American Journal of Botany 86(3): 372–386. 1999.

PHYLOGENETIC RELATIONSHIPS OF LAND PLANTS USING MITOCHONDRIAL SMALL-SUBUNIT rDNA SEQUENCES[1]

R. JOEL DUFF AND DANIEL L. NICKRENT[2]

Department of Plant Biology, Southern Illinois University, Carbondale, Illinois 62901-6509

Phylogenetic relationships among embryophytes (tracheophytes, mosses, liverworts, and hornworts) were examined using 21 newly generated mitochondrial small-subunit (19S) rDNA sequences. The ''core'' 19S rDNA contained more phylogenetically informative sites and lower homoplasy than either nuclear 18S or plastid 16S rDNA. Results of phylogenetic analyses using parsimony (MP) and likelihood (ML) were generally congruent. Using MP, two trees were obtained that resolved either liverworts or hornworts as the basal land plant clade. The optimal ML tree showed hornworts as basal. That topology was not statistically different from the two MP trees, thus both appear to be equally viable evolutionary hypotheses. High bootstrap support was obtained for the majority of higher level embryophyte clades named in a recent morphologically based classification, e.g., Tracheophyta, Euphyllophytina, Lycophytina, and Spermatophytata. Strong support was also obtained for the following monophyletic groups: hornworts, liverworts, mosses, lycopsids, leptosporangiate and eusporangiate ferns, gymnosperms and angiosperms. This molecular analysis supported a sister relationship between *Equisetum* and leptosporangiate ferns and a monophyletic gymnosperms sister to angiosperms. The topologies of deeper clades were affected by taxon inclusion (particularly hornworts) as demonstrated by jackknife analyses. This study represents the first use of mitochondrial 19S rDNA for phylogenetic purposes and it appears well-suited for examining intermediate to deep evolutionary relationships among embryophytes.

Key words: 19S rDNA; embryophyte; hornwort; liverwort; lycophyte; molecular phylogeny; ribosomal RNA; tracheophyte.

The origins and affinities of the major lineages of land plants have remained one of the fundamental questions in plant evolutionary biology. In spite of considerable effort from researchers using varied approaches, the early phylogenetic history of the land plants (embryophytes) is a major unresolved problem. During the past decade, understanding the relationships among the major lineages, i.e., tracheophytes, mosses, hornworts, and liverworts, has been advanced by comparative studies using characters from morphology (Mishler et al., 1994; Kenrick and Crane, 1997) and development (Garbary, Renzaglia, and Duckett, 1993; Garbary and Renzaglia, 1998) as well as results from molecular phylogenetic analyses. For the latter, data have been obtained from genes in the nucleus (Waters et al., 1992; Mishler et al., 1994; Kranz et al., 1995; Hedderson, Chapman, and Rootes, 1996), chloroplast (Manhart, 1994, 1995; Crowe et al., 1997; Lewis, Mishler, and Vilgalys, 1997), and mitochondrion (Malek et al., 1996).

Until recently, mitochondrial genes have been relatively ignored by plant systematists as a source of phylogenetic information. This is not unexpected given that mtDNA genome organization is highly dynamic (e.g., gene order is poorly conserved), thereby thwarting attempts to use restriction fragment length polymorphism analysis beyond the familial and even generic level (Palmer, 1992). In addition, the use of protein-coding mtDNA gene sequences for resolving angiosperm phylogeny has been hindered by substitution rates lower than that of nuclear and chloroplast genes. Based upon small-subunit rDNA sequence comparisons in angiosperms, it is generally recognized that the following trend in sequence variability exists: 18S (nuclear) $>$ 16S (plastid) $>$ 19S (mitochondrion) (Palmer et al., 1990). This trend was confirmed for photosynthetic flowering plants by Duff and Nickrent (1997) who characterized the 19S rDNAs of several nonphotosynthetic (holoparasitic) and photosynthetic angiosperms. In addition to documenting low sequence variability among the latter, this work showed that, among the holoparasitic plants, there existed increased substitution rates, transversion biases, and novel higher order rRNA structural features. Given the above trend in variability among rDNA genes, it might be expected that mitochondrial 19S rDNA sequences would provide even fewer variable sites than chloroplast 16S rDNA. This is not the case as is shown by comparing the ''core'' 19S rDNA sequence (see below) of *Marchantia* to that of an angiosperm such as *Zea* where pairwise genetic distances for 19S are at least twice that of 16S rDNA. Therefore, we reasoned that 19S rDNA may provide a useful phylogenetic marker for examining early divergences in land plants.

Congruent topologies obtained from multiple independent sources of data are considered the strongest support for phylogenetic relationships (Miyamoto and Cracraft, 1991; Penny, Hendy, and Steel, 1991). Given that relationships among the major groups of land plants remain

[1] Manuscript received 21 April 1998; revision accepted 8 September 1998.

The authors thank Barbara Crandall-Stotler, James Thompson, and Andrew Wood for providing living specimens used in this study, Karen Renzaglia for insightful discussions on embryophyte phylogeny, Carey Krajewski for helpful advice on data analysis, and David Swofford for granting permission to publish results obtained using the prerelease version of PAUP*. This study was supported in part by National Science Foundation Grant DEB-9407984 to DLN.

[2] Author for correspondence (phone: 618-453-3223; FAX: 618-453-3441; e-mail: nickrent@plant.siu.edu).

uncertain, the introduction of additional data is needed to test phylogenetic hypotheses and generate classifications. The broad objectives of this study were to: (1) delimit the phylogenetically useful partitions present in 19S rDNA sequences among all major land plant lineages, (2) characterize the amount and type of sequence variation present within this gene, (3) generate phylogenetic trees from the aligned sequences, (4) compare the results of these phylogenetic analyses to those obtained from other genes, and (5) place these results in context with the other hypotheses concerning land plant evolutionary relationships. Specifically, we will determine the utility of mitochondrial 19S rDNA in addressing a number of fundamental questions in plant phylogeny such as: (1) which land plant lineage is most basal among embryophytes? (2) are mosses the sister group to tracheophytes? (3) are bryophytes monophyletic? (4) are lycophytes the most basal member of the tracheophyte lineage? and (5) what are the relationships among the pteridophytes, a group widely recognized to be paraphyletic? This study includes 21 newly generated mitochondrial 19S rDNA sequences and represents the first use of this phylogenetically conservative molecular marker to address relationships among embryophytes.

MATERIALS AND METHODS

Representatives of the major land plant lineages were chosen for mitochondrial 19S rDNA sequencing (Table 1). In addition, sequences of two angiosperms (*Zea* and *Glycine*), a liverwort (*Marchantia*), and an algal outgroup (*Prototheca*) were obtained from GenBank. Other than the above sequences (and additional angiosperms), no other complete 19S rDNA sequences of embryophytes are present in GenBank. Attempts to amplify 19S rDNA from *Coleochaete* by the polymerase chain reaction (PCR) were unsuccessful, hence this charophyte sequence was not used as an outgroup. Instead, the sequence from the nonphotosynthetic green alga *Prototheca wickerhamii* was used.

Total genomic DNA was extracted from fresh tissue samples as described by Nickrent (1994). Primers for PCR amplification and sequencing of plant mitochondrial 19S rDNA are shown in Table 2. Standard double-stranded PCR amplifications were performed (Nickrent, 1994) using the following reaction conditions: preincubation at 94°C for 3 min followed by 35 cycles, each consisting of 1 min at 94°C, 1 min at 50°C, and 2.5 min at 72°C plus 2 s added to each extension step. The PCR products were gel purified using diethylaminoethyl (DEAE) membranes as described (Nickrent, 1994). The 19S rDNA products were sequenced directly using Sequenase v.2.0 (U.S. Biochemical, Cleveland, Ohio) or SequiTherm Excel II (Epicentre Technologies Corp., Madison, Wisconsin). Attempts to obtain 19S rDNA sequences from some taxa met with only partial success or failure. These included *Coleochaete* sp. (Coleochaetales), *Psilotum nudum* (Psilotales), *Ephedra antisyphilitica* and *Gnetum leyboldii* (Gnetales), and *Selaginella* spp. (Selaginales). For *Selaginella,* PCR products were not obtained despite attempted amplifications from genomic DNAs derived from four species. Partial sequences were obtained for *Psilotum, Ephedra,* and *Gnetum,* but acquisition of complete sequences was precluded by the presence of large introns and sequence divergence at primer positions. The 21 newly generated 19S rDNA sequences are deposited in Genbank under accession numbers listed in Table 1.

