*L only. Again, the parameters were unlinked and we assumed a GTR+Γ substitution model. For each partitioning strategy, we initiated five separate MCMC chains, each consisting of 10 to 50 million generations with convergence monitored by Tracer (Ver. 1.4). We determined the number of runs to conduct based on the effective samples sizes of each estimated parameter, where we required the posterior, prior, and likelihood to be at least 200. We heuristically removed a percentage of each run as burn-in and the resulting trees for each replicate were combined. Trees were summarized with TreeAnnotator and represent the maximum clade credibility tree. Ninty-five percent HPD were estimated using the R package (79) Bayesm (Ver. 2.2–1).

Ancestral Life-History Reconstructions. We used the Bayesian implementation of the program MultiState in BayesTraits (80) to reconstruct the probable life-history of crown seed plants and crown angiosperms across the posterior distribution of dated trees. Multistate implements a reversible-jump MCMC procedure for single multistate characters. Predominantly herbaceous clades were scored as 0, and predominantly woody clades as 1; any clade for which ancestral life form was judged to be equivocal was scored as missing data. We used an exponential hyperprior on the rate coefficients and sampled every 1,000th point from 10 million total generations. We discarded the first 2.5 million iterations as burn-in. The probability distributions obtained from the reversible-jump MCMC were examined using Tracer (Ver. 1.4).

ACKNOWLEDGMENTS. We thank J. Oliver and D. Tank for helpful suggestions on the manuscript, J. Doyle for discussions about fossils, P. Crane and S. Mathews for helpful reviews, and S. Magallón, C. Bell, D. Soltis, and P. Soltis, for kindly sharing the results of their parallel studies of angiosperm diversification. S.A.S. was partially supported by the National Evolutionary Synthesis Center (NSF EF-0905606 and NSF EF-0423641). M.J.D. and J.M.B. have been supported through a National Science Foundation Angiosperm "Tree of Life" (ATOL) award (NSF EF-0431258).

1. Schneider H, et al. (2004) Ferns diversified in the shadow of angiosperms. *Nature* 428: 553–557.
2. Davis CC, Webb CO, Wurdack KJ, Jaramillo CA, Donoghue MJ (2005) Explosive radiation of Malpighiales supports a mid-cretaceous origin of modern tropical rain forests. *Am Nat* 165:E36–E65.
3. Labandeira CC, Sepkoski JJ, Jr (1993) Insect diversity in the fossil record. *Science* 261: 310–315.
4. Grimaldi D (1999) The co-radiations of pollinating insects and angiosperms in the Cretaceous. *Ann Miss Bot Gard* 86:373–406.
5. Hunt T, et al. (2007) A comprehensive phylogeny of beetles reveals the evolutionary origins of a superradiation. *Science* 318:1913–1916.

