Int. J. Plant Sci. 164(5):737–751. 2003.

1058-5893/2003/16405-0008$15.00

PHYLOGENETIC RELATIONSHIPS AND EVOLUTION OF EXTANT HORSETAILS, *EQUISETUM*, BASED ON CHLOROPLAST DNA SEQUENCE DATA (*rbcL* AND *trnL-F*)

David L. Des Marais, Alan R. Smith, Donald M. Britton, and Kathleen M. Pryer[1]

Department of Biology, Duke University, Durham, North Carolina 27708-0338, U.S.A.; University Herbarium, University of California, Berkeley, California 94720-2465, U.S.A.; Department of Molecular Biology and Genetics, University of Guelph, Guelph, Ontario N1G 2W1, Canada; and Department of Botany, Field Museum of Natural History, 1400 South Lake Shore Drive, Chicago, Illinois 60605-2496, U.S.A.

Equisetum is a small and morphologically distinct genus with a rich fossil record. Two subgenera have been recognized based principally on stomatal position and stem branching: subg. *Equisetum* (eight species; superficial stomates; stems branched) and subg. *Hippochaete* (seven species; sunken stomates; stems generally unbranched). Prior attempts at understanding *Equisetum* systematics, phylogeny, and character evolution have been hampered by the high degree of morphological plasticity in the genus as well as by frequent hybridization among members within each subgenus. We present the first explicit phylogenetic study of *Equisetum*, including all 15 species and two samples of one widespread hybrid, *Equisetum* × *ferrissii*, based on a combined analysis of two chloroplast markers, *rbcL* and *trnL-F*. Our robustly supported phylogeny identifies two monophyletic clades corresponding to the two subgenera recognized by earlier workers. The phylogenetic placement of *Equisetum bogotense*, however, is ambiguous. In maximum likelihood analyses, it allies with subg. *Hippochaete* as the most basal member, while maximum parsimony places it as sister to the rest of the genus. A consensus phylogeny from the two analyses is presented as a basal trichotomy (*E. bogotense*, subg. *Hippochaete*, subg. *Equisetum*), and morphological character evolution is discussed. We detected rate heterogeneity in the *rbcL* locus between the two subgenera that can be attributed to an increased rate of nucleotide substitution (transversions) in subg. *Hippochaete*. We calculated molecular-based age estimates using the penalized likelihood approach, which accounts for rate heterogeneity and does not assume a molecular clock. The *Equisetum* crown group appears to have diversified in the early Cenozoic, whereas the Equisetaceae total group is estimated to have a Paleozoic origin. These molecular-based age estimates are in remarkable agreement with current interpretations of the fossil record.

Keywords: divergence times, *Equisetum*, fossil record, horsetails, likelihood ratio test, molecular clock, morphological character evolution, penalized likelihood, phylogeny.

Introduction

The Equisetopsida (Archaeocalamitaceae, Calamitaceae, Equisetaceae) are an ancient lineage of "seed-free" vascular plants distinguished in part by the regular alternation of whorled appendages at successive nodes, highly reduced leaves, a characteristic stele morphology, and sporangiophores (Page 1972*b*; Bateman 1991; Kenrick and Crane 1997*a*). The arborescent Calamitaceae and Archaeocalamitaceae are likely to have been prominent components of Paleozoic wetland habitats (DiMichele and Hook 1992; DiMichele et al. 2001), but by the mid-Permian these woody calamites were extinct (Bateman 1991). These two families shared such features as the presence of bracts in the cones between the whorls of sporangiophores and a vascular cambium (Crane 1989), which were presumably lost in Equisetaceae (*Equisetum, Equisetites*).

The Sphenophyllales (e.g., *Sphenophyllum*, *Bowmanites*) are thought to be closely related to the Equisetopsida because they also have whorled appendages, secondary xylem, and sporangiophores (Stein et al. 1984; Kenrick and Crane 1997*a*), though some debate surrounds these putative homologies. Based on the exarch development of the metaxylem, Stewart and Rothwell (1993) suggested a possible relationship of the Sphenophyllales to the Lycopsida, but many similarities with the Equisetopsida make this relationship appear unlikely (Kenrick and Crane 1997*a*). The early fossil genus *Ibyka* from the middle Devonian (Skog and Banks 1973) was shown by Kenrick and Crane (1997*a*) to be sister to Equisetopsida based on whorled branching (Stein et al. 1984) and protoxylem disintegration to form lacunae (Skog and Banks 1973). Other middle and upper Devonian taxa, such as *Hyenia*, *Calamophyton*, and *Pseudobornia*, have been variously associated with members of the Equisetopsida and Sphenophyllales, but their phylogenetic status is more problematic (Berry and Fairon-Demaret 2001; Pryer et al. 2004).

The sole living genus in the Equisetaceae, *Equisetum*, is herbaceous and thought to have diverged—or arisen by anagenesis—in the Tertiary from an older genus, *Equisetites*, which dates to the mid-Permian (Stewart and Rothwell 1993), with

[1] Author for correspondence; current address: Department of Biology, Duke University, Durham, North Carolina 27708-0338; e-mail pryer@duke.edu.

Manuscript received December 2002; revised manuscript received March 2003.

possible records dating back to the Carboniferous (Taylor and Taylor 1993). Historically, workers have placed fossil forms that are indistinguishable from modern horsetails in *Equisetum*, while *Equisetites* has been used for Mesozoic fossils that cannot be positively allied with extant species (Kelber and van Konijnenburg-van Cittert 1998). *Equisetum* is a small, easily recognized, and highly distinctive genus of vascular plants with a subcosmopolitan distribution. Most species are found between 40° and 60° north latitude and are generally confined to seasonally or sometimes perennially wet ground (Kenrick and Crane 1997*a*).

Taxonomic treatments of *Equisetum* have identified a variable number of species. We follow Hauke (1983, 1993) in recognizing 15 species worldwide, with 11 of them found in North America, north of Mexico. All species of *Equisetum* have the same chromosome number ($n = 108$; Manton 1950). Most authors (Hauke 1963, 1978) divide the genus into two subgenera: subg. *Hippochaete* (seven species), or scouring rushes, characterized by stomates sunken relative to the epidermis and generally unbranched stems, and subg. *Equisetum* (eight species), or horsetails, having stomates flush with the epidermis and stems frequently branched. Several subgeneric (sectional) classification schemes have been suggested by past workers; we present Hauke's (1963, 1978) in table 1.

The broad range of gross morphological variation within *Equisetum* and the cryptic nature of a few stable characters have resulted in a proliferation of described varieties and subspecies. Rampant hybridization and the persistence of sterile hybrids and stress-induced mutants made previous rigorous systematic analysis of the genus challenging (Hauke 1983). Early work by Schaffner (1921, 1925*a*, 1925*b*, 1930*a*, 1930*b*) on *Equisetum* identified between 20 and 23 species in five sections, which he placed along an evolutionary continuum. Schaffner's comparative studies led him to suggest that the large, perennial tropical American species, *Equisetum giganteum* and *Equisetum myriochaetum*, represented the "primitive" condition (Schaffner's [1925*b*] "Primitiva" group) and that all other *Equisetum* species represented progressively more advanced forms. Schaffner considered *Equisetum arvense*, a north-temperate annual with sterile and fertile (dimorphic) aerial stems, to be the most derived *Equisetum*.

Hauke (1963, 1978), recognizing that members of the genus *Equisetum* exhibited considerable morphological plasticity and hybridized readily, undertook a rigorous reassessment of the genus. In addition to reducing Schaffner's 23 species to 15, Hauke (1963, 1978) identified 15 widespread and common hybrids, most of which were sterile, and attributed their ecological success to the ease of persistence of the rhizomatous growth habit and propagation by vegetative fragments. Hauke (1963) presented a character-based phylogeny of subg. *Hippochaete* employing Wagner's (1961; reviewed in Wagner 1980) ground-plan divergence scheme. In his monograph of subg. *Equisetum*, Hauke (1978) calculated a phenetic dendrogram to represent relationships among the eight species, using similarity percentages based on 18 morphological characters. Unimpressed by the results, Hauke (1978, p. 401) stated, "I do not believe we can trace any sort of phylogency [*sic*] for the subgenus."

Duckett (1979*b*) presented diagrammatic representations of interspecific relationships in *Equisetum* based on artificial hybridization frequencies and also on published data on the frequency of putative natural hybrids. The absence of hybridization between subgenera in Duckett's (1979*b*) study supports the division of *Equisetum* into two subgenera. DNA hybridization analyses (Bendich and Anderson 1983) also supported the division of the genus into two subgenera but provided no further details on phylogenetic relatedness among species.

The molecular and morphological phylogeny of Pryer et al. (2001*b*) indicates that extant horsetails, together with the eusporangiate and leptosporangiate ferns, form a monophyletic group (Infradivision Moniliformopses; Kenrick and Crane 1997*a*) that is sister to the seed plants. Our study presents a phylogeny of *Equisetum* based on DNA sequence data from two chloroplast markers: *rbcL* and the *trnL-F* region, which includes the quickly evolving *trnL* intron and the *trnL-F* spacer. Based on this molecular phylogeny, we reconsider the taxonomy and morphological evolution of the group. In addition, we compare the fossil record with the timing of historical diversification based on molecular data to determine whether the *Equisetum* crown group is a relatively recent (Cenozoic) genus or if it could indeed be ancient, with its origins in the Paleozoic.

Material and Methods

Taxon Sampling

Voucher information for all *Equisetum* taxa sampled in this study is presented in table 1. All 15 species of *Equisetum* and two specimens of one widespread hybrid, *Equisetum* × *ferrissii*, were included. Our taxonomy follows Hauke (1963, 1978, 1983). Based on Pryer et al. (2001*b*), *Psilotum*, *Botrychium*, *Danaea*, *Angiopteris*, and *Osmunda* were selected as outgroup taxa (table 2); *Psilotum* and *Botrychium* were used to root the topologies.