An initial alignment of *Prototheca, Marchantia,* and representative angiosperms was obtained from the Ribosomal Database Project web site (http://www.cme.msu.edu/RDP). This alignment was imported into SeqApp (Gilbert, 1993) and the newly generated sequences added. Given the presence of novel insertions and deletions in the new sequences, higher order structural models were required to guide alignment. At present, the only published higher order mitochondrial 19S rRNA structure of a plant present in the Ribosomal RNA Comparative Structural Database (http://pundit.colorado.edu:8080/RNA/16S/16s.html) is that of *Zea* (Gutell, 1994). A structural model was also proposed by the authors of the *Prototheca* 19S rDNA sequence (Wolff and Kück, 1990). This green algal taxon was selected over *Chlamydomonas* because the latter shows several major deletions that precluded its use in multiple sequence alignments. Structural models were prepared for *Phaeoceros laevis* (Fig. 1) as well as the following other land plant taxa: *Glycine, Zea, Pinus, Adiantum, Equisetum, Isoëtes, Sphagnum,* and *Marchantia.* These models are available as PDF files at the following web site: http://www.science.siu.edu/landplants/rRNA/rRNA.html.

Portions of the 19S rDNA alignment were excluded from analysis. All land plant mitochondrial 19S rRNA sequences examined contain regions, variable in sequence and in length, associated with helix 6 and 43, i.e., the V1 and V7 domains (Duff and Nickrent, 1997), hence these were not included in the data matrix. These regions correspond to positions 71–240 and 1214–1447 on the *Phaeoceros* model (Fig. 1). A region corresponding to positions 307–309 on *Phaeoceros* varies widely in length among the sampled land plants [0–119 nucleotides (nt); helix 10] and was therefore excluded. Positions 916–935 (V5 region) varies from 3 to 117 nt and positions 1767–1803 (V9 region) varies from 25 to 57 nt, hence these were similarly excluded. Nucleotides corresponding to positions 699–725 on *Phaeoceros* were only alignable among embryophytes because the *Prototheca* sequence contains an additional 42 nt in this region. This insertion precluded unambiguous alignment with the outgroup but was treated as missing data, hence was only used to infer ingroup relationships. Our results also identified several putative introns in 19S rDNA that were excluded from analysis. In *Sphagnum, Takakia, Atrichum,* and *Isoëtes* this variable-length intron occurs following nt 870 on the *Phaeoceros* model. The previously characterized 2.4-kb intron of *Marchantia* (following nt 998) was not included. Finally, the extreme 5′ and 3′ ends of the molecule (nt 1–40 and 1862–1901, respectively) corresponding to PCR priming sites were eliminated from phylogenetic analysis. The final alignment can be obtained at http://www.science.siu.edu/landplants/Alignments/Alignments.html.

To estimate the amount of phylogenetic signal in the 19S rDNA data, the skewness test (Hillis and Huelsenbeck, 1992) was implemented using the RANDOM TREES (10 000 trees) option in PAUP*. Phylogenies were reconstructed using maximum parsimony (MP) and maximum likelihood (ML) as implemented in the test version of PAUP* version 4.0d63 (by permission of D. Swofford, Smithsonian Institution, personal communication). MP analyses were conducted on a 100 MHz Power Macintosh 8100 and ML analyses on a 266 MHz Power Macintosh G3. For MP analyses, the full 27-taxon matrix was analyzed using the heuristic search option (MULPARS on, with TBR branch-swapping, gaps coded as missing data). The topologies of the eight resulting trees differed only within the angiosperms and ferns. A shortened 20-taxon data set was then produced by removing the sequences of *Nicotiana, Lindera, Glycine, Phegopteris, Diplazium, Adiantum,* and *Huperzia.* Given the similarity of these to their respective exemplars, the exclusion of these sequences did not affect global topological relationships (i.e., the resulting 20-taxon tree was fully congruent with that derived from the 27-taxon matrix). The 20-taxon matrix was then used in MP searches using the branch-and-bound option. The resulting strict consensus tree (of two equally parsimonious, minimum-length trees—A and B) was then used to estimate parameters for ML searches. The resulting empirical nucleotide frequencies were: A = 0.26864, C = 0.21791, G = 0.29365, and T = 0.21980. Furthermore, the transition/transversion ratio used was 1.6286 (kappa = 3.1616), the estimated proportion of invariant sites was 0.1948, the starting branch lengths were obtained using the Rogers-Swofford approximation method, and the estimated value for the gamma shape parameter was 0.7071 (four rate categories). Given these parameters, a ML tree corresponding to the Hasegawa, Kishino, and Yano (1985) model was estimated using a heuristic search strategy (MULPARS on, one tree held at each step). To avoid restricting the

TABLE 1. Taxa included in phylogenetic analyses of mitochondrial 19S rDNA sequences.

Species	Voucher or publication	Collection source	GenBank accession no.[a]
Angiosperms			
Zea mays L.	Chao et al. (1984)	Mexico (cultigen)	GBANX00794
Nicotiana tabacum L.	D. Nickrent 2917[b]	Tropical America[c]	GBANU82638
Glycine max (L.) Merr.	Grabau (1995)	Asia (cultigen)	GBANM16589
Lindera benzoin (L.) Blume	D. Nickrent 2901	Jackson Co., IL, USA	GBANU82646
Gymnosperms			
Dioon edule Lindl.	D. Nickrent 4125	Mexico[c]	GBANAF058657
Juniperus virginiana L.	D. Nickrent 4069	Jackson Co., IL, USA[c]	GBANAF058658
Pinus strobus L.	D. Nickrent 4066	Jackson Co., IL, USA[c]	GBANAF058659
Pteridophytes			
Adiantum pedatum (Tourn.) L.	J. Duff 9718	Jackson Co., IL, USA	GBANAF058660
Botrychium dissectum var. *obliquum* (Muhl.) Clute	D. Nickrent s.n.	Jackson Co., IL, USA	GBANAF058661
Diplazium pycnocarpon (Spreng.) M. Broun.	J. Duff 9720	Jackson Co., IL, USA	GBANAF058662
Equisetum hyemale L.	D. Nickrent 4056	Jackson Co., IL, USA[c]	GBANAF058663
Ophioglossum vulgatum L. var. *pycnostichum* Fern.	D. Nickrent 4068	Jackson Co., IL, USA	GBANAF058664
Polypodium aureum L.	D. Nickrent 4059	New World tropics[c]	GBANAF058665
Phegopteris hexagonoptera (Michx.) Fee	J. Duff 9719	Jackson Co., IL, USA	GBANAF058666
Lycophytes			
Diphasiastrum digitatum (Dillen.) Holub.	J. Duff s.n.	Knox Co., TN, USA	GBANAF058667
Huperzia lucidula (Mich.) Trevisan	J. Duff 9717	Jackson Co., IL, USA	GBANAF058668
Isoëtes hystrix Bory	D. Nickrent 4072	Cádiz, Spain	GBANAF058669
Mosses			
Atrichum angustatum (Brid.) BSG	J. Duff 9715	Jackson Co., IL, USA	GBANAF058670
Sphagnum palustre L.	J. Duff s.n.	Knox Co., TN, USA[e]	GBANAF058671
Takakia lepidozioides Hattori & Inoue	B. Crandall-Stotler s.n., via W. Schofield	British Columbia, Canada	GBANAF058672
Tortula ruralis (Hedw.) Smith	A. Wood s.n., via M.J. Oliver[f]	Alberta, Canada	GBANAF058673
Liverworts			
Calypogeia muelleriana (Schiffn.) K. Muell.	J. Duff 9703	Jackson Co., IL, USA	GBANAF058674
Marchantia polymorpha L.	Odo et al. (1992)	—	GBANM68929
Hornworts			
Megaceros tosanus Steph.	Z. Iwatsuki 1598	Japan[d]	GBANAF058675
Notothylas orbicularis (Schwien.) Sull.	B. Crandall-Stotler 271	Jackson Co., IL, USA[d]	GBANAF058676
Phaeoceros laevis (L.) Prosk.	J. Thompson 47	Jackson Co., IL, USA	GBANAF058677
Green alga (outgroup)			
Prototheca wickerhamii Tubaki and Soneda	Wolff, Soltis, and Soltis (1994)	—	GBANU02970

[a] The prefix GBAN has been added for linking the on-line version of *AJB* to GenBank and is not part of the actual GenBank accession number.
[b] Voucher specimens of Nickrent and Duff deposited at SIU.
[c] Living plants cultivated at SIUC (under greenhouse conditions or on campus).
[d] Living plants being maintained by B. Crandall-Stotler, SIUC.
[e] Living plants cultivated at University of Tennessee.
[f] Voucher specimen deposited at JEPS.

TABLE 2. Plant mitochondrial 19S rDNA primers.