6. McKenna DD, Farrell BD (2006) Tropical forests are both evolutionary cradles and museums of leaf beetle diversity. *Proc Natl Acad Sci USA* 103:10947–10951.
7. Moreau CS, Bell CD, Vila R, Archibald SB, Pierce NE (2006) Phylogeny of the ants: diversification in the age of angiosperms. *Science* 312:101–104.
8. Matheny PB, et al. (2009) Out of the Palaeotropics? Historical biogeography and diversification of the cosmopolitan ectomycoorhizal mushroom family Inocybaceae. *J Biogeogr* 36:577–592.
9. Hibbett DS, Matheny PB (2009) The relative ages of ectomycorrhizal mushrooms and their plant hosts estimated using Bayesian relaxed molecular clock analyses. *BMC Biol* 7:13.
10. Barrett PM, Willis KJ (2001) Did dinosaurs invent flowers? Dinosaur-angiosperm coevolution revisited. *Biol Rev Camb Philos Soc* 76:411–447.
11. Tavaré S, Marshall CR, Will O, Soligo C, Martin RD (2002) Using the fossil record to estimate the age of the last common ancestor of extant primates. *Nature* 416: 726–729.
12. Bromham L, Penny D (2003) The modern molecular clock. *Nat Rev Genet* 4:216–224.
13. Crane PR, Friis EM, Pedersen KR (1995) The origin and early diversification of angiosperms. *Nature* 374:27–33.
14. Sanderson MJ (2003) Molecular data from 27 proteins do not support a Precambrian origin of land plants. *Am J Bot* 90:954–956.
15. Foote M, Hunter JP, Janis CM, Sepkoski JJ, Jr (1999) Evolutionary and preservational constraints on origins of biologic groups: divergence times of eutherian mammals. *Science* 283:1310–1314.
16. Heckman DS, et al. (2001) Molecular evidence for the early colonization of land by fungi and plants. *Science* 293:1129–1133.
17. Sanderson MJ (1997) A nonparametric approach to estimating divergence times in the absence of rate constancy. *Mol Biol Evol* 14:1218–1231.
18. Sanderson MJ (2002) Estimating absolute rates of molecular evolution and divergence times: a penalized likelihood approach. *Mol Biol Evol* 19:101–109.
19. Thorne JL, Kishino H (2002) Divergence time and evolutionary rate estimation with multilocus data. *Syst Biol* 51:689–702.
20. Yang Z (2007) PAML 4: phylogenetic analysis by maximum likelihood. *Mol Biol Evol* 24:1586–1591.
21. Anderson CL, Bremer K, Friis EM (2005) Dating phylogenetically basal eudicots using *rbcL* sequences and multiple fossil reference points. *Am J Bot* 92:1737–1748.
22. Bell CD, Soltis DE, Soltis PS (2005) The age of the angiosperms: a molecular timescale without a clock. *Evolution* 59:1245–1258.
23. Magallón SA, Sanderson MJ (2005) Angiosperm divergence times: the effect of genes, codon positions, and time constraints. *Evolution* 59:1653–1670.
24. Moore MJ, Bell CD, Soltis PS, Soltis DE (2007) Using plastid genome-scale data to resolve enigmatic relationships among basal angiosperms. *Proc Natl Acad Sci USA* 104:19363–19368.
25. Magallón SA, Castillo A (2009) Angiosperm diversification through time. *Am J Bot* 96: 349–365.
26. Soltis DE, Gitzendanner MA, Soltis PS (2007) A 567-taxon data set for angiosperms: the challenges posed by Bayesian analyses of large data sets. *Int J Plant Sci* 168: 137–157.
27. Marshall CR (1998) in *The Adequacy of the Fossil Record*, eds Donovan SK, Paul CRC (Wiley, New York), pp 23–53.
28. Doyle JA, Donoghue MJ (1993) Phylogenies and angiosperm diversification. *Paleobiology* 19:141–167.
29. Bell CD, Donoghue MJ (2005) Dating the Dipsacales: comparing models, genes, and evolutionary implications. *Am J Bot* 92:284–296.
30. Hughes NF, McDougall AB (1990) Barremian-Aptian angiospermid pollen records from southern England. *Rev Palaeobot Palynol* 65:145–151.
31. Doyle JA (1992) Revised palynological correlations of the lower Potomac Group (USA) and the Cocobeach sequence of Gabon (Barremian-Aptian). *Cretaceous Res* 13: 337–349.
32. Brenner GJ (1996) in *Flowering Plant Origin, Evolution, and Phylogeny*, eds Taylor DW, Hickey LJ (Chapman & Hall, New York), pp 91–115.
33. Friis EM, Pedersen KR, Crane PR (2006) Cretaceous angiosperm flowers: innovation and evolution in plant reproduction. *Palaeogeogr Palaeoclimatol Palaeoecol* 232: 251–293.
34. Sanderson MJ, Thorne JL, Wikstrom N, Bremer K (2004) Molecular evidence on plant divergence times. *Am J Bot* 91:1656–1665.
35. Sanderson MJ, Doyle JA (2001) Sources of error and confidence intervals in estimating the age of angiosperms from *rbcL* and 18s rDNA data. *Am J Bot* 88:1499–1516.
36. Soltis PS, Soltis DE, Savolainen V, Crane PR, Barraclough TG (2002) Rate heterogeneity among lineages of tracheophytes: integration of molecular and fossil data and evidence for molecular living fossils. *Proc Natl Acad Sci USA* 99:4430–4435.
37. Gaut BS, Muse SV, Clark WD, Clegg MT (1992) Relative rates of nucleotide substitution at the *rbcL* locus of monocotyledonous plants. *J Mol Evol* 35:292–303.
38. Kay KM, Whittall JB, Hodges SA (2006) A survey of nuclear ribosomal internal transcribed spacer substitution rates across angiosperms: an approximate molecular clock with life history effects. *BMC Evol Biol* 6:36.
39. Smith SA, Donoghue MJ (2008) Rates of molecular evolution are linked to life history in flowering plants. *Science* 322:86–89.
40. Britton T, Anderson CL, Jacquet D, Lundqvist S, Bremer K (2007) Estimating divergence times in large phylogenetic trees. *Syst Biol* 56:741–752.
41. Drummond AJ, Ho SYW, Phillips MJ, Rambaut A (2006) Relaxed phylogenetics and dating with confidence. *PLoS Biol* 4:e88.
42. Drummond AJ, Rambaut A (2007) BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evol Biol* 7:214.