DNA Extraction, Amplification, and Sequencing

Total genomic DNA was extracted from silica-dried tissue following Dubuisson (1997), with the exception of using 5.5% DTAB buffer rather than 2% CTAB. Approximately 1.4 kb of the *rbcL* gene and 0.5–0.8 kb (depending on indel lengths) of the *trnL-F* region were amplified by the polymerase chain reaction (PCR). The amplification primers for *rbcL* are shown in figure 1; *trnL-F* primers followed Taberlet et al. (1991). PCR mixtures were 25 μL, and each contained 0.125 μL Taq polymerase (Roche Diagnostics, Indianapolis), 2.5 μL of 10X PCR Buffer (PerkinElmer, Foster City, Calif.), 0.5 μL MgCl (25 mM), 0.5 μL bovine serum albumin (BSA), 2.5 μL dNTP (10 mM), 1.25 μL of each amplification primer (10 μM), 0.5 μL genomic DNA, and purified water to volume.

PCR cycles were programmed on a PTC-200 DNA Engine (MJ Research, Watertown, Mass.) starting with a 5-min 94°C denaturation and finishing with a 4-min 72°C extension step. Specific PCR conditions used for *rbcL* were 94°C denaturation for 45 s, 54°C annealing for 30 s, and 72°C extension for 90 s; for *trnL-F*, we used 94°C denaturation for 45 s, 52°C annealing for 30 s, and 72°C extension for 70 s. All reactions were held at 4°C before visualization with ethidium bromide on a 1.5% agarose gel. Amplicons were purified using Millipore spin columns (Millipore, Bedford, Mass.). Both strands

Table 1

***Equisetum* Taxa Examined, Taxonomic Section, Vouchers, and Database Accession Numbers**

Taxon	Section (Hauke 1963, 1978)	Voucher information	GenBank accession number (*rbcL*/*trnL-F*)	Fern DNA database number[a]
Equisetum L. subg. *Equisetum*:				
E. arvense L.	*Heterophyadica*	France: North Belledonne Range, *P.G. Wolf 907* (UTC; UC 1742206)	AY226140/AY226125	1035
E. bogotense Kunth	*Equisetum*	Colombia: La Conche, *Bruce McAlpin s.n.* (UC 1733471)	AY226139/AY226124	765
E. diffusum D. Don	*Equisetum*	India: cult. United States: New York Botanical Garden (acc. no. 1352/77), *A.R. Smith 2637* (UC 1763851)	AY226141/AY226126	2015
E. fluviatile L.	*Equisetum*	Canada: Ontario, Hunt-Wellington Co., *K.M. Pryer 00-05 & D.M. Britton* (DUKE)	AY226142/AY226121	1038
E. palustre L.	*Equisetum*	Canada: Ontario, Hunt-Wellington Co., *K.M. Pryer 00-10 & D.M. Britton* (DUKE)	AY226138/AY226123	1043
E. pratense Ehrh.	*Heterophyadica*	Canada: Ontario, Hunt-Wellington Co., *K.M. Pryer 00-06 & D.M. Britton* (DUKE)	AY226137/AY226122	1039
E. sylvaticum L.	*Heterophyadica*	France: Belledonne Mountain Range, *P.G. Wolf 908* (UTC; UC 1742205)	AY226136/AY226120	1003
E. telmateia Ehrh. subsp. *braunii* (J. Milde) Hauke	*Equisetum*	United States: California, Alameda Co., *A.R. Smith 2575* (UC 1733467)	AY226135/AY226119	768
Equisetum L. subg. *Hippochaete* (J. Milde) Baker:				
E. giganteum L.	*Incunabula*	Ecuador: Napo, *R. Aquinda s.n.* (DUKE)	AY226127/AY226118	1048
E. hyemale L. subsp. *affine* (Engelm.) Calder & R.L. Taylor	*Hippochaete*	United States: California, Alameda Co., *A.R. Smith s.n.* (UC 1733472)	AY226128/AY226110	764
E. laevigatum A. Braun	*Ambigua*	United States: Utah, Logan, *W. Speer s.n.* (UTC; UC 1748060)	AY226130/AY226112	1050
E. myriochaetum Schltdl. & Cham.	*Ambigua*	United States: Illinois, Chicago, *D.L. Des Marais 00-47 & K.M. Pryer* (DUKE)	AY226131/AY226114	1045
E. ramosissimum Desf. subsp. *debile* (Roxb.) Hauke	*Ambigua*	Taiwan: *R. Cranfill s.n.* (UC)	AY226132/AY226115	2018
E. scirpoides Michx.	*Hippochaete*	Canada: Ontario, Hunt-Wellington Co., *K.M. Pryer 00-04 & D.M. Britton* (DUKE)	AY226133/AY226116	1037
E. variegatum Schleich. ex F. Weber & D. Mohr	*Hippochaete*	United States: Alaska, Prince of Wales Island, *B. Hoshizaki 592* (UC 1733475)	AY226134/AY226117	761
E. ×*ferrissii* Clute (*E. hyemale* {female} × *laevigatum* {male})	(Hybrid)	Canada: Ontario, Hunt-Wellington Co., *K.M. Pryer 00-07 & D.M. Britton* (DUKE)	AY226129/AY226111	1040
E. ×*ferrissii* Clute (*E. laevigatum* {female} × *hyemale* {male})	(Hybrid)	United States: California, San Francisco, *P. Hammond s.n.* (UC 1733476)	AF313579/AY226113	760

[a] Permanent record numbers in http://www.biology.duke.edu/pryerlab/ferndb.

of purified PCR products were directly sequenced in 10-μL reactions using the *rbcL* primers shown in figure 1 and the *trnL-F* primers in Taberlet et al. (1991). Cycle sequencing used BigDye version 2 sequencing chemistry, and reaction mixtures were run on an ABI 377 automated DNA sequencer (Applied Biosystems, Foster City, Calif.).

Sequencing reads from each amplicon were compared with GenBank databases using a BLASTN similarity search to reduce the possibility of including contaminant sequence data. Sequence data were edited and assembled into alignments using the program Sequencher 4.0 (Gene Codes, Ann Arbor, Mich.). All the *Equisetum rbcL* (except AF313579) and *trnL-F* sequences obtained in this study are new and have been deposited in GenBank (table 1). For the outgroups, a new *trnL-F* sequence was obtained for *Psilotum* (table 2) and deposited in GenBank.

Sequence Alignment and Phylogenetic Analyses

The *rbcL* and *trnL-F* sequences were aligned by eye using Sequencher 4.0. The *rbcL* alignment contained no gaps and no ambiguities. Compared to *rbcL*, the *trnL-F* sequences were highly variable, and thus the final alignment contained numerous gaps and positional ambiguities. Due to the high degree of sequence divergence between the ingroup and outgroup taxa, *Psilotum* was the only outgroup that could be reliably included in the *trnL-F* alignment. Ambiguously aligned regions (e.g., regions of uncertain homology among taxa) of the *trnL* intron and *trnL-F* spacer were delimited following the method of Lutzoni et al. (2000). Because of their large size and dubious phylogenetic importance, these regions of questionable homology were excluded entirely from all analyses.

Maximum parsimony and maximum likelihood analyses were performed using PAUP* version 4.0b8 (Swofford 2000) on Macintosh Power PCs and on the High Performance Computer Cluster facility at the Field Museum (a Microway Linux cluster of six dual Alpha processor compute nodes/500 MHz and four UltraSparcII/400 MHz). In the maximum parsimony analyses, *rbcL* and *trnL-F* character-state changes were initially weighted equally. In subsequent analyses of *rbcL*, character-state changes for each codon position were weighted using a symmetric step matrix with costs calculated using the negative natural logarithm of the observed average frequency of each possible symmetric character-state change (A↔C, A↔G, A↔T, C↔G, C↔T, G↔T) at that codon position across all sites (Felsenstein 1981; Wheeler 1990; Maddison and Maddison 2000). The character-state changes of the nonprotein coding *trnL-F* region were weighted using a single symmetric step matrix with costs calculated as above, but across all included sites. Step matrices were implemented in the Assumptions block of the NEXUS files and tested for triangle inequality. Alignments (including step matrices) are available in TreeBASE (http://www.treebase.org/treebase/).

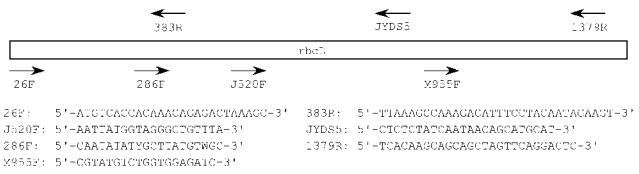

Fig. 1 Sequences and approximate annealing sites for primers used in amplifying (external) and sequencing (both external and internal) *rbcL* in *Equisetum*.

We tested for homogeneity of the *rbcL* and *trnL-F* data by using the partition-homogeneity test (Farris et al. 1995) as implemented in PAUP*. In the combined data set analyses, all the outgroup taxa (except *Psilotum*) that had been excluded from the *trnL-F* alignment because of ambiguous positional homology were scored as missing data for *trnL-F*. Parsimony heuristic searches for both equally and unequally weighted analyses were conducted on the *trnL-F* and *rbcL* alignments separately and on a combined data set with 1000 random-addition-sequence replicates, TBR branch swapping, MULTrees on, and zero-length branches collapsed. Branch support was determined by 1000 bootstrap replicates with 100 random-addition-sequence replicates, using the same parameters as for the heuristic searches.

Maximum likelihood analyses for *rbcL* were run using the TrN model and heterogeneous rates of change (four rate categories) across all sites following a discrete gamma (Γ) distribution. *TrnL-F* analyses were run using the F81 + Γ model and four rate categories. Combined data set (*rbcL* + *trnL-F*) analyses used the TrN + Γ model with four rate categories. The appropriate maximum likelihood evolutionary models, Γ-

Table 2

Outgroup Taxa Examined, Vouchers, and Database Accession Numbers

Taxon	Voucher information[a]	GenBank accession number (*rbcL*/*trnL-F*)	Fern DNA database number[b]
Psilotum nudum (L.) P. Beauv	…	L11059/…	627
Psilotum nudum (L.) P. Beauv	United States: Utah State University Greenhouse, P.G. Wolf 608	…/AY241586	623
Botrychium biternatum (Sav.) Underw.	…	L13474/…	480
Danaea elliptica Sm.	…	AF313578/…	451
Angiopteris lygodiifolia Rosenst.	…	X58429/…	447
Osmunda cinnamomea L.	…	D14882/…	497

[a] Only for newly sequenced taxon.