Primer sequence (5′→3′)	Primer length	Position on *Glycine*	Primer name	Specificity[a]
GAG TTT GAT CCT GGC TCA GA	20	9–28	9 for	mt specific
CCA CAC TRG GAC TG	14	322–335	282 for	mt, cp, and bac
GCC GCT TGT AAA GCT C	16	434–450	434 for	mt specific
GTG CCA GCA GCC GCG G	16	505–520	530 for	mt, cp, and bac
CCA AAA GCG AAG GCA	15	707–721	707 for	mt specific
GCC GTA AAC GAT G	13	797–809	756 for	mt, cp, and bac
AAC TCA AAG GAA TTG	15	893–907	856 for	mt, cp, and bac
CTG CAT GGC TGT CGT C	16	1035–1050	1090 for	mt specific
CAC ACG TGC TAC AAT	15	1622–1636	1622 for	mt specific
GTA CAC ACC GCC CGT CAC ACC	21	1787–1807	1787 for	mt specific
AGT YGC AGT GTG GCT G[b]	16	318–333	318 rev	mt specific
CCT ACG TGC CCT TTA CGC [b]	18	558-575	558 rev	mt specific
CAC ACG AAA TTC CAC T	16	658–673	658 rev	mt specific
CCA GGC GGA GTG TTT	15	857–871	857 rev	mt specific
GCC CCC GYC AAT TCC T	16	900–915	915 rev	mt, cp, and bac
GCT GGT AAG GTT TTG CG	17	957–973	957 rev	mt specific
CCA TGC ACC ACC TG[b]	14	1029–1042	993 rev	mt, cp, and bac
CCA CCT TCC TCC AGT	15	1567–1581	1567 rev	mt specific
CAT GCG GAC TTG ACG TCA	18	1585–1602	1585 rev	mt specific
GCC ACA GGT TCC CCT ACG GCT[b]	21	1949–1969	1949 rev	mt specific

[a] mt = mitochondrial 19S rDNA, cp = chloroplast 16S rDNA, and bac = bacterial 16S rDNA.
[b] Erroneously reported as reverse complements in Duff and Nickrent (1997).

search to a single island of trees (Maddison, 1991), a random stepwise addition sequence was specified for 100 replicates. The Kishino and Hasegawa (1989) test was then used to estimate the significance of the differences in the log likelihoods of two suboptimal trees (A and B, from the MP) to the best tree (C, from ML).

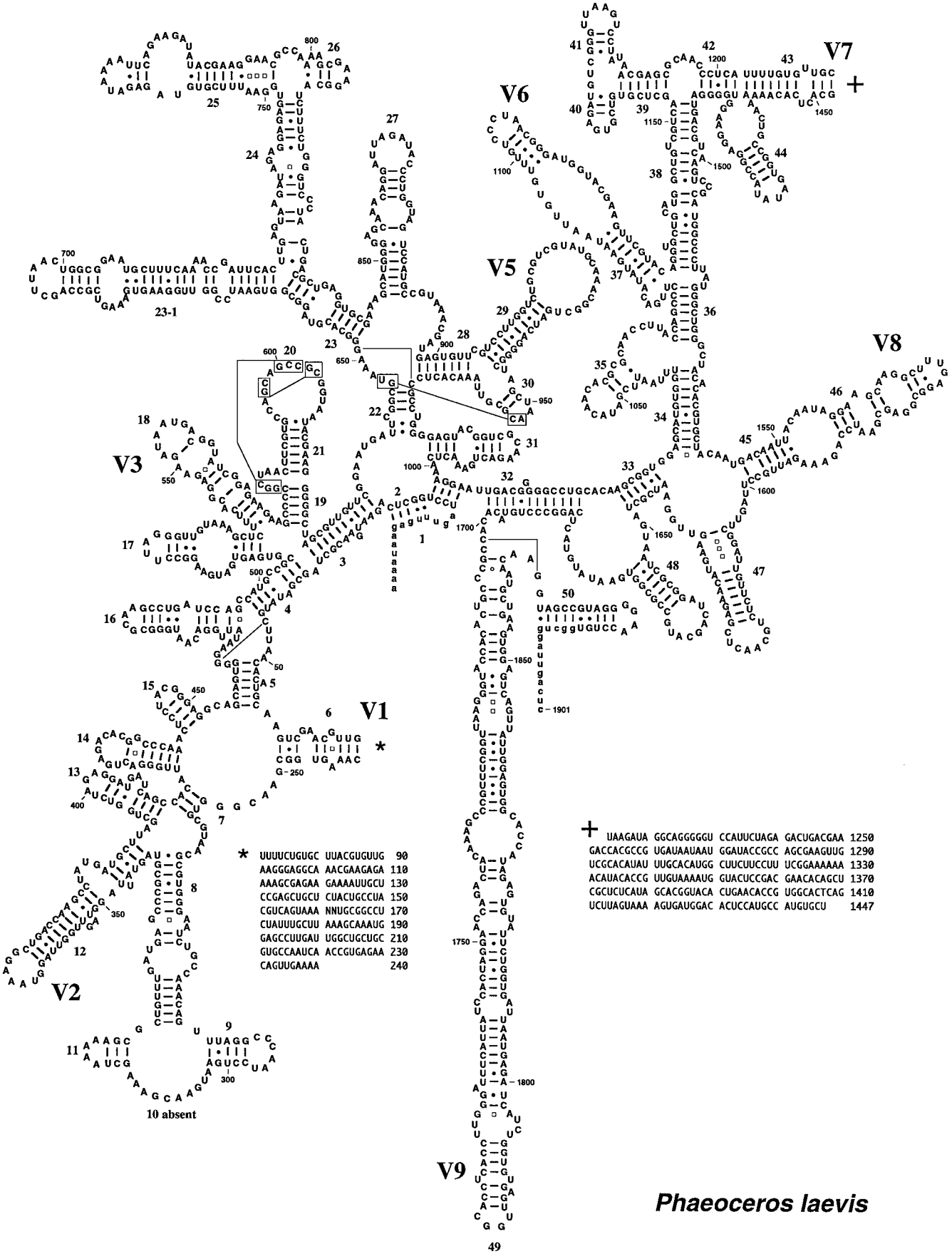
The effects of taxon inclusion/exclusion were tested using the Lanyon (1985) delete-one jackknife procedure. A matrix with one ingroup taxon removed (18 taxa plus outgroup) was then subjected to 100 bootstrap replications (Felsenstein, 1985) using MP heuristic searches. This procedure was repeated for all 19 ingroup taxa. The ranges in bootstrap values (those greater than 50%) were then plotted on the appropriate nodes of the strict consensus branch and bound MP tree. Additional tests were conducted by deleting one or two hornwort taxa and examining the effects on the robustness of clades using both bootstrap and Bremer decay (Bremer, 1988) analyses.

RESULTS

Among the 1716 nt present in the 27-taxon alignment, 449 sites (26%) were variable and among these 268 (15.6 %) were phylogenetically informative. These values were obtained following removal of the outgroup (*Prototheca*) given its genetic distance from the ingroup taxa inflates variability measures. Previous comparisons of transition/transversion ratios in angiosperms showed a marked reduction in the number of transitions or even transversion biases (Duff and Nickrent, 1997). This trend is not apparent in the present data set, which includes representatives of all major embryophyte groups. The average of all pairwise comparisons, including the outgroup *Prototheca,* resulted in a transition/transversion ratio of 1.62. A slight bias toward transversions is seen in those pairwise comparisons involving *Prototheca,* which ranged from 0.91 to 1.07. Comparisons with *Sphagnum* yielded a high bias toward transitions with ratios often greater than 2:1. Nucleotide composition varied no more than 3% over all land plants with an average A+T composition (excluding *Prototheca*) of 48%. Only *Prototheca* differed significantly in nucleotide composition with a strong bias toward A+T (60%).

The distribution of lengths for the 10 000 random trees evaluated was strongly skewed to the left ($g_1 = -0.722$) compared to the critical value of -0.09 ($P < 0.01$) for 25 taxa and 500 characters (Hillis and Huelsenbeck, 1992). This g_1 value indicates that the data are significantly more structured than are random data and implies the presence of strong phylogenetic signal in the 19S rDNA data set.