43. Moles AT, et al. (2005) Factors that shape seed mass evolution. *Proc Natl Acad Sci USA* 102:10540–10544.
44. Beaulieu JM, et al. (2007) Correlated evolution of genome size and seed mass. *New Phytol* 173:422–437.
45. Soltis PS, Soltis DE, Chase MW (1999) Angiosperm phylogeny inferred from multiple genes as a tool for comparative biology. *Nature* 402:402–404.
46. Pryer KM, et al. (2001) Horsetails and ferns are a monophyletic group and the closest living relatives to seed plants. *Nature* 409:618–622.
47. Cantino PD, et al. (2007) Towards a phylogenetic nomenclature of *Tracheophyta*. *Taxon* 56:822–846.
48. Burleigh JG, Mathews S (2004) Phylogenetic signal in nucleotide data from seed plants: implications for resolving the seed plant tree of life. *Am J Bot* 91:1599–1613.
49. Zanis MJ, Soltis DE, Soltis PS, Mathews S, Donoghue MJ (2002) The root of the angiosperms revisited. *Proc Natl Acad Sci USA* 99:6848–6853.
50. Wellman CH, Gray J (2000) The microfossil record of early land plants. *Philos Trans R Soc Lond B Biol Sci* 355:717–731, discussion 731–732.
51. Kenrick PK, Crane PR (1997) *The Origin and Early Diversification of Land Plants* (Smithsonian Institution Press, Washington, London).
52. Doyle JA (1998) Molecules, morphology, fossils, and the relationship of angiosperms and Gnetales. *Mol Phylogenet Evol* 9:448–462.
53. Rothwell GW, Scheckler SE, Gillespie WH (1989) Elkinsia gen. nov., a late Devonian gymnosperm with cupulate ovules. *Bot Gaz* 150:170–189.
54. Gerrienne P, Meyer-Berthaud B, Fairon-Demaret M, Streel M, Steemans P (2004) Runcaria, a middle devonian seed plant precursor. *Science* 306:856–858.
55. Dunn MT (2004) The Fayetteville flora I: Upper Mississippian (middle Chesterian/lower Namurian A) plant assemblages of permineralized and compression remains from Arkansas, USA. *Rev Palaeobot Palynol* 132:79–102.
56. Moore BR, Donoghue MJ (2007) Correlates of diversification in the plant clade Dipsacales: geographic movement and evolutionary innovations. *Am Nat* 170(Suppl 2):S28–S55.
57. Bousquet J, Strauss SH, Doerksen AH, Price RA (1992) Extensive variation in evolutionary rates of *rbcL* gene sequence among seed plants. *Proc Natl Acad Sci USA* 89:7844–7848.
58. Gaut BS, Morton BR, McCaig BC, Clegg MT (1996) Substitution rate comparisons between grasses and palms: synonymous rate differences at the nuclear gene Adh parallel rate differences at the plastid gene *rbcL*. *Proc Natl Acad Sci USA* 93: 10274–10279.
59. Taylor DW, Hickey LJ (1992) Phylogenetic evidence for the herbaceous origin of angiosperms. *Plant Syst Evol* 180:137–156.
60. Feild TS, Arens NC, Doyle JA, Dawson TE, Donoghue MJ (2004) Dark and disturbed: a new image of early angiosperm ecology. *Paleobiology* 30:82–107.
61. Feild TS, Chatelet DS, Brodribb TJ (2009) Ancestral xerophobia: a hypothesis on the whole plant ecophysiology of early angiosperms. *Geobiology* 7:237–264.
62. Doyle JA (2006) Seed ferns and the origin of angiosperms. *J Torrey Bot Soc* 133: 169–209.
63. Doyle JA (2008) Integrating molecular phylogenetic and paleobotanical evidence on origin of the flower. *Int J Plant Sci* 169:816–843.
64. Friis EM, et al. (2007) Phase-contrast X-ray microtomography links Cretaceous seeds with Gnetales and Bennettitales. *Nature* 450:549–552.
65. Wang H, et al. (2009) Rosid radiation and the rapid rise of angiosperm-dominated forests. *Proc Natl Acad Sci USA* 106:3853–3858.
66. Marshall CR (1997) Confidence intervals on stratigraphic ranges with nonrandom distributions of fossil horizons. *Paleobiology* 23:165–173.
67. Ho SYW, Phillips MJ (2009) Accounting for calibration uncertainty in phylogenetic estimation of evolutionary divergence times. *Syst Biol* 58:367–380.
68. Grimaldi D, Engel MS (2005) *Evolution of the Insects* (Cambridge University Press, New York).
69. Reik EF (1955) Fossil insects from the Triassic beds at Mt. Crosby, Queensland. *Aust J Zool* 3:654–691.
70. Poinar GO, Jr, Danforth BN (2006) A fossil bee from Early Cretaceous Burmese amber. *Science* 314:614.
71. Danforth BN, Brady SG, Sipes SD, Pearson A (2004) Single-copy nuclear genes recover cretaceous-age divergences in bees. *Syst Biol* 53:309–326.
72. Ren D, et al. (2009) A probable pollination mode before angiosperms: Eurasian, long-proboscid scorpionflies. *Science* 326:840–847.
73. Edgar RC (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res* 32:1792–1797.
74. Smith SA, Beaulieu JM, Donoghue MJ (2009) Mega-phylogeny approach for comparative biology: an alternative to supertree and supermatrix approaches. *BMC Evol Biol* 9:37.
75. Smith SA, Dunn CW (2008) Phyutility: a phyloinformatics tool for trees, alignments and molecular data. *Bioinformatics* 24:715–716.
76. Stamatakis A (2006) RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics* 22:2688–2690.
77. Huelsenbeck JP, Ronquist F (2001) MRBAYES: Bayesian inference of phylogenetic trees. *Bioinformatics* 17:754–755.
78. Ronquist F, Huelsenbeck JP (2003) MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* 19:1572–1574.
79. R Development Core Team (2009) R: A language and environment for statistical computing. (R Foundation for Statistical Computing, Vienna, Austria) http://www.R-project.org. Accessed March 2, 2010.
80. Pagel M, Meade A (2006) Bayesian analysis of correlated evolution of discrete characters by reversible-jump Markov chain Monte Carlo. *Am Nat* 167:808–825.
81. Qiu Y, et al. (2006) The deepest divergences in land plants inferred from phylogenomic evidence. *Proc Natl Acad Sci USA* 103:15511–15516.