[b] Permanent record numbers in http://www.biology.duke.edu/pryerlab/ferndb.

distribution shape parameters, nucleotide frequencies, and substitution-rate parameters were estimated empirically for each data partition via maximum likelihood using a hierarchical likelihood ratio test implemented with the program Modeltest version 3.06 (Posada and Crandall 1998). For each separate analysis and the combined analysis, 1000 random-addition-sequence searches were conducted to find the most likely topology. Branch support using the likelihood criterion was determined by 1000 bootstrap replicates with 10 random-addition-sequence replicates per bootstrap replicate. Branch support for the combined data set analysis was also measured with the Bayesian Markov Chain Monte Carlo (BMCMC) method, using the program MrBayes version 1.11 (Huelsenbeck 2000). Four parallel MCMC samplers (three of which were "heated") of one million generations each were run using the same maximum likelihood parameters used for the combined data set analysis as starting priors. Following an initial "burn-in" phase of one hundred thousand generations to optimize the likelihood of tree space, we sampled one tree every hundredth generation and then built a majority-rule consensus tree of these 9000 trees. For the purpose of comparison with bootstrapping, we chose to consider nodes with posterior probabilities greater than 0.90 (e.g., the node appears in greater than 90% of sampled trees) as being well supported.

Divergence Times Estimated Using rbcL

A likelihood ratio test was conducted using PAUP* to determine whether a global molecular clock hypothesis could be applied to the sampled ingroup *rbcL* sequences. Using the program HYPHY (Pond and Muse 2001), we tested for between-taxa rate heterogeneity (multiple local molecular clock hypotheses) using pairwise likelihood ratio tests following the method of Muse and Weir (1992). All pairs of *Equisetum* species were tested for branch-length heterogeneity against each pair's nearest sister taxon. For pairwise comparisons between subgenera, *Danaea* was used as the outgroup. All tests were run under the TrN + Γ likelihood model with four rate categories. Rate heterogeneity was detected at a 5% significance level.

Owing to the nonclocklike behavior of *rbcL* observed across the land plants (Bousquet et al. 1992; Gaut et al. 1992; Sanderson 1997; Soltis et al. 2002), an approach that allows for rate heterogeneity must be employed to separate rate and time in estimated branch lengths. Because a molecular clock could not be enforced for our *rbcL* data set, we estimated approximate lineage divergence times using the penalized likelihood method of Sanderson (2002*a*), which is a semiparametric approach using rate smoothing, to allow for robust estimation of node ages in the presence of rate variation between lineages. A "supertree" (Sanderson et al. 1998) topology was assembled based on the *rbcL* sequences from this study and the complete sample of *rbcL* sequences from Pryer et al. (2001*b*). The green alga *Chara* was used to root the tree (Karol et al. 2001) and to partition branch lengths among the basalmost taxa; *Chara* was then excluded in subsequent analyses. Branch lengths were estimated via maximum likelihood using the TrN + Γ model with four rate categories and invariants. The maximum likelihood model, proportion of invariable sites, Γ-distribution shape parameters, nucleotide frequencies, and substitution-rate parameters were estimated via a hierarchical likelihood ratio test using Modeltest version 3.06 (Posada and Crandall 1998). All penalized likelihood analyses were implemented with the program r8s version 1.5 (Sanderson 2002*b*) on a Linux platform. A cross-validation procedure estimated the optimal rate-smoothing value to be 1000. In all analyses, the root of the land plant tree was fixed to 476 Ma, the presumed time of origin of land plants based on fossil evidence (Kenrick and Crane 1997*b*). In the constrained analyses, the origin of the Equisetaceae was held to a minimum of 380 Ma, seed plants to 364 Ma, and leptosporangiate ferns to 340 Ma (Kenrick and Crane 1997*b*). Both constrained and unconstrained analyses were run using the Truncated Newton (TN) algorithm (Sanderson 2002*b*). Confidence intervals (95%) for the divergence times of clades were approximated by inspecting the likelihood surface of the best-fit TN solutions. All ages were also estimated using the Powell algorithm to check for consistency.

Results

The *rbcL* alignment contained 1326 bp for 17 ingroup taxa and five outgroup taxa without positional homology ambiguities or gaps. Of these, 905 characters were constant, 140 were variable but parsimony uninformative, and 279 characters were parsimony informative. Implementation of the three codon-position step matrices in the unequally weighted analysis increased the number of parsimony-informative characters to 281. The *trnL-F* sequences varied considerably in length within *Equisetum* alone, from 519 bp in *Equisetum hyemale* to 765 bp in *Equisetum palustre*. Including gaps, the final *trnL-F* alignment contained 1016 characters, of which 594 were excluded due to positional homology ambiguities. Of the 422 included characters, 321 were constant, 68 were variable but parsimony uninformative, and 33 were parsimony informative. Implementation of the step matrix in the unequally weighted analysis increased the number of parsimony-informative characters to 39. The partition homogeneity test (Farris et al. 1995) confirmed that the *rbcL* and *trnL-F* data were congruent, allowing us to combine the data sets for simultaneous analysis; hence, we report here only on the results of the combined data set analyses.

The same single most parsimonious reconstruction was recovered from both the equally and unequally weighted analyses of the combined *rbcL* + *trnL-F* data (fig. 2). The most likely tree recovered using the maximum likelihood optimization criterion is shown in figure 3. Both parsimony and likelihood analyses revealed exceptional support for the monophyly of *Equisetum* (figs. 2, 3). Within *Equisetum*, two monophyletic clades corresponding to subg. *Hippochaete* and subg. *Equisetum* (minus *Equisetum bogotense*) were resolved with parsimony with robust support (100% and 94% bootstrap values, respectively); *E. bogotense* was isolated (80% bootstrap value) as sister to all other species of *Equisetum* (fig. 2). The maximum likelihood analysis yielded essentially the same topology as parsimony but with two major differences: weak support (62% bootstrap value) for subg. *Equisetum* (minus *E. bogotense*) and weak resolution (60% bootstrap value/

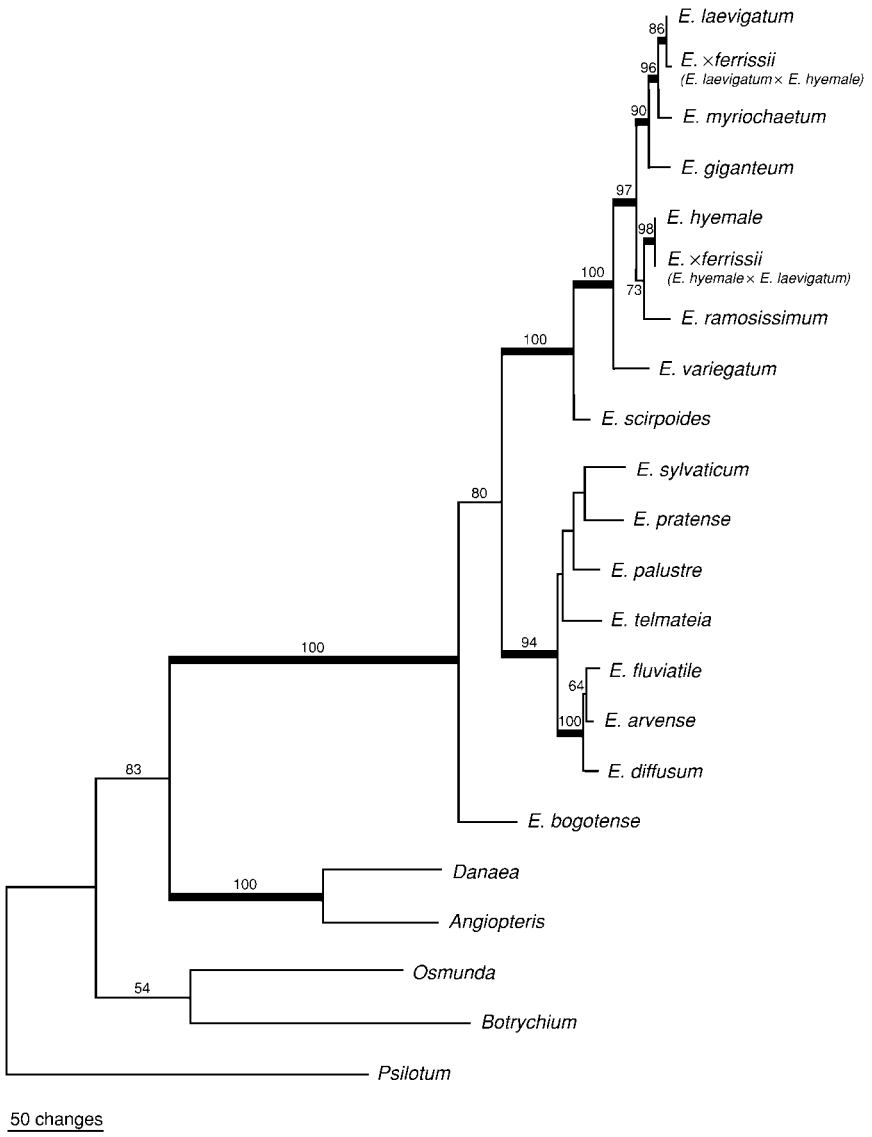

Fig. 2 Single most parsimonious reconstruction of *Equisetum* relationships recovered from both equally weighted and unequally weighted analyses of combined *rbcL* + *trnL-F* DNA sequence data. Tree statistics shown are from unequally weighted analysis: tree length = 1281.56 steps, Consistency Index (excluding uninformative characters) = 0.6580, Retention Index = 0.7977. Numbers at nodes are bootstrap values >50%. Branches with thick lines are the most robustly supported (bootstrap value >85%). Scale bar indicates a branch length corresponding to 50 character-state changes.