The heuristic analysis of the 27-taxon matrix resulted in eight most-parsimonious trees of length 1122, one of which is shown as a phylogram in Fig. 2. Topological differences among the eight trees occur only in the angiosperm and leptosporangiate fern clades, both of which collapse to a polytomy in the strict consensus tree. All of the major embryophyte clades are monophyletic, i.e., mosses, liverworts, hornworts, lycophytes, pteridophytes (including *Equisetum*), gymnosperms, and angiosperms. Bootstrap and Bremer support were high for the majority of these clades including mosses (94%), liverworts (100%), pteridophytes (100%), eusporangiate ferns (93%), leptosporangiate ferns (100%), seed plants (100%), gymnosperms (92%), and angiosperms (100%). Moderately high bootstrap support was obtained for the tracheophyte clade (78%), the *Equisetum*–fern–seed plant clade (Euphyllophytina) (84%), hornworts (75%), and lycophytes (Lycophytina) (66%). The latter (low) bootstrap value derives from sequence divergence between *Isoëtes* and the two members of Lycopodiales. The sister relationship between *Equisetum* and the leptosporangiate ferns results in a paraphyletic Filicopsida. *Takakia* emerges as the most basal member of the mosses and *Sphagnum* occurs on a relatively long branch that is sister to the clade composed of *Tortula* and *Atrichum.* Despite the topology shown in Fig. 2, the relationships between the three major bryophyte clades (mosses, liverworts, and hornworts) and the remaining embryophytes are not well resolved, i.e., with bootstrap values <50%. Trees only one or two steps longer result in a polytomy involving the moss, liverwort, hornwort, and tracheophyte clades.


V1
V2
V3
V5
V6
V7
V8
V9
* UUUUCUGUGC UUACGUGUUG 90
AAGGGAGGCA AACGAAGAGA 110
AAAGCGAGAA GAAAAUUGCU 130
CCGAGCUGCU CUACUGCCUA 150
CGUCAGUAAA NNUGCGGCCU 170
CUAUUUGCUU AAAGCAAAUG 190
GAGCCUUGAU UGGCUGCUGC 210
GUGCCAAUCA ACCGUGAGAA 230
CAGUUGAAAA 240
+ UAAGAUA GGCAGGGGGU CCAUUCUAGA GACUGACGAA 1250
GACCACGCCG UGAUAAUAAU GGAUACCGCC AGCGAAGUUG 1290
UCGCACAUAU UUGCACAUGG CUUCUUCCUU UCGGAAAAAA 1330
ACAUACACCG UUGUAAAAUG GUACUCCGAC GAACACAGCU 1370
CGCUCUCAUA GCACGGUACA CUGAACACCG UGGCACTCAG 1410
UCUUAGUAAA AGUGAUGGAC ACUCCAUGCC AUGUGCU 1447
10 absent
Phaeoceros laevis

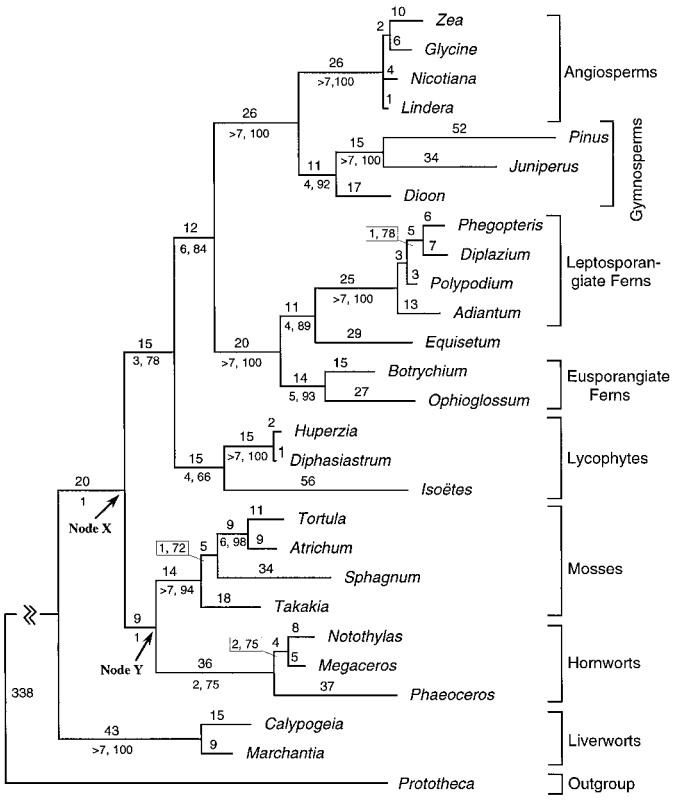


Fig. 2. One of eight shortest, equally parsimonious trees obtained using the maximum parsimony optimality criterion. The length is 1122 steps for 26 ingroup embryophyte taxa and the green algal outgroup taxon *Prototheca.* Branch lengths are indicated above the lines. The number of steps required to collapse a branch (Bremer decay index) followed by bootstrap percentages (based upon 100 replicates) is shown below the lines. Lines with no bootstrap values below denote nodes supported in <50% of the replications. Nodes X and Y collapse in trees one step longer, thus resulting in a polytomy for liverworts, hornworts, mosses, and tracheophyes. This tree was constructed using 299 phylogenetically informative characters out of the total 620 variable characters. CI (excluding uninformative characters) = 0.5892, RI = 0.7393, RC = 0.5344.

For the 20-taxon alignment (again, excluding the outgroup), 435 sites (29.9%) were variable and among these 236 (16.2 %) were phylogenetically informative. The branch-and-bound analysis of the 20-taxon matrix resulted in two most-parsimonious trees (A and B) of length 1074. These trees were present on two distinct islands as determined by conducting heuristic searches with random taxon addition sequence. Tree A was fully congruent with the one obtained from the heuristic search using the 27-taxon matrix (Fig. 2) and showed liverworts as the most basal embryophyte clade. Tree B, however, placed hornworts as the most basal lineage and mosses sister to liverworts (Fig. 3). The strict consensus of these two trees (not shown) again results in a polytomy involving the moss, liverwort, hornwort, and tracheophyte clades. Bootstrap support (using heuristic search strategies) was generally equal to or higher than values obtained from the 27-taxon analysis. For example, the tracheophyte clade was now supported in 91% of the bootstrap replicates. Similar increases occurred for the Euphyllophytina clade (92%) and the lycophytes (80%). Decreased bootstrap support was seen for the hornwort clade (66%).

All 100 replicates of the heuristic search using the ML optimality criterion resulted in the same tree (−ln likelihood = 7195.07335) using a random addition sequence for the 20-taxon data matrix. This tree had a topology similar to tree B obtained using MP, however the ML tree (inset, Fig. 3) differed in that the three hornwort genera were paraphyletic. Whether significant differences existed between the –ln likelihood values of trees A and B (MP) and C (ML) were determined by the Kishino and Hasegawa (1989) test. The −ln likelihood of tree A was 7201.07382 and 7198.59727 for tree B. Results of this test showed the likelihood values obtained from the MP trees were not significantly different than the optimal tree (P = 0.5574 and 0.5088, respectively, for trees A and B).

Results of the delete-one jackknife experiments are shown as ranges in bootstrap values on Fig. 3. Generally these values deviated little from the bootstrap values obtained when the 20-taxon data matrix was analyzed, but exceptions occurred. For example, the range in bootstrap values for the Euphyllophytina clade is 58–98% with the lowest value attributable to the removal of *Diphasiastrum.* Indeed, deletion of this taxon resulted in four of the lowest scores recorded across all nodes. Similarly, *Tortula* and *Juniperus* together accounted for six other lowest scores. It might be expected that clades would be most affected by removal of taxa that are components of that clade. More frequently, the deletion of a taxon influences bootstrap values for distant clades—a somewhat nonintuitive result.

The effects on tree topology, bootstrap, and decay index values were determined after excluding particular combinations (delete-one and delete-two variations) of hornwort and liverwort taxa from the 20-taxon matrix (Fig. 4A-H). With the exception of the taxon combination in Fig. 4C, all the major embryophyte lineages in Fig. 4 can be reduced to a polytomy when clades with low bootstrap support are collapsed. Deletion of one hornwort retained the topology with liverworts basal when *Phaeoceros* or *Notothylas* were excluded. Removal of *Megaceros* from the matrix, however, resulted in the hornwort clade being basal (Fig. 4D). The most dramatic effect upon tree topology and bootstrap support values is achieved by removing both *Notothylas* and *Megaceros,* leaving *Phaeoceros* as the sole anthocerote representative (Fig. 4E). This combination of taxa results in strong bootstrap support for a basal hornwort (*Phaeoceros*) and well-

←

Fig. 1. Higher order structural model for the ''core'' mitochondrial-encoded small-subunit ribosomal RNA for *Phaeoceros laevis* (Anthocerophyta). Base pairing follows the covariance model proposed for *Zea* by Gutell (1994). Lowercase bases at the 5′ and 3′ end of the molecule were not determined because they are near priming sites (*Zea* sequence shown for clarity). Secondary structures are not known for the V1 and V7 sequences, hence only the sequence is shown.