Supporting Information

Smith et al. 10.1073/pnas.1001225107

Fig. S1. The maximum-likelihood gene tree for three genes (*18S*, *atpB*, and *rbcL*) with identical topology to the tree shown in Fig 1*A*. Note that studies focused on the root of the land plants, including outgroups, place the root along the liverwort branch (Marchantia+Haplomitrium; "bryophytes" paraphyletic).

Fig. S2. Maximum-likelihood gene tree for 18S.

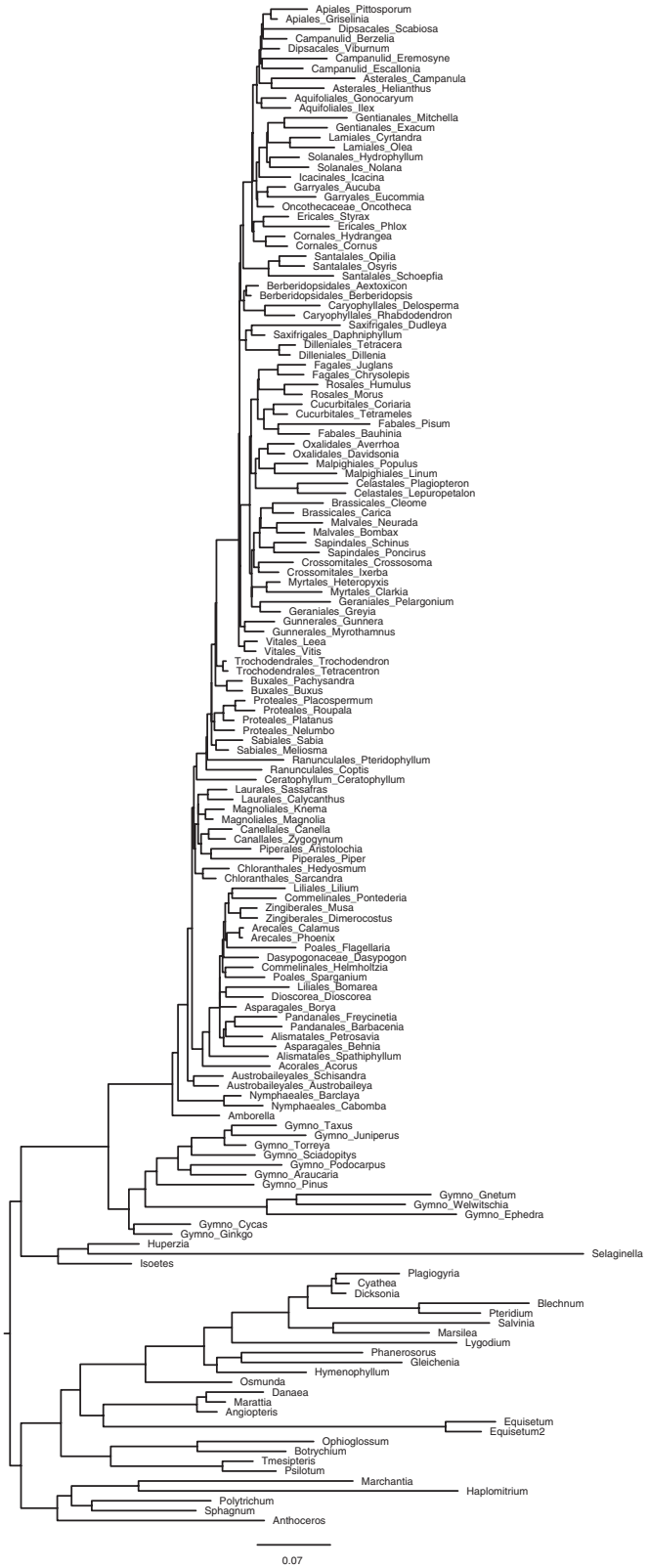

Fig. S3. Maximum-likelihood gene tree for atpB.