72% BMCMC posterior probability) of *E. bogotense* as sister to subg. *Hippochaete* (fig. 3). Both the parsimony and the likelihood analyses are identical in their resolution of a mostly pectinate arrangement of the seven species within subg. *Hippochaete*, with *Equisetum scirpoides* as the earliest diverging taxon and *Equisetum laevigatum* and *Equisetum myriochaetum* being the most derived species pair (figs. 2, 3). All branches within subg. *Hippochaete* are very strongly supported. The two hybrid individuals of *Equisetum* × *ferrissii* (table 1) that were included in this study do not come together within subg. *Hippochaete*; instead, one pairs with *E. laevigatum* and the other with *E. hyemale* (figs. 2, 3).

Relationships within subg. *Equisetum* (minus *E. bogotense*) are less well defined in both analyses, except for one clade (*E. diffusum*, (*E. arvense*, *E. fluviatile*)), which comes together with high support in both the parsimony and likelihood analyses (100% bootstrap value and 90% bootstrap/94% BMCMC posterior probability, respectively) (figs. 2, 3). Although very weakly supported, the other taxa that make up subg. *Equisetum* come together as a clade (*E. telmateia*, (*E. palustre*, (*E. sylvaticum*, *E. pratense*))) using either optimization criterion (figs. 2, 3).

The likelihood ratio test rejected the global molecular clock hypothesis for the *rbcL* data set ($P < 0.001$). The results of the pairwise relative rate tests are summarized in figure 4. Note that because of the ambiguous position of *E. bogotense* in the

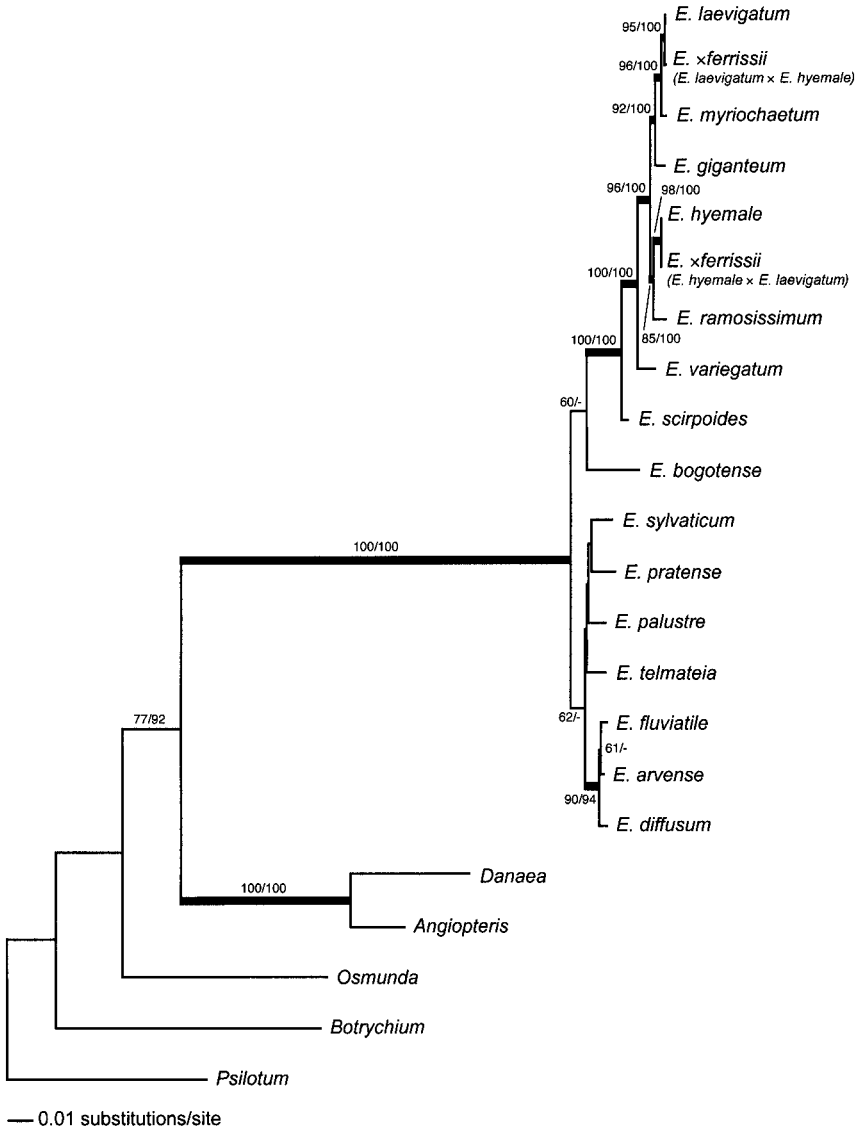

Fig. 3 Best estimate of topology for *Equisetum* relationships based on maximum likelihood analysis of combined *rbcL* + *trnL-F* DNA sequence data; ln likelihood = −6579.6525. Numbers at nodes before the slash are bootstrap values >50%; numbers following the slash are Bayesian Markov Chain Monte Carlo (BMCMC) posterior probabilities >90% (a dash indicates a BMCMC value <90% for a node with bootstrap support >50%). Branches with thick lines are the most robustly supported (bootstrap value ≥85% or Bayesian posterior probability ≥94%). Scale bar indicates a branch length corresponding to 0.01 substitutions per site.

parsimony and likelihood analyses (cf. figs. 2, 3), the overall topology in figure 4 is depicted as a basal trichotomy (*E. bogotense*, subg. *Hippochaete*, subg. *Equisetum*). Statistically significant rate acceleration was revealed most notably on the lineage leading to subg. *Hippochaete* (as compared to the lineage leading to subg. *Equisetum*), after the divergence of *E. scirpoides* and before the divergence of *E. variegatum* (fig. 4). Using MacClade (Maddison and Maddison 2000) to determine the most parsimonious ancestral character state reconstruction for each variable nucleotide position, we were able to determine that the significant rate difference between the two subgenera was due exclusively to an increased rate of transversions in subg. *Hippochaete*. The transition/transversion ratio is 1.87 in subg. *Hippochaete*, compared to 3.40 in subg. *Equisetum*. Furthermore, the ratio of nonsynonymous substitutions (Ka) to synonymous substitutions (Ks) (=Ka/Ks ratio) was slightly higher in subg. *Hippochaete* (0.173) compared to subg. *Equisetum* (0.100).

Our estimates of divergence times for *Equisetum* using penalized likelihood are presented as a chronogram (fig. 5). Based on the constrained analysis (data not shown), which held the (*Equisetum*) + (*Danaea*, *Angiopteris*) split to 381 Ma, extant *Equisetum* species were estimated to have diverged 48.9 Ma (95% confidence interval = 36.3 to 67.7 Ma). Unconstrained analyses estimated the same divergence to be 36.3 (95% confidence interval = 26.9 to 48.1 Ma) and the divergence of the

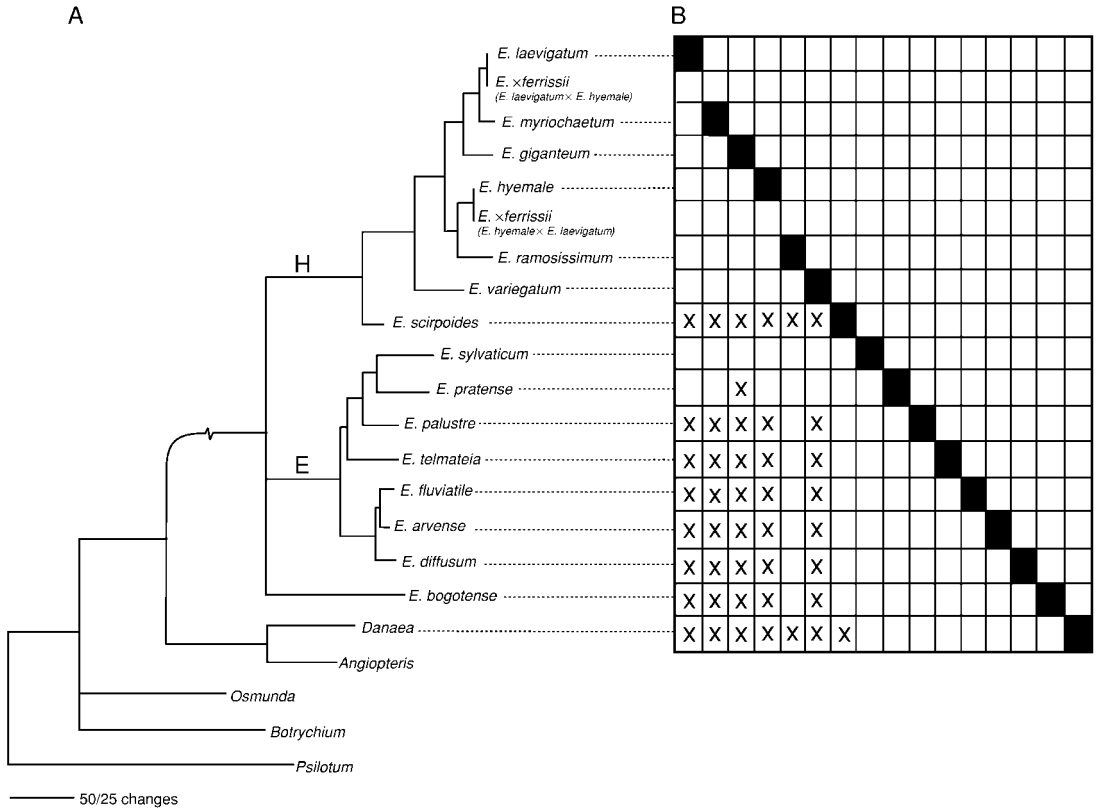

Fig. 4 Results of the pairwise relative rate tests. *A*, Overall summary topology of *Equisetum* relationships based on results in figures 2 and 3, with *Equisetum bogotense*, subg. *Hippochaete* (*H*), and subg. *Equisetum* (*E*) depicted as a basal trichotomy. Length of branch leading to *Equisetum* was reduced by half to accommodate figure size; basal to the break, scale bar indicates 50 character-state changes, whereas distal to the break, scale bar indicates 25 character-state changes. *B*, Each column represents all pairwise comparisons between one species (black box) and the remaining species (white boxes). Note that the two hybrid samples of *Equisetum* × *ferrissii* were not included in the comparisons (no black box in their respective rows). A white box for a species in any column indicates that the rate of substitution for that particular species is not significantly different from the rate for the species with the black box in that same column. A box with an x for a species in any column indicates that the rate of substitution for that particular species is significantly slower than the rate for the species with the black box in that same column. For example, *Equisetum giganteum* is not significantly faster than any other species in subg. *Hippochaete*, except for *Equisetum scirpoides*, but it is significantly faster than *Equisetum bogotense* and all species in subg. *Equisetum*, except for *Equisetum sylvaticum*. Lineages were considered to be evolving at significantly different rates at $P \leq 0.05$.