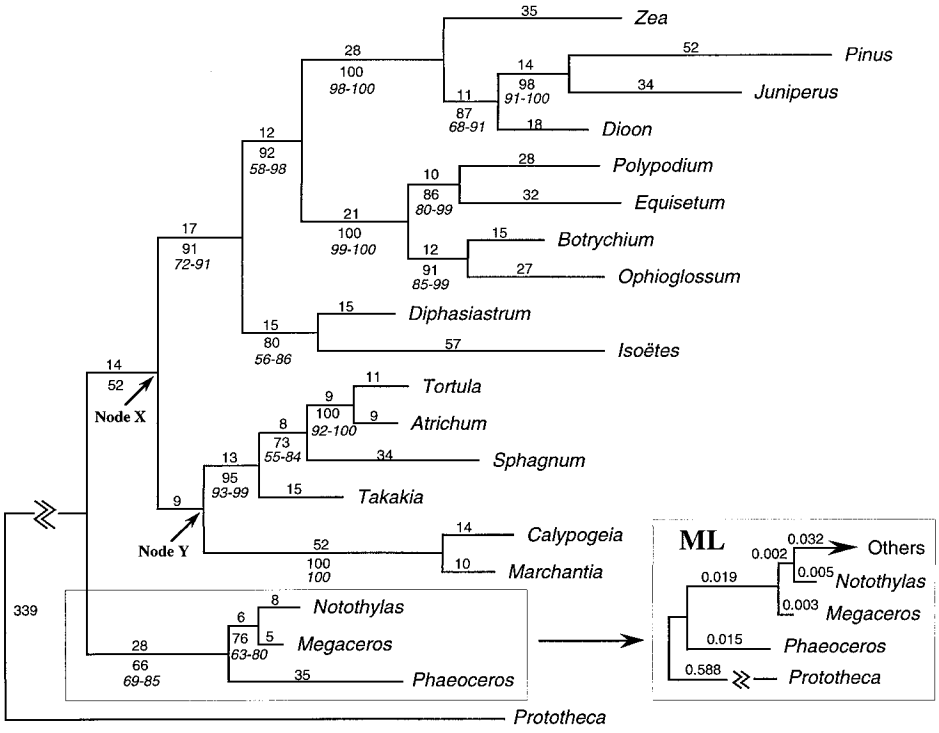


Fig. 3. Tree B of two shortest, equally parsimonious trees obtained from a branch-and-bound search using maximum parsimony. The length is 1072 steps for 19 ingroup embryophye taxa plus the outgroup *Prototheca.* Branch lengths are indicated above the lines. Bootstrap percentages (based upon 100 replicates) are shown immediately below the lines. The italicized values represent ranges of bootstrap percentages obtained via a delete-one jackknife operation (see text). Lines with no bootstrap values below denote nodes supported in <50% of the replications. Nodes X and Y collapse in trees one step longer, thus resulting in a polytomy for liverworts, hornworts, mosses and tracheophyes. This tree was constructed using 273 phylogenetically informative characters out of the total 606 variable characters. CI (excluding uninformative characters) = 0.5863, RI = 0.6458, RC = 0.4753. The boxed inset shows the portion of the maximum likelihood tree (tree C) that differed in topology from the parsimony tree B. Numbers above the lines indicate branch lengths. The score of the best ML tree was 7195.07335.

supported clades for the remaining embryophyte lineages.

DISCUSSION

Interest in the early evolution of land plants has increased tremendously over the past decade, owing to the influx of new comparative data from morphology, ultrastructure, and paleobotany, as well as from increasingly powerful analytical methods (Kenrick and Crane, 1997). Acquisition of gene sequence data from all three subcellular compartments has allowed traditional hypotheses to be tested using molecular phylogenetic methods. The data presented here represent the first published molecular phylogenetic study using mitochondrial-encoded small-subunit rDNA sequences in plants. After removal of length-variable regions, it appears that substitutions within the core 19S rDNA are occurring at an appropriate rate to examine divergence events among extant land plants.

Background on land plant relationships—Modern cladistic analyses of both morphological and molecular data have resulted in a surprisingly diverse array of conflicting topologies in regards to relationships among the major embryophyte lineages (Fig. 5). Early morphologically based cladistic studies of green plant evolution include Parenti (1980) and Bremer and Wanntorp (1981), which were followed by those of Mishler and Churchill (1984, 1985). These analyses led to the hypothesis that the bryophytes are paraphyletic (Fig. 5A). Despite controversy (Robinson, 1985; Whittemore, 1987) this idea gained additional support (Bremer et al., 1987; Kenrick, and Crane, 1991; Mishler et al., 1994) and remained relatively unchallenged for nearly a decade. Based on spermatogenesis (Garbary, Renzaglia, and Duckett, 1993; Maden et al., 1997) an alternate phylogenetic hypothesis was obtained that supported bryophyte monophyly and postulated a primary dichotomy between bryophytes and vascular plants at the base of the embryophytes (Fig. 5B). A more extensive character matrix involving general morphological, morphogenetic, and ultrastructural data was recently analyzed by Garbary and Renzaglia (1998). The resulting tree (Fig. 5C) suggests that the hornworts are basal to a clade containing a monophyletic moss and liverwort assemblage plus tracheophytes, however this topology appears dependent upon which charophyte outgroup taxon is used. The strict consensus of 54 trees resulting from the parsimony analysis of morphological data conducted by Kenrick and Crane (1997) resolved only a sister relationship between mosses and tracheophytes (Fig. 5D).

Molecular phylogenetic methods using DNA sequences derived from all three subcellular compartments have been used to address embryophyte relationships. One of

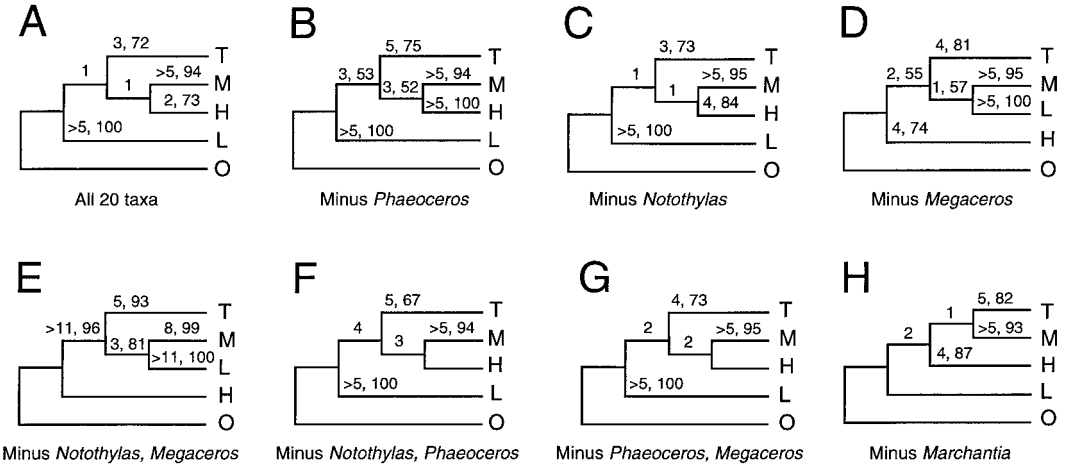


Fig. 4. Results of delete-two jackknife involving the hornwort and liverwort taxa. Shown above the lines are the decay indices followed by bootstrap percentages (based upon 100 replicates). Lines with no bootstrap values denote nodes supported in <50% of the replications. No decay or bootstrap values are shown when only one hornwort (E, F, and G) or one liverwort (H) are included.

the earliest studies used partial 18S and 26S rRNA sequences (Waters et al., 1992). The single most-parsimonious cladogram contained hornworts and mosses on a clade that was sister to the tracheophytes with liverworts basal, however only one additional step was required to constrain the topology to that of Mishler and Churchill (1984; Fig. 5A). The study by Mishler et al. (1994) examined morphological and molecular data sets (partial 18S and 26S rRNA) separately and in combination. The combined molecular data resulted in a strict consensus tree that provided no resolution among the bryophytes or tracheophytes, and separate analyses of these partitions gave conflicting topologies. The combination of 26S rDNA sequence and morphological characters resulted in the same topology as derived from morphology alone (Fig. 5A), likely because of insufficient phylogenetic signal present in the partial 26S rRNA sequences. In contrast, the combination of 18S rDNA sequences and morphological characters resulted in a basal position for hornworts (Fig. 5E). A more recent study used complete 18S rDNA sequences for 24 land plants: six tracheophytes, nine mosses, seven liverworts, and two hornworts (Hedderson, Chapman, and Rootes, 1996). One most-parsimonious tree was retrieved following a branch-and-bound search that placed mosses and liverworts on a clade sister to the tracheophytes, and hornworts sister to that entire assemblage (Fig. 5F). Constraining this solution to the topology of Mishler and Churchill (1984) added eight steps, i.e., the highest among all constraint options tried. Nuclear 18S rDNA has also been used in phylogenetic studies of bryophytes (Bopp and Capesius, 1996) and liverworts (Capesius and Bopp, 1997). The former study included only one hornwort representative and no tracheophytes, and the latter was focused on liverworts and used mosses as the outgroup. In both cases, mosses (Bryopsida) were sister to leafy liverworts (Jungermanniopsida) and that clade sister to thalloid liverworts (Marchantiopsida), thus making liverworts paraphyletic. The sister relationship of the hornwort *Anthoceros* with Jungermanniopsida (Bopp and Capesius, 1996) is incongruent with other molecular phylogenetic analyses (Fig. 5E-H).