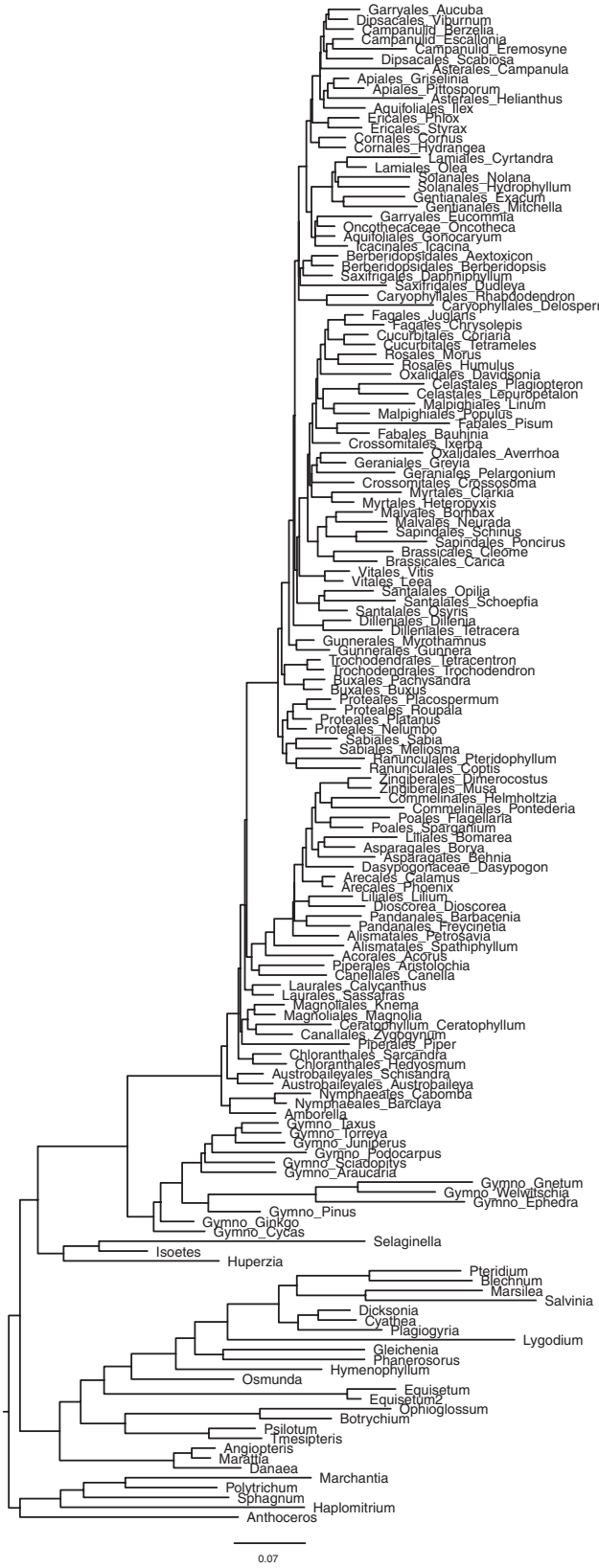

Fig. S4. Maximum-likelihood gene tree for rbcL.

Table S1. List of taxa included in the phylogenetic analysis

Taxon	18S	*atpB*	*rbcL*
Bryophyta			
Anthoceros	X80984	D43695	D43695
Haplomitrium	U18504	AF313555	U87071
Marchantia	X75521	X04465	X04465
Polytrichum	X80982	AF313556	U87087
Sphagnum	Y11370	AF313557	L13485
Lycopodiophyta			
Huperzia	AF313567	U93819	Y07934
Isoetes	AF313577	AF313544	L11054
Selaginella	AF313575	AF313554	AJ010854
Monilophyta			
Angiopteris	D85301	AF313546	X58429
Blechnum	AF313570	U93838	U05909
Botrychium	AF313566	U93825	L13474
Cyathea	AF313574	AF313553	AF313585
Danaea	AF313561	AF313543	AF313578
Dicksonia	U18624	U93829	U05919
Equisetum	AF313576	AF313541	AF313579
Equisetum2	AF313562	AF313542	AF313580
Gleichenia	AF313572	AF313550	AF313584
Hymenophyllum	AF313559	AF313538	AF275645
Lygodium	AB001538	AF313549	L13479
Marattia	AF313564	AF313540	AF313581
Marsilea	AF313573	AF313551	L13480
Ophioglossum	AF313565	U93826	AF313582
Osmunda	AF313560	AF313539	D14882
Phanerosorus	AF313571	AF313548	AF313583
Plagiogyria	AF313568	AF313547	U05643
Psilotum	X81963	U93822	L11059
Pteridium	AF313569	U93835	U05939
Salvinia	X90413	AF313552	U05649
Tmesipteris	AF313563	AF313545	U30836
Acrogymnospermae			
Gymno_Araucaria	D38240	EF490503	EF490510
Gymno_Cycas	D85297	AF313558	L12674
Gymno_Ephedra	D38242	AF239779	D10732
Gymno_Ginkgo	D16448	AJ235481	D10733
Gymno_Gnetum	U42416	AF187060	U72819
Gymno_Juniperus	D38243	AY664834	AY988260
Gymno_Pinus	D38245	D17510	X58134
Gymno_Podocarpus	D38473	AF469661	AF462414
Gymno_Sciadopitys	D85292	AF239792	L25753
Gymno_Taxus	EF017310	AJ235619	AF456388
Gymno_Torreya	D38249	AY664832	AY664858
Gymno_Welwitschia	AF207059	AJ235645	AY296573
Angiospermae			
Acorales_Acorus	AF197584	AJ235381	AY298815
Alismatales_Petrosavia	AF206987	AF209649	AF206806
Alismatales_Spathiphyllum	AF207023	AJ235606	AJ235807
Amborella	U42497	D89556	L12628
Apiales_Griselinia	AF206922	AF209595	L11225
Apiales_Pittosporum	L28142	AJ235561	L11202
Aquifoliales_Gonocaryum	AF206919	AJ235483	AJ235779
Aquifoliales_Ilex	AF206938	AJ235502	X98735
Arecales_Calamus	AF168828	AF233081	M81810
Arecales_Phoenix	AF206991	AF209652	M81814
Asparagales_Behnia	AF206864	AF209542	AF206740
Asparagales_Borya	AF206872	AF209543	AF206741
Asterales_Campanula	U42510	AJ235423	DQ356118
Asterales_Helianthus	AF107577	AJ236205	L13929
Austrobaileyales_Austrobaileya	U42503	AJ235403	L12632
Austrobaileyales_Schisandra	L75842	AJ235599	L12665
Berberidopsidales_Aextoxicon	AF206839	AJ235384	X83986

Table S1. Cont.