Equisetaceae and Marattiales (*Danaea*, *Angiopteris*) to have occurred 266.1 Ma (95% confidence interval = 237.7 to 295.2 Ma). The ages estimated by the Powell algorithm were similar, so only the results estimated via the TN algorithm are presented.

Discussion

Phylogeny and Evolution of Equisetum

Fifteen horsetail species are the sole surviving link with a once ecologically prominent and diverse group of early land plants (Page 1972*b*; Bateman 1991; Kenrick and Crane 1997*a*). The great antiquity of the lineage together with its highly distinctive morphology has made it of considerable interest to botanists and paleobotanists alike. Though few in species number, the ubiquitous distribution of *Equisetum* has also made it a familiar curiosity to many people. This study is the first to consider all 15 living species of *Equisetum* in an explicit phylogenetic framework. The phylogenetic relationships presented here (figs. 2, 3) support the monophyly of *Equisetum*, which has not been questioned. Milde (1865), however, regarded the stomatal character (stomates flush with epidermis vs. stomates sunken) to be so important that he split the genus into two genera: *Equisetum* and *Hippochaete*. This taxonomic disagreement has persisted and has been reflected in the classification of *Equisetum* for more than 100 yr, with some botanists agreeing with and finding support for the generic distinction (Farwell 1916) and others arguing against it, preferring to recognize the two groups as subgenera of a single genus (Schaffner 1921; Hauke 1963). To Hauke (1963), the differences (stomatal structure, chromosomes of *Hippochaete* larger than those of *Equisetum* [Manton 1950], no hybrids known to occur between *Hippochaete* and *Equisetum*) were far outweighed by the numerous similarities the two groups shared in general morphology, anatomy, and reproductive structure. Our study supports the monophyly of subg. *Hippochaete* (*sensu* Hauke 1963) but rejects the continued rec-

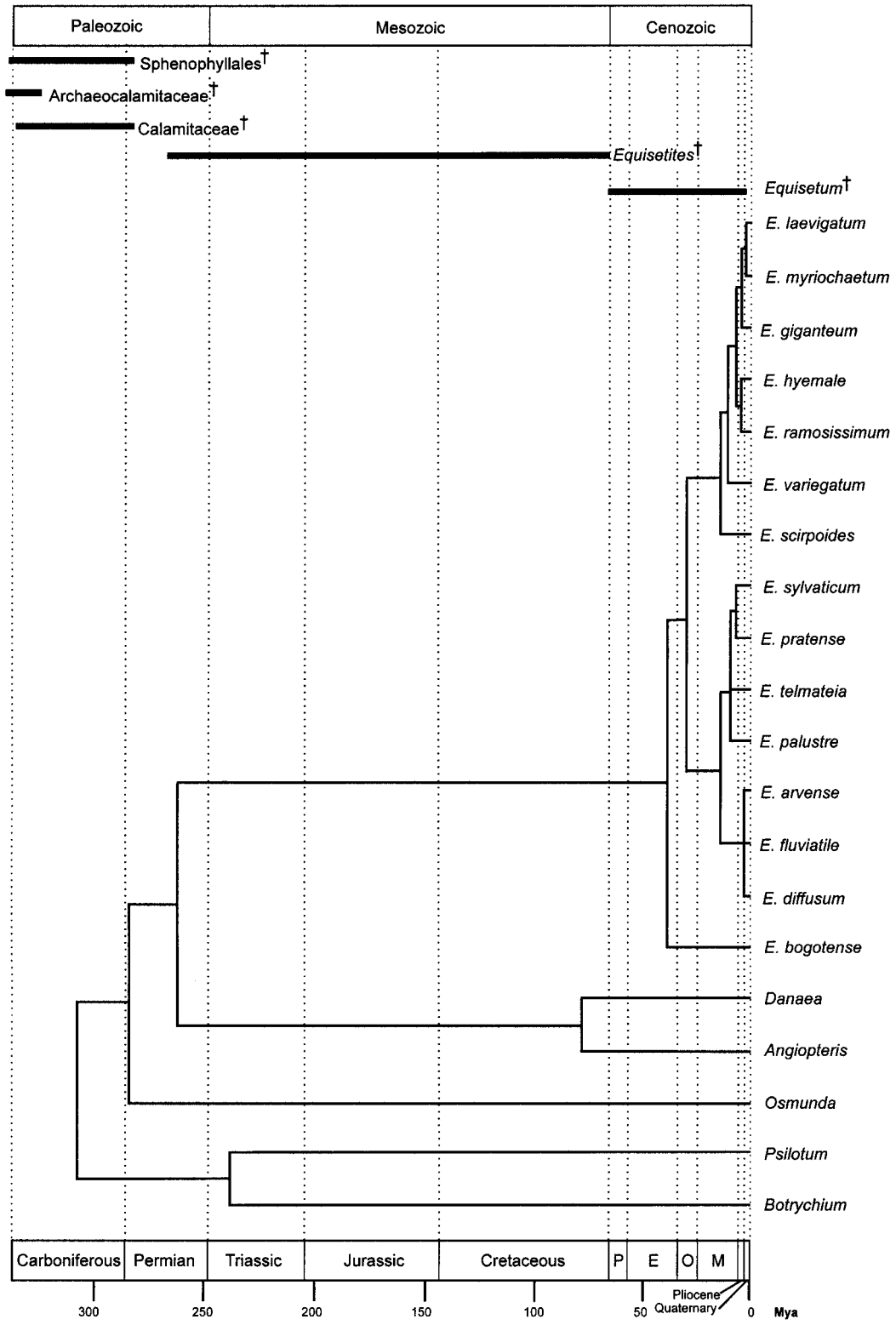

Fig. 5 The topology of the *Equisetum* crown group relationships from the maximum parsimony analysis of the *rbcL* data only (not shown in earlier figures) is presented as a chronogram calibrated against the geological timescale (Geological Society of America 1999). Geologic timescale abbreviations for the Tertiary: *P* = Paleocene, *E* = Eocene, *O* = Oligocene, *M* = Miocene. Divergence times indicated by the chronogram were estimated using penalized likelihood (Sanderson 2002*a*) employing the Truncated Newton algorithm with no internal age constraints enforced. Stratigraphic ranges for fossil *Equisetum* and fossil groups with possible affinities to *Equisetum* are indicated by thick solid bars and are based on dates in Stewart and Rothwell (1993). *Mya* = million years ago.

ognition of subg. *Equisetum* (*sensu* Hauke 1978) because of the phylogenetic lability of *Equisetum bogotense*. We suggest that its taxonomic status be reconsidered in future studies.

Although support for the other members of subg. *Equisetum* (minus *E. bogotense*) forming a monophyletic group was strong (94% bootstrap value) in our parsimony analysis (fig. 2), this node was not well supported in the likelihood analysis (62% bootstrap value; fig. 3). The generally low support for relationships within this clade stands in stark contrast to the highly supported phylogeny we obtained for subg. *Hippochaete*, regardless of the optimization criterion used (cf. figs. 2, 3). This finding is in agreement with Hauke's (1978, p. 398) observations that "in the subg. *Equisetum* there does not appear to be any obvious phylogeny." Because the species that make up subg. *Equisetum* are so well characterized and easily distinguished in comparison to those that make up subg. *Hippochaete*, it is indeed surprising that relationships among species in subg. *Equisetum* are so elusive compared to those in subg. *Hippochaete*. Though the support for relationships was weak, the topology within subg. *Equisetum* (minus *E. bogotense*) was identical in both the parsimony and likelihood analyses (figs. 2, 3). The sister relationship we observed between *Equisetum sylvaticum* and *Equisetum pratense* (two taxa previously regarded as closely related because of their semidimorphic habit, with fertile stems persisting rather than dying back) is the only alliance that is consistent with most former studies (Page 1972*a*; Hauke 1978). Our analyses provide high support for a single subclade (*Equisetum diffusum*, *Equisetum arvense*, *Equisetum fluviatile*) within subg. *Equisetum* (minus *E. bogotense*) (figs. 2, 3). Only Page (1972*a*) previously considered *E. arvense* and *E. fluviatile* to be closely related (he assigned them to his sect. *Equisetum*), in spite of their overall different growth habits, because of similarities in their micromorphology. This relationship is at odds with Schaffner (1925*b*, 1930*b*), who regarded *E. fluviatile* to be the most "primitive" species within subg. *Equisetum* and *E. arvense* to be the most advanced.

Both Hauke (1963) and Schaffner (1930*b*) considered the large, branched, Neotropical species *Equisetum giganteum* to be the most "primitive" *Equisetum* and proposed for the genus a Gondwanan origin, subsequent dispersal, and diversification in northern latitudes. The other large, branched, Neotropical species, *Equisetum myriochaetum*, was also regarded by Hauke (1963) and Schaffner (1930*a*, 1930*b*) as "primitive." Schaffner (1930*a*) in particular regarded the modern geographic distribution of *Equisetum* species to reflect, for the most part, their phylogenetic history, with the more derived taxa being more northerly. The only other Neotropical species, the diminutive *E. bogotense*, was the exception to Schaffner's (1930*a*) hypothesis, since he considered it to be more derived. Its geographic distribution was explained as a reinvasion of South America by the otherwise completely north-temperate subg. *Equisetum* (Hauke 1978). Molecular evidence presented here suggests that the ancestral condition in *Equisetum* was more likely a small, unbranched plant, and that the large size and regular branching observed in the Neotropical *E. giganteum* and *E. myriochaetum* is a derived condition.