In addition to nuclear rRNA genes, chloroplast-encoded genes have also been sequenced from representatives of all embryophyte groups. Small-subunit (16S) rDNA genes appear to contain insufficient numbers of variable

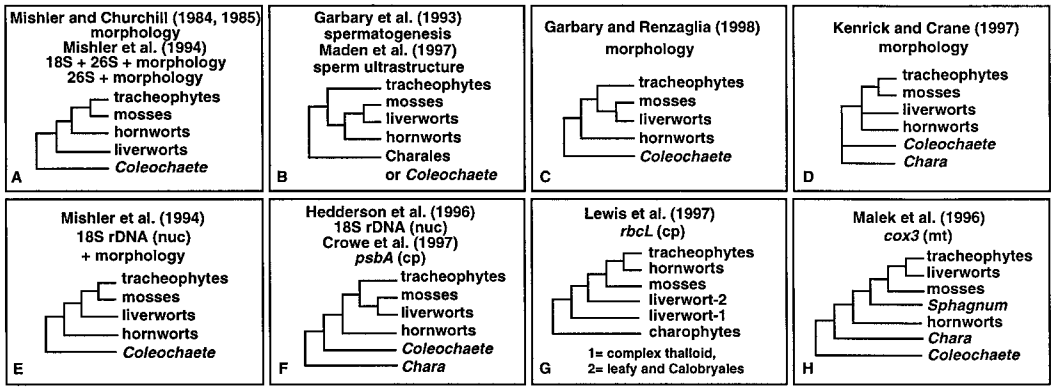


Fig. 5. Results of various cladistic analyses of land plants (embryophytes) using morphological (B-D), molecular (F-H), or combinations of both (A and E). Tracheophytes include angiosperms, gymnosperms, pteridophytes, and lycophytes. Only those studies that included a representative of tracheophytes, mosses, hornworts, and liverworts are shown.

TABLE 3. Comparison of plant small-subunit rDNA genes.[a]

Features	Plastid 16S rDNA	Nuclear 18S rDNA	Mitochondrial 19SrDNA
Number of taxa in matrix	13	20	20
Placeholder (missing) taxa	1 (7)[b]	4 (0)[c]	0 (0)
Total characters	1432	1702	1453
Constant characters	1141 (79.7%)	1161 (68.1%)	1018 (70%)
Uninformative characters	180 (12.5%)	334 (19.7%)	199 (13.7%)
Informative characters	111 (7.7%)	207 (12.1%)	236 (16.2%)
Total variable characters	281 (19.6%)	541 (31.8%)	435 (29.9%)
Informative/total characters	39.5%	38.2%	54.2%
CI (minus uninformative sites)	0.5691	0.4893 (0.5553)[d]	0.5863 (0.6167)
HI (minus uninformative sites)	0.4309	0.5107 (0.4447)	0.4137 (0.3833)
Retention Index	0.5032	0.4811 (0.4806)	0.6458 (0.5503)
Rescaled Consistency Index	0.3898	0.3109 (0.3474)	0.4753 (0.4412)

[a] Calculations exclude the outgroups (*Coleochaete* for 18S and 16S, *Prototheca* for 19S) because their distances to ingroup taxa inflate variability measures.

[b] *Doodia* was used as a placeholder for *Polypodium*. The *Takakia* 16S rDNA sequence was generated for this study (GenBank accession number GBANAF058678). Missing taxa with no available placeholders in GenBank include: *Dioon*, *Tortula*, *Atrichum*, *Calypogeia*, *Notothylas*, *Megaceros*, and *Phaeoceros*.

[c] Placeholder (19S taxon): *Cycas (Dioon)*, *Adiantum (Polypodium)*, *Andreaea (Tortula)*, *Anthoceros (Megaceros)*.

[d] Number in parentheses determined from a 13-taxon matrix to allow better comparisons to 16S rDNA.

sites to resolve land plant phylogenetic relationships (Manhart, 1995). Sequences of *rbcL* have proven useful in examining relationships within particular embryophyte clades such as angiosperms (reviewed in Chase and Albert, 1998), gymnosperms (Price, 1996; Hasebe et al., 1992; Goremykin et al., 1996), pteridophytes (Wolf, Soltis, and Soltis, 1994; Hasebe et al., 1995; Pryer, Smith, and Skog, 1995), and liverworts (Lewis, Mishler, and Vilgalys, 1997). Fewer studies have used this gene to examine deep relationships among all embryophytes, an exception being the study by Manhart (1994) that generated *rbcL* sequences for 23 embryophytes (one hornwort, three mosses, two liverworts, and 17 tracheophytes). Several unusual relationships resulted from this analysis, and low levels of support were seen for many clades. Although nucleotide substitutions rates for some of the tracheophyte clades may be saturated (Goremykin, et al., 1996), Lewis, Mishler, and Vilgalys (1997) concluded that third positions in *rbcL* retain phylogenetic signal across green plants. That study, focused mainly upon liverwort relationships, concluded that bryophytes and possibly liverworts were paraphyletic, hornworts were sister to tracheophytes, and liverworts were the basal embryophyte lineage (Fig. 5G). Poor resolution among certain branches was attributed to rapid radiation of those clades. Another chloroplast gene that appears to be evolving at the appropriate rate to examine deep relationships among embryophytes is *psbA* (Crowe et al., 1997). This gene is highly conserved (>80%) and preliminary phylogenetic studies have resulted in topologies similar to those obtained using 18S rDNA that place hornworts as sister to the remaining land plants (Hedderson, Chapman, and Rootes, 1996; Fig. 5F).

Among the ~90 genes present in the mitochondrial genome, only *cox3* has been used broadly in comparative phylogenetic analyses. This gene has been used to demonstrate the effects of RNA editing on phylogenetic reconstruction (Hiesel, von Haeseler, and Brennicke, 1994; Bowe and dePamphilis, 1996; Malek et al., 1996). Analyses of partial *cox3* sequences using parsimony and maximum likelihood methods for 23 embryophytes placed the hornwort *Anthoceros* at the base of the land plant clade with high bootstrap support (Malek et al., 1996; Fig. 5H). The small size of this gene (~700 nt), however, limits the number of phylogenetically informative sites available to resolve relationships among all groups of land plants. Recent phylogenetic analyses of the mitochondrial *nad5* gene (V. Knoop, unpublished data) resulted in the same topology derived from 18S rDNA + morphology (Fig. 5E).

Phylogenetic signal in 19S rDNA—For the molecular phylogenetic method to reflect accurately evolutionary history, it is important to match features of sequence divergence (e.g. substitution rate, proportion of sites free to vary, etc.) with the time frame of divergence of the organisms under study (Graybeal, 1994; Hillis, Mable, and Moritz, 1996). With respect to embryophyte phylogeny, Wolf (1997) stated ''finding a single gene that has a strong phylogenetic signal for ancient divergence events is unlikely.'' Although it is unlikely that a single molecule exists that is effective in resolving *both* ancient and recent divergences, rDNA genes retain phylogenetic signal over very long time periods. For example, the conserved 16S rDNA genes have been widely used to infer cladogenic events among eukaryotes and prokaryotes that occurred greater than 500 million years ago (mya) (Olsen, 1987; Woese, 1987). As evidenced by microfossils, embryophytes emerged in the mid-Ordovician, that is ~450 mya (Kenrick and Crane, 1997), thus genes such as the plastid-encoded 16S rDNA *should be* appropriate for probing relationships at this level. Such was not the case as was shown in the study of Manhart (1995), which demonstrated that too few substitutions were present for adequate resolution of the major lineages.