Taxon	18S	*atpB*	*rbcL*
Berberidopsidales_Berberidopsis	AF206866	AJ235409	AJ235773
Brassicales_Carica	U42514	AF035901	M95671
Brassicales_Cleome	U42511	AF209565	M95755
Buxales_Buxus	X16599	AF092110	DQ182333
Buxales_Pachysandra	AF094533	AF092111	AF093718
Campanulid_Berzelia	U42508	AF095731	L14391
Campanulid_Eremosyne	U42807	AJ236215	L47969
Campanulid_Escallonia	U42544	AJ235467	L11183
Canellales_Canella	AF206879	AJ235424	AJ131928
Canellales_Zygogynum	AF206865	AJ235408	L12633
Caryophyllales_Delosperma	AF206899	AJ235452	AJ235778
Caryophyllales_Rhabdodendron	AF207007	AJ235578	Z97649
Celastales_Lepuropetalon	L28141	AF209616	L11192
Celastales_Plagiopteron	AF206993	AJ235562	AJ235787
Ceratophyllum_Ceratophyllum	D85300	D89552	M77030
Chloranthales_Hedyosmum	AF206925	AJ235490	L12649
Chloranthales_Sarcandra	AF207012	AJ235593	L12663
Commelinales_Helmholtzia	U42074	AF168922	AY465691
Commelinales_Pontederia	AF168871	AF209657	L20128
Cornales_Cornus	X16602	AJ235444	L11216
Cornales_Hydrangea	U42781	AJ235497	L11187
Crossomitales_Crossosoma	U42529	AF209571	L11179
Crossomitales_Ixerba	AF084476	AF209606	AF084475
Cucurbitales_Coriaria	AF206891	AJ235443	L01897
Cucurbitales_Tetrameles	U41502	AF209689	AF206828
Dasypogonaceae_Dasypogon	AJ417898	AF168907	AF206758
Dilleniales_Dillenia	AY788163	AY788268	L01903
Dilleniales_Tetracera	AJ235982	AJ235622	AJ235796
Dioscorea_Dioscorea	AF069203	AJ235456	AY904802
Dipsacales_Scabiosa	AJ236006	AJ236207	AF156734
Dipsacales_Viburnum	AJ236007	AJ235640	AJ420869
Ericales_Phlox	AJ235996	AJ236221	AF206809
Ericales_Styrax	U43296	AJ235615	L12623
Fabales_Bauhinia	U42537	AF209540	AM234266
Fabales_Pisum	U43011	X03852	X03853
Fagales_Chrysolepis	AF206886	AF209563	AF206750
Fagales_Juglans	AF206943	AF209609	AF206785
Garryales_Aucuba	U42522	AJ235402	L11210
Garryales_Eucommia	L54066	AJ235469	L01917
Gentianales_Exacum	AJ236023	AJ236195	L11684
Gentianales_Mitchella	U42802	AF209630	Z68805
Geraniales_Greyia	U43151	AF209594	L11185
Geraniales_Pelargonium	AF206982	AF035911	L01919
Gunnerales_Gunnera	U43787	EU002162	EU002279
Gunnerales_Myrothamnus	AF094555	AF093386	AF060707
Icacinales_Icacina	AF206935	AF209603	AF206780
Lamiales_Cyrtandra	AJ236053	AJ236172	AF206757
Lamiales_Olea	L49289	AJ236163	AJ001766
Laurales_Calycanthus	U38318	AJ235422	L14291
Laurales_Sassafras	AF233091	AF209668	AF206790
Liliales_Bomarea	AF206871	AJ235413	Z77255
Liliales_Lilium	AF206952	AF209618	Z77295
Magnoliales_Knema	AF206946	AF209611	L12653
Magnoliales_Magnolia	AF206956	AJ235526	AY008934
Malpighiales_Linum	L24401	AJ235521	Z75681
Malpighiales_Populus	AF206999	AF209658	AF206812
Malvales_Bombax	U42507	AJ233051	AF022118
Malvales_Neurada	AF206970	AF209637	U06814
Myrtales_Clarkia	U67930	AF209564	L10896
Myrtales_Heteropyxis	AF206927	AF209597	AF206775
Nymphaeales_Barclaya	AF096692	AF209536	M77028
Nymphaeales_Cabomba	AF206878	AF187058	M77027

Table S1. Cont.

Taxon	18S	*atpB*	*rbcL*
Oncothecaceae_Oncotheca	AF206976	AJ235549	AJ131950
Oxalidales_Averrhoa	AF206859	AJ235404	L14692
Oxalidales_Davidsonia	AF206897	AF209574	AF206759
Pandanales_Barbacenia	AF206861	AJ235406	AJ131946
Pandanales_Freycinetia	AF206915	AF209590	AF206770
Piperales_Aristolochia	AF206855	AJ235399	L12630
Piperales_Piper	AF206992	AJ235560	EF450311
Poales_Flagellaria	AF168845	AF209589	AF206769
Poales_Sparganium	L24419	AF209678	M91633
Proteales_Nelumbo	L75835	AF093387	M77032
Proteales_Placospermum	L75837	AF060391	AF093729
Proteales_Platanus	U42794	U86386	AY858644
Proteales_Roupala	AF094559	AF060416	AF093728
Ranunculales_Coptis	L75838	AF093393	AF093730
Ranunculales_Pteridophyllum	AF094560	U86400	U86631
Rosales_Humulus	AF206931	AF209599	U02729
Rosales_Morus	L24398	AJ235536	L01933
Sabiales_Meliosma	AF206961	AF209626	AF206793
Sabiales_Sabia	L75840	AF093395	AM183414
Santalales_Opilia	U42790	AJ235550	AJ131773
Santalales_Osyris	L24409	AF209641	L11196
Santalales_Schoepfia	AF207017	AF209671	L11205
Sapindales_Poncirus	AF206997	AJ238409	AJ235806
Sapindales_Schinus	AF207015	AF035914	U39270
Saxifrigales_Daphniphyllum	U42531	AF092118	L01901
Saxifrigales_Dudleya	U42526	AJ235461	L11182
Solanales_Hydrophyllum	AJ236019	AJ235498	L01927
Solanales_Nolana	AJ236017	AF209638	U08616
Trochodendrales_Tetracentron	U42814	AF093422	L12668
Trochodendrales_Trochodendron	AF094565	AF093423	L01958
Vitales_Leea	AF206951	AJ235520	AJ235783
Vitales_Vitis	AF207053	AJ235643	L01960
Zingiberales_Dimerocostus	AF168839	AF168909	AF243838
Zingiberales_Musa	U42083	AF168931	L05455

List of taxa included in the phylogenetic analysis with GenBank accession numbers, which were compiled from Soltis et al. (1) and Pyer et al. (2). In many cases, all three genes were not sampled for a particular taxon, and so sequences from closely related species of the same genus were used. The taxon labels are consistent with the labels shown in Figs. S1, S2, S3, and S4.