Hybridization

Prior attempts at understanding *Equisetum* systematics, phylogeny, and character evolution have been hampered by the proliferation of morphological intermediates caused by rampant hybridization among members within each subgenus. We were very careful to avoid hybrids, yet the widespread and prolific sterile hybrid between *Equisetum laevigatum* and *Equisetum hyemale*, *Equisetum* × *ferrissii*, was a persistent nuisance in our sampling for this study in that it was often confused for one or the other parent (Moran 1983). In our final analyses we intentionally included two such samples (table 1) with interesting results. Though morphologically indistinguishable from each other, these two samples differed in their sampled chloroplast sequences: one is strongly supported as sister to *E. laevigatum* while the other is strongly supported as sister, and is identical in its sequence, to the other parent species, *E. hyemale* (figs. 2, 3). Because the chloroplast is maternally inherited, the hybrid sequence data may be an excellent indicator as to which species served as the maternal parent in any given hybrid (Gastony and Yatskievych 1992; Renzaglia et al. 2002). This observation may be useful in future studies that attempt to include *Equisetum* hybrids and predict parentage using a phylogenetic criterion.

Rates of Evolution

The increased rates of nucleotide substitution (due to transversions) that we observed in subg. *Hippochaete* were surprising and warrant further study (fig. 4). Several other studies have revealed a similar rate difference between two basally divergent clades at the *rbcL* locus (Lewis et al. 1997; Pryer et al. 2001*a*). One possible explanation for the significantly different substitution rates between the two subgenera might involve the subgeneric distinction of flush stomates (subg. *Equisetum*) versus sunken stomates (subg. *Hippochaete*) and how these "hardwired" structural differences might have differentially affected photosynthetic pathways, metabolic activity, and CO_2 availability in the two subgenera (Hibberd and Quick 2002; Raven 2002). This possibility of variation in the photosynthetic machinery within the genus is currently under study (K. M. Pryer, A. R. Smith, and T. S. Feild, unpublished data). The higher ratio of nonsynonymous substitutions (Ka) to synonymous substitutions (Ks) in subg. *Hippochaete* is also of interest, as it may indicate a relaxation of purifying selection or positive selection at the *rbcL* locus in subg. *Hippochaete*. Sampling more loci from both the nuclear and plastid genomes might allow for a more direct assessment of the evolutionary history accompanying the observed rate change in *rbcL* between these two subgenera.

An additional character correlated with the increased rate of substitution that we observed in subg. *Hippochaete* is genome size, as estimated by cell-flow cytometry (Grime et al. 1988; Obermayer et al. 2002). Sampled members of subg. *Equisetum* have C-values in the range of 12.5 to 14.2 pg of DNA per nucleus (Grime et al. 1988), whereas species in subg. *Hippochaete* have C-values ranging from 21.3 to 30.4 pg of DNA per nucleus (Obermayer et al. 2002), which may account for the larger chromosome size observed in subg. *Hippochaete* (Manton 1950). Without C-values for *E. bogotense* or for

closely related outgroups, and given that our current best knowledge for relationships at the base of *Equisetum* is a polytomy, we cannot polarize this character; that is, we cannot hypothesize whether the genome has expanded in subg. *Hippochaete* or contracted in subg. *Equisetum*. A possible explanation for the expansion theory is that subg. *Hippochaete* might have experienced a bottleneck after diverging from the rest of the genus, thereby exposing a large number of deleterious recessive mutations in its genome to selection. Segmental duplications and genome rearrangements might have been favored subsequently as a mechanism to "cover" these recessive mutations with new dominant modifiers, thereby also leaving a signal of increased rate of molecular evolution. Until there is better resolution of our current phylogeny, an equally compelling evolutionary argument could also be put forth to explain why selection might have favored the contraction theory and why subg. *Equisetum* may have jettisoned a large part of its genome.

Recent Divergence of Equisetum

One hypothesis we wished to address in pursuing a phylogeny of the *Equisetum* crown group was whether it represented a recent radiation or if it "may be regarded as the oldest of living vascular plants" (Arnold 1947, p. 151). Paleobotanists reserve the name *Equisetum* for Cenozoic fossils that cannot be distinguished from modern species (Stewart and Rothwell 1993). The genus *Equisetites* is usually applied to Mesozoic casts or compression-impression fossils that are very similar in their leaf sheath and sporangiophore arrangement to modern *Equisetum*, but that cannot (usually because of their considerably larger size) be assigned to a modern species. Because of this there has been some question as to whether the fossil genus might be congeneric with the extant genus (Hauke 1963).

A recent comparative study (Renzaglia et al. 2002) of the detailed architectural features of the sperm cells between *E. arvense* (subg. *Equisetum*) and *E. hyemale* (subg. *Hippochaete*) demonstrated that they are remarkably similar. The conservation of cellular features between the two subgenera may indicate a recent radiation for the genus. This observation agrees well with our dates (Tertiary) for diversification of the *Equisetum* crown group, estimated using penalized likelihood (fig. 5). The DNA sequence-based age estimates for the Equisetaceae total group in our unconstrained analyses indicate a Paleozoic origin (fig. 5), an estimate in remarkable accord with current interpretations of the fossil record. This lends support to the suggestion of an ancient relationship between Equisetaceae and Calamitaceae (Kenrick and Crane 1997*a*). The date of radiation for the *Equisetum* crown group in both our constrained and unconstrained analyses also corroborates the traditional placement of Cenozoic Equisetaceae taxa in *Equisetum*, reserving membership in *Equisetites* for Mesozoic fossils that are morphologically distinct from extant species. *Equisetum* likely diverged in the Tertiary from *Equisetites*. The age estimates from our unconstrained analyses for the Equisetaceae total group and *Equisetum* crown group nodes are also well within, or close to, the standard deviations of the maximum parsimony age estimates calculated by Soltis et al. (2002; their table 4, *rbcL*): Equisetaceae = 297.3 ± 43.1; *Equisetum* = 64.8 ± 12.7. Their ages were estimated using an earlier approach (nonparametric rate smoothing [NPRS]) devised by Sanderson (1997) to accommodate rate heterogeneity among lineages. The penalized likelihood approach (Sanderson 2002*a*), at least for the nodes of interest to us, appears to provide age estimates that are in line with how the fossil history has been interpreted for the group.

Because we sampled all the living species from the Cenozoic crown group radiation, there are no additional taxa that can be included (from a molecular perspective) to reduce the very long branch that leads from *Equisetum* to other vascular plants. This further emphasizes the analytical challenges one faces in determining the relationships of *Equisetum* to other plants.

Morphological Character Evolution in Equisetum

Our gene-estimated species tree allows an independent hypothesis with which we can address morphological character evolution in *Equisetum*. Here we consider several morphological characters thought by past workers to be of phylogenetic utility. For example, several sporophytic characters have been used to split the genus into subgenera *Equisetum* and *Hippochaete*. The position of the stomatal apparatus—sunken into the epidermis in subg. *Hippochaete* and flush with the epidermis in subg. *Equisetum*—shows good phylogenetic correlation, as does chromosome size (larger in subg. *Hippochaete*; Manton 1950). Interestingly, there are also clear-cut distinctions in gametophyte morphology between the two subgenera. Column lamellae and sunken antheridia with two opercular cells characterize subg. *Hippochaete*, whereas unistratose plates and projecting antheridia, generally with more than two opercular cells, prevail in subg. *Equisetum* (Duckett 1979*a*). The absence of known hybrids between species of the two subgenera further substantiates this subgeneric division (Hauke 1974; Duckett 1979*b*).

There are no known hybrids involving *E. bogotense*, an observation that has been attributed to the absence of any sympatric members of subg. *Equisetum* in its native South and Central American habitats (Duckett 1979*b*). Our results suggest that *E. bogotense* is phylogenetically the most isolated species among extant congeners and might, therefore, be incapable of hybridization with other species in the genus. The ambiguous position of *E. bogotense* is the only ingroup difference between the parsimony and likelihood analyses of the combined *rbcL* + *trnL-F* data set. Because the phylogenetic relationship of *E. bogotense* was not resolved in this study, we summarize the topology used to reconstruct character-state evolution (fig. 6) as a basal trichotomy among *E. bogotense*, subg. *Hippochaete*, and subg. *Equisetum*.

Persistence of leaf sheath teeth was regarded by Hauke (1963) as a "primitive" character for the genus, and our study confirms that it is indeed plesiomorphic (fig. 6*A*). *Equisetum bogotense*, subg. *Equisetum*, and the basal members of subg. *Hippochaete* (*Equisetum scirpoides* and *Equisetum variegatum*) all retain their leaf sheath teeth throughout the life of the upright stems. A transition to withered or caducous sheath teeth took place among the derived species of subg. *Hippo-*