It is of interest to compare sequence variability characteristics and phylogenetic utility of the small-subunit ribosomal RNA genes derived from the plastid (16S), nucleus (18S), and mitochondrion (19S). Excluding the outgroups, the number of phylogenetically informative characters for 16S, 18S, and core 19S are 111, 207, and 236, respectively (Table 3). Expressed as a percentage of the

number of variable characters, the proportion of phylogenetically useful characters in mitochondrial 19S rDNA is greater than either 18S or 16S. The presence of strong phylogenetic signal was also apparent from the results of skewness tests (see Results). The amount of homoplasy (determined from data sets of comparable size) is lowest for 19S, intermediate for 16S, and highest in 18S rDNA. These three rDNAs compare well to the theoretical patterns of gene evolution proposed by Graybeal (1994), whereby percentage sequence difference is plotted against time since divergence. The first pattern, which is comparable to 16S rDNA, provides phylogenetic information for ancient divergences (500 mya) but has low numbers of phylogenetically informative sites that change relatively slowly. The second pattern provides phylogenetic information up to ~100 mya but is homoplasious for deeper divergences. Here, a moderate proportion of sites are free to change, but the accumulation of multiple hits (saturation) eventually occurs, resulting in a flattening of the curve. That this pattern explains the behavior of nuclear 18S rDNA is shown by its comparably higher level of homoplasy and its inability to resolve deeper embryophyte relationships (Kranz et al., 1995; P. Soltis, unpublished data). The third pattern shows a moderate but continuous accumulation of sequence divergence over time and thus yields proportionately greater numbers of phylogenetically informative characters per length of the molecule. Such a gene shows lower homoplasy and allows resolution of deeper divergences. This pattern may explain the behavior of mitochondrial 19S rDNA, at least during early and intermediate divergence times prior to eventual saturation. The problem of resolving intermediate-level divergences (i.e., in the 50–300 mya range) is not unique to plants (Graybeal, 1994). For this reason, the discovery that mitochondrial 19S rDNA sequences have high bootstrap support for intermediate nodes in the embryophyte tree is important in that few molecules are available that are useful at this level. Given that phylogenetic signal is lost when attempting to use 19S rDNA sequences among closely related plants (e.g., angiosperms, ferns, Lycopodiales), additional resolution at this level might be attained by combining 19S rDNA with a more rapidly evolving gene.

Phylogenetic information content of 19S rRNA structural features—In addition to substitutional mutations that contribute phylogenetically informative characters, we have discovered a number of structural features of 19S rRNA that clearly are associated with particular clades. Complete 19S rRNA higher order structural diagrams have been prepared for nine land plant representatives. These rRNAs will be more fully characterized in a separate publication. The following represents a summary of those features that support the molecular phylogenetic relationships reported here.

Novel structural features add additional support to the already well-supported (100% bootstrap) pteridophyte plus *Equisetum* clade. All pteridophytes (but not *Equisetum*) have a V1 region 50 nt in length or shorter. Whereas most land plants lack helix 10 (see *Phaeoceros*, Fig. 1), all ferns and *Equisetum* have additional nucleotides that can be readily paired to form a stem-loop structure. Additional pteridophyte-specific features include a longer than average helix 17 and a longer V6 region. The two liverworts have V5 regions much larger than other land plants (117 nt), which are more similar in length to *Phaeoceros* (37 nt, Fig. 1). The V5 region is identical in sequence and in length for *Marchantia* and *Calypogeia* and is also highly similar to that found in two other genera for which only partial sequences are currently available (*Conocephalum* and *Riccia*). It has been reported previously that the *Marchantia* 19S rDNA contains a 2.4-kb intron positioned at the base of helix 31 on the mature rRNA (Odo et al., 1992). The complete sequence of *Calypogeia* (leafy liverwort) and PCR amplifications flanking this region in *Pellia* (simple thalloid) and *Riccia* (complex thalloid) suggest they lack this large intron. Further sampling is required to see whether this intron occurs only in members of Marchantiales. A putative intron of 1.1 kb in length was discovered in three of the four mosses surveyed (absent in *Tortula*). This intron occurs in the apical loop portion of helix 27, an extremely conserved portion of the core 19S rRNA. Its absence in *Tortula* may represent a recent loss. Curiously, an insertion of 226 nt occurs at the same position in *Isoëtes* and its sequence is 90% similar to the moss intron. Insertions of identical length were also seen in six other *Isoëtes* species but were absent in *Diphasiastrum* and *Huperzia.* Finally, a putative intron (greater than 1 kb) was identified in the helix-10 region of *Psilotum.* This intron partly accounts for our inability to obtain a complete sequence for this taxon. Similarly, the lack of complete sequences from *Gnetum* and *Ephedra* can also be attributed to large, putative introns in their 19S rDNA sequences. Although these structural features introduce methodological complications, once they are fully characterized they offer potentially valuable sources of phylogenetic data.

Relationships among land plants deduced from 19S rDNA—The most recent review of relationships among land plants considered morphological features of both extant and extinct plants as well as ultrastructural and molecular phylogenetic data (Kenrick and Crane, 1997). Their resulting morphologically based classification (cf. Chapter 7) of Chlorobiota (green plants) will be used as a framework for discussion of relationships derived from the present study (Fig. 6). As discussed in the Introduction and as shown in Fig. 5, considerable disagreement exists as to the branching order for the major embryophyte clades. Despite this, a number of common features have emerged. For example, there is strong support for the concept that the grade taxon Charophyceae contains the closest algal relatives of embryophytes (Bremer, 1985; Graham, Delwiche, and Mishler, 1991; Graham, 1993). In the present study, we were not able to test this directly given our inability to obtain a charophycean 19S rDNA sequence. Future work would benefit from the inclusion of such sequences, which could be used as outgroups in studies of embryophyte phylogeny, as well as to further address which clade is sister to the land plants.

Although traditional classifications divide embryophytes into two groups (bryophytes and tracheophytes), the majority of morphological and molecular phylogenetic studies show the former (mosses, liverworts, and hornworts) to be paraphyletic. Furthermore, most analyses result in monophyletic embryophytes (compared with

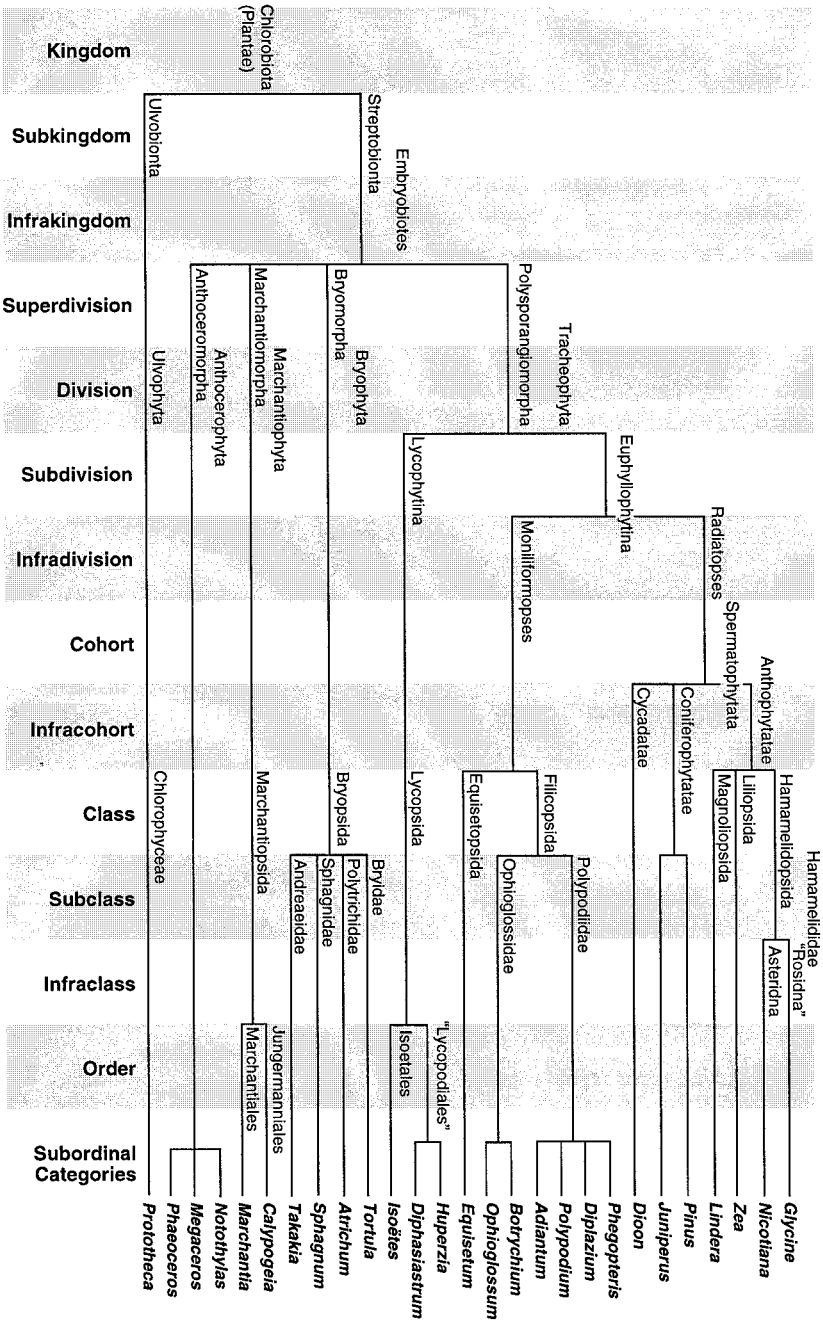


Fig. 6. Morphologically based classification of Kenrick and Crane (1997) superimposed upon the suite of taxa used in this study. In general, the topology is supported by results from mitochondrial 19S rDNA (cf. Figs. 2 and 3). Taxa more closely related than the lowest category given in that classification are joined under subordinal categories. Taxa in quotes are paraphyletic. See Kenrick and Crane (1997) for further details.