1. Soltis PS, Soltis DE, Chase MW (1999) Angiosperm phylogeny inferred from multiple genes as a tool for comparative biology. *Nature* 402:402–404.
2. Pryer KM, et al. (2001) Horsetails and ferns are a monophyletic group and the closest living relatives to seed plants. *Nature* 409:618–622.

Table S2. Fossil information, minimum age estimates, and associated lognormal prior-probability distribution parameters for the clades calibrated in our divergence-time analysis of land plants

	Clade	Fossil	Fossil type	Ref.	Stem/Crown	Min. age	Mean (SD)
1	*Tracheophyta*	*Leclerquia* sp.	Plant	1	Stem	377.4	3.5 (0.5)
2	Marattiales	*Marattia, Angiopteris*	Spore	2,3	Crown	166.1	1.5 (0.5)
3	Cyatheales	*Cniopteris, Kylikipteris, Eboracia*	Plant	2	Crown	166.1	1.6 (0.5)
4	*Euphyllophyta*	Unamed	Seeds	4	Crown	365	3.5 (0.5)
5	*Acrogymnospermae*	*Emporia lockardii*	Cones	5,6	Crown	290	2.4 (0.5)
6	Nymphaeales	*Scutifolium jordanicum*	Leaves	7	Crown	105.0	1.5 (0.5)
7	Laurales	Unamed	Flower	8	Crown	108.8	2.1 (0.5)
8	Magnoliales	Unamed	Flower	8	Stem	108.8	1.5 (0.5)
9	Chloranthales	*Hedyosmum* sp.	Flower	9	Crown	121.0	1.5 (0.5)
10	Canellales	Unamed	Pollen	10	Stem	122.5	2.0 (0.5)
11	Pandanales	*Pandanus* sp.	Pollen	11,12	Crown	65.0	1.8 (0.5)
12	Arecales	*Dicolpopollis malesianus*	Pollen	13	Crown	65.0	1.8 (0.5)
13	Poales	*Restio* sp.	Pollen	12,14	Crown	68.1	1.8 (0.5)
14	Zingiberales	*Spirematospermum chandlerae*	Seeds	15	Crown	83.5	1.8 (0.5)
15	Proteales	*Platanocarpus brookensis*	Flower	16	Crown	108.8	1.5 (0.5)
16	Buxales	Unamed	Fruit, flower	17	Stem	112.0	1.5 (0.5)
17	Gunnerales	*Retitricolpites microreticulatus*	Pollen	12	Crown	88.2	1.5 (0.5)
18	Saxifragales	*Divisestylus* sp.	Fruit, flower	18	Crown	89.3	1.5 (0.5)
19	Caryophyllales	Unamed	Seeds	19	Crown	83.5	1.5 (0.5)
20	Dilleniales	*Dillenites* sp.	Seeds	19	Crown	51.9	1.5 (0.5)
21	Santanales	Unamed	Seeds	19	Crown	51.9	1.5 (0.5)
22	Ericales	Unamed	Flower	20	Crown	91.2	1.5 (0.5)
23	Cornales	Unamed	Flower	21	Crown	85.8	1.5 (0.5)
24	Lamiales	*Fraxinus wilcoxiana*	Fruit	22	Crown	44.3	1.5 (0.5)
25	Solanales	*Cantisolanum daturoides*	Fruit	19	Crown	44.3	1.5 (0.5)
26	Aquifoliales	*Ilexpollenites* sp.	Pollen	12,23	Crown	85.0	1.5 (0.5)
27	Vitales	Unamed	Seeds	19	Crown	57.9	1.5 (0.5)
28	Myrtales	*Esqueiria futabensis*	Flower	24	Crown	88.2	1.5 (0.5)
29	Sapindales	Unamed	Fruit	25	Crown	65.0	1.5 (0.5)
30	Malvales	Unamed	Pollen	26	Crown	69.7	1.5 (0.5)
31	Fabales	Unamed	Fruit	27	Crown	59.9	1.5 (0.5)
32	Fagales	*Tenerina* sp.	Pollen	28	Stem	96.0	1.5 (0.5)