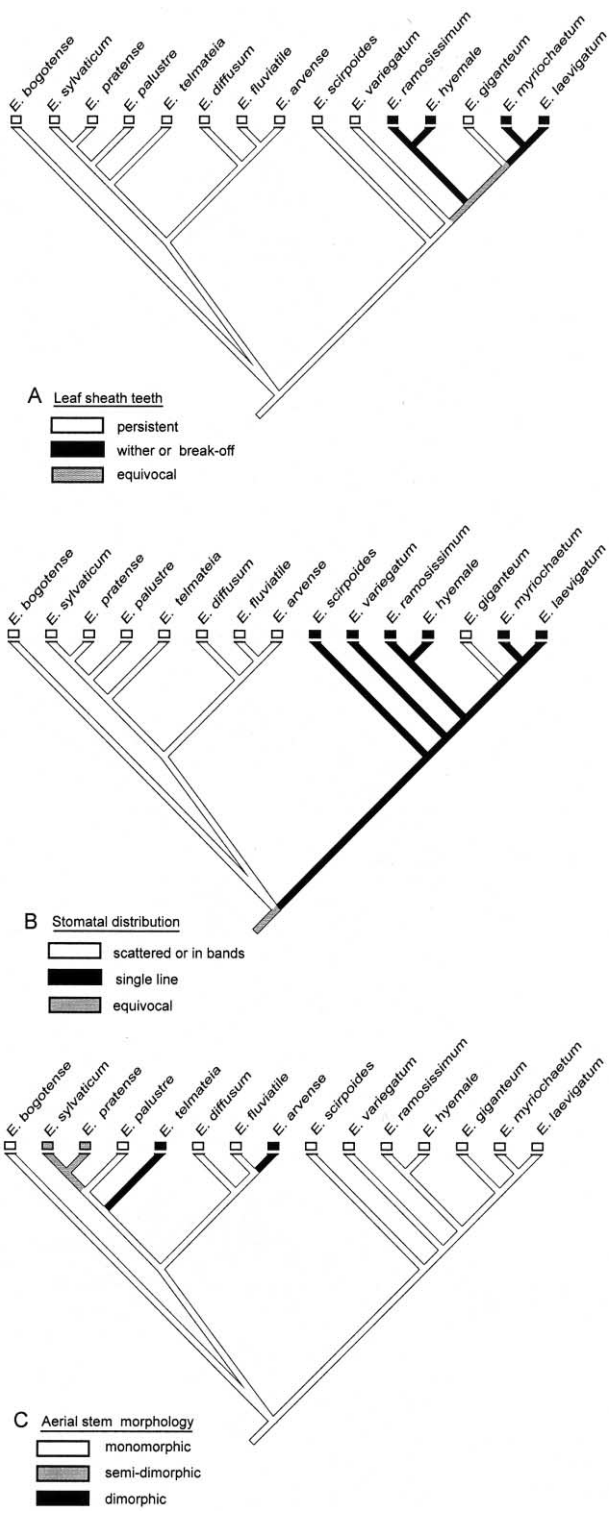

Fig. 6 Evolution of selected sporophyte characters in *Equisetum*

chaete (*Equisetum ramosissimum* to *E. laevigatum* clade). The exact nature of the character transition is ambiguous: either sheath teeth were lost independently in the clades (*E. ramosissimum* + *E. hyemale*) and (*E. myriochaetum* + *E. laevigatum*), or they were lost only once at the base of the (*E. ramosissimum* to *E. laevigatum*) clade, and there was a reversal to persistent teeth in *E. giganteum*. With the exception of *E. giganteum*, all species in the (*E. ramosissimum* to *E. laevigatum*) clade usually shed their leaf sheath teeth, sometimes by an abscission layer (*E. hyemale*), either prior to maturity or on mature plants. This character shows considerable variation in *E. ramosissimum*, with the sheath teeth in subsp. *debile* drying and regularly caducous and those in subsp. *ramosissimum* withering but not always shed.

The persistent sheath teeth in *E. giganteum* was one of several homoplastic characters that misled Hauke (1963) into believing that this species was the most "primitive" member of *Equisetum* (fig. 6*A*). Hauke (1963) also proposed that stomatal distribution in the stem grooves was indicative of phylogenetic position. He postulated that stomates in bands or scattered was the "primitive" condition and that stomates in single files was advanced. Mapping this character on our phylogeny indicates that it correlates almost perfectly with subgeneric circumscription, the sole exception being *E. giganteum* (fig. 6*B*). Its large size, persistent sheath teeth, and stomates in bands led Hauke (1963) and Schaffner (1925*b*, 1930*a*, 1930*b*) to believe that these were good indicators of a species not so far removed from fossil members of the genus and therefore of early origin. *Equisetum giganteum*, however, is well nested within subg. *Hippochaete*, and these sheath teeth and stomatal distribution characters are clearly homoplastic.

Four species of *Equisetum* (all within subg. *Equisetum*) exhibit some degree of aerial stem dimorphism (Hauke 1993), producing both vegetative and fertile stems. *Equisetum arvense* and *Equisetum telmateia* are dimorphic with green, branched, and persistent vegetative stems and fertile stems that are tan (achlorophyllous), lacking stomates, unbranched, and ephemeral. *Equisetum sylvaticum* and *E. pratense* are semi-dimorphic with green, branched, and persistent vegetative stems and fertile stems that are tan, with stomates, initially unbranched, and persisting to become branched and green after spore discharge. Schaffner (1925*b*, 1930*a*, 1930*b*) placed considerable importance on aerial stem dimorphism in his classification of subg. *Equisetum* with his sect. *Heterophyadica* including all four of these species. Although he downplayed the importance of dimorphism, Hauke (1978) maintained sect. *Heterophyadica* in his monograph of subg. *Equisetum* but excluded *E. telmateia* from it based on other characters that he determined aligned it more closely with members of sect. *Equisetum*. In an earlier study, Page (1972*a*) had dismissed dimorphism and semi-dimorphism as being unreliable phylogenetic markers, an assessment slightly at odds with his inclusion of the two semi-dimorphic species, *E. pratense* and *E. sylvaticum*, as the sole members of his sect. *Subvernalia*. Our analyses suggest that semi-dimorphism is, in fact, a synapomorphy for this clade (fig. 6*C*). Though this node is weakly supported in our analyses, this grouping was supported by Page (1972*a*) on the basis of micromorphological characters of the epidermis and by Duckett (1979*a*, 1979*b*) on the basis of similarities in gametophyte morphology and the successful synthesis of hybrids between the two species. The fertile stems of the two dimorphic taxa, *E. arvense* and *E. telmateia*, produce no stomates (Hauke 1978), hence their inability to photosynthesize and persist. Our molecular phylogeny places *E. arvense* in a different clade from *E. telmateia*, with strong bootstrap support, suggesting that

dimorphism and the loss of stomates on fertile stems are homoplastic (fig. 6*C*).

Perhaps one of the most fascinating and widely studied aspects of *Equisetum* biology is the gametophyte and its highly varied sexual behavior (Hauke 1963, 1969, 1978, 1980; Duckett 1970, 1972, 1977, 1979*a*, 1979*b*; Duckett and Duckett 1980; Duckett and Pang 1984). The gametophyte is terrestrial, thalloid, and photosynthetic. Though there is debate on the sexual sequence of individual species (Hauke 1963, 1969; Duckett and Pang 1984), all extant taxa produce gametophytes that initially are unisexual, producing either antheridia or archegonia (Duckett 1970, 1972, 1979*a*, 1979*b*); subsequently, the females produce antheridia and give rise to bisexual gametophytes (to varying degrees) by lamellar proliferation (Duckett and Pang 1984). The tendency for initially female gametophytes to become rapidly bisexual is apparently more marked in subg. *Equisetum* than in subg. *Hippochaete* (Duckett 1972). Archegonial development on these bisexual gametophytes eventually ceases, and the initially female gametophytes ultimately become functionally male (Hauke 1978). The formation of archegonia on initially male gametophytes is a rarer event in species in both subgenera and is strongly influenced by the culture medium, the age of the parent tissue, and the particular species (Duckett 1972). *Equisetum bogotense* is alone in having dimorphic gametophytes that apparently remain strictly unisexual (Hauke 1969), though Duckett (1972) suggested that Hauke's experiments may have been too short for the full sexual behavior to become manifest.

Hauke's early observations (1963, 1969) that *E. giganteum* produced gametophytes that form archegonia and antheridia at the same time and from the same region of the meristem (i.e., initially bisexual) provided additional fuel for his hypothesis that this was a "primitive species" (he presumed that unisexuality was a derived condition). In their experiments, Duckett and Pang (1984) clearly demonstrated that the gametophytes of *E. giganteum* follow the same pattern in their sexual behavior as observed in other species of subg. *Hippochaete*, thereby refuting the rationale used by Hauke. The nested position of *E. giganteum* within subg. *Hippochaete* in our phylogeny (figs. 2, 3) adds support to this argument by Duckett and Pang (1984).

Conclusion

Our findings from DNA sequence analysis point to the need for a thorough reexamination of morphological characters and character states in *Equisetum*, an analysis that ideally should include fossils as well. The interpretation of many characters in *Equisetum*, including cones, leaves, and other structures, remains difficult because of their uniqueness among tracheophytes, a paucity of developmental studies, and the consequent difficulty of assessing homologies. It is our hope that hypotheses of species relationships, such as we present here, are a step toward eventually better understanding and interpreting morphological evolution in this fascinating group of plants.

Acknowledgments

We thank F. Lutzoni for advice with *trnL-F* sequence alignment, Bayesian analyses, and likelihood ratio tests; S. Zoller for advice on implementing analyses using UNIX; J. Hunt and A. Driskell for advice with molecular techniques; R. Cranfill, R. Foster, P. Hammond, B. Hoshizaki, R. Moran, W. Speer, and P. G. Wolf for providing dried, vouchered specimens and/or genomic DNA; and M. W. Hahn for stimulating discussions. E. A. Zimmer and J. E. Skog kindly permitted us access to their unpublished *trnL-F* sequence data for *Osmunda*, *Angiopteris*, and *Ophioglossum*; although we did not use these data in our analyses because of the alignment difficulties they posed with our data set, they were useful initially. We are grateful to two anonymous reviewers for their useful comments on the manuscript. This work was supported by National Science Foundation grants DEB-9675533 and DBI-9871374 to K. M. Pryer and DEB-9616260 to A. R. Smith. D. L. Des Marais was supported in part by a summer (2000) internship from the Field Museum of Natural History while an undergraduate student at the University of California, Berkeley.

Literature Cited

Arnold CA 1947 An introduction to paleobotany. McGraw-Hill, New York.

Bateman RM 1991 Palaeobiological and phylogenetic implications of anatomically-preserved *Archaeocalamites* from the Dinantian of Oxroad Bay and Loch Humphrey Burn, southern Scotland. Palaeontogr B 223:1–59.

Bendich AJ, RS Anderson 1983 Repeated DNA sequences and species relatedness in the genus *Equisetum*. Plant Syst Evol 143:47–52.

Berry CM, M Fairon-Demaret 2001 The Middle Devonian flora revisited. Pages 120–139 *in* PG Gensel, D Edwards, eds. Plants invade the land: evolutionary and environmental perspectives. Columbia University Press, New York.

Bousquet JS, SH Strauss, AH Doerksen, RA Price 1992 Extensive variation in evolutionary rate of *rbcL* gene sequences among seed plants. Proc Natl Acad Sci USA 89:7844–7848.

Crane PR 1989 Patterns of evolution and extinction in vascular plants. Pages 153–187 *in* K Allen, D Briggs, eds. Evolution and the fossil record. Belhaven, London.

DiMichele WA, RW Hook 1992 Paleozoic terrestrial ecosystems. Pages 204–325 *in* AK Behrensmeyer, JD Damuth, WA DiMichele, R Potts, H-D Sues, SL Wing, eds. Terrestrial ecosystems through time: evolutionary paleoecology of terrestrial plants and animals. University of Chicago Press, Chicago.

DiMichele WA, WE Stein, RM Bateman 2001 Ecological sorting of vascular plant classes during the Paleozoic evolutionary radiations. Pages 285–335 *in* WD Allmon, DJ Bottjer, eds. Evolutionary paleoecology: the ecological context of macroevolutionary change. Columbia University Press, New York.

Dubuisson J-Y 1997 *RbcL* sequences: a promising tool for the molecular systematics of the fern genus *Trichomanes* (Hymenophyllaceae)? Mol Phylogenet Evol 8:128–138.

Duckett JG 1970 Sexual behavior of the genus *Equisetum*, subgenus *Equisetum*. Bot J Linn Soc 63:327–352.

——— 1972 Sexual behavior of the genus *Equisetum*, subgenus *Hippochaete*. Bot J Linn Soc 65:87–108.
——— 1977 Towards an understanding of sex determination in *Equisetum*: an analysis of regeneration in gametophytes of the subgenus *Equisetum*. Bot J Linn Soc 74:215–242.
——— 1979*a* Comparative morphology of the gametophytes of *Equisetum* subgenus *Hippochaete* and the sexual behaviour of *E. ramosissimum* subsp. *debile* (Roxb.) Hauke, *E. hyemale* var. *affine* (Engelm.) A. A. Eaton, and E. *laevigatum* A. Br. Bot J Linn Soc 79: 179–205.
——— 1979*b* An experimental study of the reproductive biology and hybridization in the European and North American species of *Equisetum*. Bot J Linn Soc 79:205–229.
Duckett JG, AR Duckett 1980 Reproductive biology and population dynamics of wild gametophytes of *Equisetum*. Bot J Linn Soc 80: 1–40.
Duckett JG, WC Pang 1984 The origins of heterospory: a comparative study of sexual behavior in the fern *Platyzoma microphyllum* and the horsetail *Equisetum giganteum*. Bot J Linn Soc 88:11–34.
Farris JS, M Källersjo, AG Kluge, C Bult 1995 Testing significance of incongruence. Cladistics 10:315–319.
Farwell OA 1916 The genus *Hippochaete* in North America, north of Mexico. Mem N Y Bot Gard 6:461–472.
Felsenstein J 1981 A likelihood approach to character weighting and what it tells us about parsimony and compatibility. Biol J Linn Soc 16:183–196.
Gastony GJ, G Yatskievych 1992 Maternal inheritance of the chloroplast and mitochondrial genomes in cheilanthoid ferns. Am J Bot 79:716–722.
Gaut BS, SV Muse, MT Clegg 1992 Relative rates of nucleotide substitution at the *rbcL* locus of monocotyledonous plants. J Mol Evol 35:292–303.
Geological Society of America 1999 Geologic time scale. Product code CTS004. AR Palmer, J Geissman, compilers.
Grime JP, JG Hodgsen, R Hunt 1988 Comparative plant ecology: a functional approach to common British species. Unwin Hyman, London.
Hauke R 1963 A taxonomic monograph of the genus *Equisetum* subgenus *Hippochaete*. Cramer, Weinheim.
——— 1969 Gametophyte development in Latin American horsetails. Bull Torrey Bot Club 96:568–577.
——— 1974 The taxonomy of *Equisetum*: an overview. New Bot 1: 89–95.
——— 1978 A taxonomic monograph of *Equisetum* subgenus *Equisetum*. Nova Hedwigia 30:385–455.
——— 1980 Gametophytes of *Equisetum diffusum*. Am Fern J 70: 39–44.
——— 1983 Horsetails (*Equisetum*) in North America. Fiddlehead Forum 10:39–42.
——— 1993 Equisetaceae (Horsetail Family). Pages 76–84 *in* Flora of North America Editorial Committee, ed. Flora of North America north of Mexico. Vol 2. Oxford University Press, New York.
Hibberd JM, WP Quick 2002 Characteristics of C_4 photosynthesis in stems and petioles of C_3 flowering plants. Nature 415:451–454.
Huelsenbeck JP 2000 MrBayes: Bayesian inference of phylogeny. Computer program distributed by the author, Rochester, N.Y. http://morphbank.ebc.uu.se/mrbayes/
Karol KG, RM McCourt, MT Cimino, CF Delwiche 2001 The closest living relatives of land plants. Science 294:2351–2353.
Kelber K-P, JHA van Konijnenburg-van Cittert 1998 *Equisetites arenaceus* from the Upper Triassic of Germany with evidence for reproductive strategies. Rev Palaeobot Palynol 100:1–26.
Kenrick P, PR Crane 1997*a* The origin and early diversification of land plants: a cladistic study. Smithsonian Institution, Washington, D.C.
——— 1997*b* The origin and early evolution of plants on land. Nature 389:33–39.
Lewis LA, BD Mishler, R Vilgalys 1997 Phylogenetic relationships of the liverworts (Hepaticae), a basal embryophyte lineage, inferred from nucleotide sequence data of the chloroplast gene *rbcL*. Mol Phylogenet Evol 7:377–393.
Lutzoni F, P Wagner, V Reeb, S Zoller 2000 Integrating ambiguously aligned regions of DNA sequences in phylogenetic analyses without violating positional homology. Syst Biol 49:628–651.
Maddison DR, WP Maddison 2000 MacClade 4. Sinauer, Sunderland, Mass.
Manton I 1950 Problems of cytology and evolution in the Pteridophyta. Cambridge University Press, Cambridge.
Milde J 1865 Repräsentieren die Equiseten der gegenwärtigen Schöpfungsperiode ein oder zwei Genera? Bot Z 23:297–299.
Moran RC 1983 *Equisetum* × *ferrissii* (Equisetaceae) in Illinois. Castanea 48:79–82.
Muse SV, BS Weir 1992 Testing for equality of evolutionary rates. Genetics 132:269–276.
Obermayer R, IJ Leitch, L Hanson, MD Bennett 2002 Nuclear DNA C-values in 30 species double the familial representation in pteridophytes. Ann Bot 90:209–217.
Page CN 1972*a* An assessment of inter-specific relationships in *Equisetum* subgenus *Equisetum*. New Phytol 71:355–369.
——— 1972*b* An interpretation of the morphology and evolution of the cone and shoot of *Equisetum*. Bot J Linn Soc 65:359–397.
Pond SLK, S Muse 2001 HYPHY. Computer program distributed by the authors. http://peppercat.stat.ncsu.edu/~hyphy/
Posada D, KA Crandall 1998 Modeltest: testing the model of DNA substitution. Bioinformatics 14:817–818.
Pryer KM, H Schneider, S Magallón 2004 The radiation of vascular plants. Chapter 10 *in* J Cracraft, MJ Donoghue, eds. Assembling the tree of life. Oxford University Press, London (in press).
Pryer KM, H Schneider, AR Smith, R Cranfill, PG Wolf, JS Hunt, S Sipes 2001*b* Horsetails and ferns are a monophyletic group and the closest living relative to seed plants. Nature 409:618–622.
Pryer KM, AR Smith, JS Hunt, JY Dubuisson 2001*a* *RbcL* data reveal two monophyletic groups of filmy ferns (Filicopsida: Hymenophyllaceae). Am J Bot 88:1118–1130.
Raven JA 2002 Evolutionary options. Nature 415:375–377.
Renzaglia KS, SB Dengate, SJ Schmitt, JG Duckett 2002 Novel features of *Equisetum arvense* spermatozoids: insights into pteridophyte evolution. New Phytol 154:159–174.
Sanderson MJ 1997 A nonparametric approach to estimating divergence times in the absence of rate constancy. Mol Biol Evol 14: 1218–1231.
——— 2002*a* Estimating absolute rates of molecular evolution and divergence time: a penalized likelihood approach. Mol Biol Evol 19: 101–109.
——— 2002*b* r8s, version 1.5. Computer program distributed by the author. http://ginger.ucdavis.edu/r8s/
Sanderson MJ, A Purvis, C Henze 1998 Phylogenetic supertrees: assembling the trees of life. Trends Ecol Evol 13:105–109.
Schaffner JH 1921 North American species of *Equisetum* north of Mexico. Am Fern J 11:65–75.
——— 1925*a* Main lines of evolution in *Equisetum*. I. Am Fern J 15: 8–12.
——— 1925*b* Main lines of evolution in *Equisetum*. II. Am Fern J 15:35–39.
——— 1930*a* Diagnostic analysis and phylogenetic relationship of the main groups of *Equisetum*. Am Fern J 20:11–18.
——— 1930*b* Geographic distribution of the species of *Equisetum* in relation to their phylogeny. Am Fern J 20:89–106.
Skog JE, HP Banks 1973 *Ibyka amphikoma*, gen. et sp. n., a new

protoarticulate from the Late Middle Devonian of New York state. Am J Bot 60:366–380.

Soltis PS, DE Soltis, V Savolainen, PR Crane, TG Barraclough 2002 Rate heterogeneity among lineages of tracheophytes: integration of molecular and fossil data and evidence for molecular living fossils. Proc Nat Acad Sci USA 99:4430–4435.

Stein WE, DC Wight, CB Beck 1984 Possible alternatives for the origin of Sphenopsida. Syst Bot 9:102–118.

Stewart WN, GW Rothwell 1993 Paleobotany and the evolution of plants. Cambridge University Press, Cambridge.

Swofford DL 2000 PAUP*: phylogenetic analysis using parsimony (*and other methods). Sinauer, Sunderland, Mass.

Taberlet P, L Gielly, G Pautou, J Bouvet 1991 Universal primers for amplification of three non-coding regions of chloroplast DNA. Plant Mol Biol 17:1105–1109.

Taylor TN, EL Taylor 1993 The biology and evolution of fossil plants. Prentice Hall, Englewood Cliffs, N.J.

Wagner WH 1961 Problems in the classification of ferns. Recent Adv Bot 1:841–844.

——— 1980 Origin and philosophy of the groundplan-divergence method of cladistics. Syst Bot 5:173–193.

Wheeler WC 1990 Combinatorial weights in phylogenetic analysis: a statistical parsimony procedure. Cladistics 6:269–275.