1. Grierson JD, Bonamo PM (1979) *Leclercquia complexa*: Earliest ligulate lycopod (Middle Devonian). *Am J Bot* 66:474–476.
2. Harris TM (1961) *The Yorkshire Jurassic Flora: I. Thallophyta-Pteridophyta* (British Museum Natural History, London).
3. Hill CR (1987) Jurassic *Angiopteris* (Marattiales) from North Yorkshire. *Rev Palaeobot Palynol* 51:65–93.
4. Rothwell GW, Scheckler SE (1998) *Origin and Evolution of Gymnosperms* (Columbia University Press, New York), pp 85–134.
5. Mapes G, Rothwell GW (1984) Permineralized ovulate cones of *Lebachia* from Late Paleozoic Hamilton Quarry area in southeastern Kansas. *Palaeontology* 27:69–94.
6. Mapes G, Rothwell GW (1991) Structure and relationship of primitive conifers. *Neues Jahrbuch für Geologie und Palaeontologie* 183:269–287.
7. Taylor DW, Brenner GJ, Basha SH (2008) *Scutifolium jordanicum* gen. et. sp. nov. (Cabombaceae), an aquatic fossil plant from the lower Cretaceous of Jordan, and the relationships of related leaf fossils to living genera. *Am J Bot* 95:340–352.
8. Crane PR, Friis EM, Pedersen KR (1994) Palaeobotanical evidence on the early radiation of magnoliid angiosperms. *Plant Syst Evol* 8:51–72.
9. Friis EM, Pedersen KR, Crane PR (1999) Early angiosperm diversification: the diversity of pollen associated with angiosperm reproductive structures in early Cretaceous floras from Portugal. *Ann Miss Bot Gard* 86:259–296.
10. Doyle JA, Hotton CL, Ward JV (1990) Early Cretaceous tetrads, zonasulcate pollen, and Winteraceae. II. Cladistic analysis and implications. *Am J Bot* 77:1558–1568.
11. Jarzen DM (1978) The terrestrial palynoflora from the Cretaceous-Tertiary transition, Alabama, U.S.A. *Pollen and Spores* 20:535–553.
12. Muller J (1981) Fossil pollen records of extant angiosperms. *Bot Rev* 47:1–142.
13. Pan AD, Jacobs BF, Dransfield J, Baker WJ (2006) The fossil history of palms (Arecaceae) in Africa and new records from the Late Oligocene (28–27 Mya) of north-western Ethiopa. *Bot J Linn Soc* 151:69–81.
14. Jarzen DM (1978) Some Maastrichtian palynomorphs and their phytogeographical and palaeoecological implications. *Palynology* 2:29–38.
15. Friis EM (1988) *Spirematospermum chandlerae* sp. Nov., an extinct Zingiberacaea from the North American Cretaceous. *Tert Res* 9:7–12.
16. Crane PR, Herendeen PS (1996) Cretaceous floras containing angiosperm flowers and fruits from eastern North America. *Rev Palaeobot Palynol* 90:319–337.
17. Anderson CL, Bremer K, Friis EM (2005) Dating phylogenetically basal eudicots using *rbcL* sequences and multiple fossil reference points. *Am J Bot* 92:1737–1748.
18. Hermsen EJ, Gandolfo MA, Nixon KC, Crepet WL (2003) Divisestylus gen. Nov. (aff. Iteaceae), a fossil saxifrage from the late Cretaceous of New Jersey, USA. *Am J Bot* 90:1373–1388.
19. Collinson ME, Boulter MC, Holmes PL (1993) in *The Fossil Record* (Chapman and Hall, London), Vol 2, pp 809–841.
20. Nixon KC, Crepet WL (1993) Late Cretaceous fossil flowers of ericalean affinity. *Am J Bot* 80:616–623.
21. Magallon S (1997) Affinity within Hydrangeaceae of a structurally preserved Late Cretaceous flower (Coniacian-Santonian of Georgia, U.S.A). *Am J Bot* 84 (Suppl.):215.
22. Call VB, Dilcher DL (1992) Investigations of angiosperms from the Eocene of southwestern North America: samaras of *Fraxinus wilcoxiana* berry. *Rev Palaeobot Palynol* 74:249–266.
23. Martin HA (1977) The history of Ilex (Aquifoliaceae) with special reference to Australia: evidence from pollen. *Aust J Bot* 25:655–673.
24. Takahashi M, Crane PR, Ando H (1999) *Esgueiria futabensis* sp. nov., a new angiosperm flower from the Upper Cretaceous (lower Coniacian) of northeastern Honshu, Japan. *Paleontol Res* 3:81–87.
25. Knobloch ED, Mai DH (1986) Monograph of the fruits and seeds in the Cretaceous of Central Europe. *Rozpravy Ústreõdního Ústavu Geologického*, 47:1–219 (in German).
26. Wolfe JA (1976) Stratigraphic distribution of some pollen types from the Campanian and lower Maestrichtian rocks (Upper Cretaceous) of the middle Atlantic states. *U.S. Geological Survey Professional Paper* 997:1–108.
27. Herendeen PS, Crane PR (1992) *Advances in Legume Systematics. Pt. 4. The Fossil Record* (Royal Botanic Gardens, Kew, United Kingdom), pp 57–68.
28. Friis EM, Pedersen KR, Schonenberger J (2006) Normapolles plants: a prominent component of the Cretaceous rosid diversification. *Plant Syst Evol* 260:107–140.

Table S3. Divergence-time estimates (in Myr) for major clades of land plants as estimated partitioning for codon position

Clade	Coding analysis
Land plants	442 (401–514)
Tracheophyta	411 (392–442)
Spermatophyta	331 (306–363)
Acrogymnospermae	301 (294–313)
Gnetophyta	157 (103–218)
Angiospermae	231 (189–281)
Magnoliidae	166 (147–190)
Monocotyledoneae	145 (127–168)
Eudicotyledoneae	155 (137–176)

Dates within parentheses denote the 95% HPD.