algal outgroups), mosses, liverworts, hornworts, and tracheophytes. Exceptions to this include the study by Manhart (1994), which, in an equal weighted parsimony analysis of *rbcL*, produced polyphyletic mosses, liverworts, and tracheophytes. Equal weighted parsimony resulted in the topology shown in Fig. 5G in the *rbcL* study by Lewis, Mishler, and Vilgalys (1997) that depicts liverworts as paraphyletic. Differential weighting and maximum likelihood analyses resulted in similar relationships among more terminal clades but significantly different topologies for deeper nodes.

Phylogenetic analyses of 19S rDNA using both parsimony and maximum likelihood resulted in relationships that are highly congruent with relationships depicted in the classification of Kenrick and Crane (1997). These molecular analyses resulted in high bootstrap support for many of the higher level taxa shown in Fig. 6. Monophyletic Superdivisions (Divisions) include Anthoceromorpha (Anthocerophyta), Marchantiomorpha (Marchantiophyta), Bryomorpha (Bryophyta), and Polysporangiomorpha (Tracheophyta). Molecular phylogenetic analyses also fully support the recognition of two well-established clades within Tracheophy-

ta, i.e., Euphyllophytina and Lycophytina. For the latter, prior cpDNA studies indicated a relationship with the bryophytes (Raubeson and Jensen, 1992). In agreement with our 19S rDNA results, phylogenetic analyses employing *cox3* clearly support the placement of Lycophytina at the base of the tracheophytes (Malek et al., 1996).

Within Lycophytina, extant plants occur in class Lycopsida and three orders Lycopodiales (paraphyletic), Selaginellales, and Isoëtales. With 19S rDNA, *Isoëtes* is sister to *Diphasiastrum* (and *Huperzia*) with moderately high bootstrap support. This taxon is present on a long branch, suggesting either substitution rate heterogeneity or insufficient taxon density (note no sequence of *Selaginella* could be obtained). The sequence of *I. hystrix* was used in this study because it was full length, however it differed significantly from partial sequences obtained from seven other diverse representatives of the genus. For example, comparison of a partial sequence of *I. caroliniana* (~1200 nt) with *I. hystrix* showed that the latter differed at over 30 sites but had V1 and V7 regions of nearly the same length and sequence. We have yet to fully explore these observations, however possible explanations are RNA editing (unlikely given the types of changes) and multiple, different copies of 19S rDNA within individuals.

As mentioned in Kenrick and Crane (1997), the previously unnamed taxon Euphyllophytina has been widely recognized and is well supported by the presence of at least seven synapomorphies (mainly involving vegetative anatomy). Strong support for this clade (92% bootstrap) was obtained using 19S rDNA sequences. This taxon is further divided into two Infradivisions, Radiatopses and Moniliformopses, on the basis of position and ontogeny of the protoxylem. Extant taxa in Moniliformopses occur in two classes, Filicopsida (ferns) and Equisetopsida (also called Arthrophyta—the horsetails). The exact relationship between *Equisetum* and the ferns remains controversial owing to variation in placement following different analyses. For example, *Equisetum* occurs as basal to all tracheophytes using *atpB* (Wolf, 1997) or as sister to vascular plants minus lycophytes using mitochondrial *cox3* (Malek et al., 1996). Pryer, Smith, and Skog (1995), using a 50-taxon character matrix that combined morphological and *rbcL* characters, found *Equisetum* to be more closely related to the leptosporangiate ferns than to either *Angiopteris* or the eusporangiate fern *Botrychium*. This result is supported by morphological and spermatogenesis data (Maden et al., 1997; Garbary and Renzaglia, 1998) and received moderately high bootstrap support (86%) following analysis of 19S rDNA sequences. Given these results, it may eventually be appropriate to classify *Equisetum* at a level equivalent with other subclasses within Filicopsida. As mentioned above, *Psilotum* was not included in our analysis because only a partial 19S sequence (~1000 nt) was available. Analyses using this partial sequence supports a sister relationship of this genus with the eusporangiate ferns (specifically, Ophioglossaceae), a result in agreement with Manhart (1994), Pryer, Smith, and Skog (1995), Malek et al. (1996), and Wolf (1997).

All extant members of Infradivision Radiatopses are classified within cohort Spermatophytata (seed plants) by Kenrick and Crane (1997). Four equivalent Infracohorts were proposed: Cycadatae (cycads), Coniferophytatae (conifers), Ginkgoatae (Ginkgo), and Anthophytatae (flowering plants). Although we lack a 19S rDNA sequence for *Ginkgo*, our results strongly suggest further hierarchical structure whereby conifers and cycads are sister. This is in agreement with results derived from 18S rDNA and *rbcL*, which show gymnosperms to be monophyletic (Goremykin et al., 1996; Chaw et al., 1997). Additional sequencing within representative gymnosperms will undoubtedly yield further resolution of relationships within Radiatopses. Partial 19S rDNA sequences of *Ephedra* and *Gnetum* (data not shown) exhibited long branches, thus suggesting substitution rate heterogeneity for this gene. Rate heterogeneity (or long branch lengths) have also been noted among gnetophytes in other molecular phylogenetic studies using nuclear, plastid, and mitochondrial genes (Manhart, 1995; Goremykin et al., 1996; Price, 1996; Malek et al., 1996). As shown in Fig. 2, too few substitutions exist to resolve relationships within angiosperms (here Infracohort Anthophytatae), although as shown by Duff and Nickrent (1997), increased rates of evolution for 19S rDNA in parasitic flowering plants mirror similar accelerations documented for their other genomes (Nickrent et al., 1998).

Liverworts or hornworts as the basalmost embryophyte lineage?—Despite overall high bootstrap support for most clades using 19S rDNA, the question as to which bryophyte lineage (liverworts or hornworts) is basal remains unresolved. In contrast with the majority of molecular analyses that support the hornworts basal hypothesis, our 19S parsimony and maximum likelihood analyses indicate that either topology is an equally viable hypothesis. Although mosses never occupied the basal position, collapsing the two poorly supported internal nodes (X and Y, Figs. 2 and 3) results in a polytomy involving tracheophytes, mosses, liverworts, and hornworts. In no cases were mosses sister to tracheophytes, a relationship reported in a number of morphological (Mishler and Churchill, 1984; Mishler et al., 1994; Kenrick and Crane, 1997) and molecular (Mishler et al., 1994; Lewis, Mishler, and Vilgalys, 1997) analyses. The 19S rDNA topologies obtained in this study are most similar to those derived from analysis of nuclear 18S rDNA (Hedderson, Chapman, and Rootes, 1996; Hedderson, Chapman, and Cox, 1998), although taxon sampling differed substantially between the two studies.

The liverworts basal hypothesis was first championed by Mishler and Churchill (1984, 1985) and later by Mishler et al. (1994). In the cladistic study of Kenrick and Crane (1997; their Fig. 3.35), mosses were resolved as the sister group to vascular plants, but results were conflicting on relationships among liverworts, hornworts, and the tracheophyte/moss clade. For this reason, their summary classification (Table 7.1) shows a basal polytomy involving liverworts and hornworts. In the first study of *rbcL* (Manhart, 1994), *Marchantia* was sister to all remaining embryophytes, however liverworts, mosses, and tracheophytes were polyphyletic and only one hornwort (*Megaceros*) was included. A later study using the same gene (Lewis, Mishler, and Vilgalys, 1997) resulted in liverworts being paraphyletic. Support for the hornworts basal hypothesis comes from morphological studies

(Garbary and Renzaglia, 1998), nuclear 18S rDNA (Mishler et al., 1994; Hedderson, Chapman, and Rootes, 1996), plastid-derived *psbA* (Crowe et al., 1997), and mitochondrial *cox3* (Malek et al., 1996). In a recent study, Qiu et al. (1998) tested for the presence of mitochondrial group II introns in three genes among all land plant lineages. The complete absence of all three introns in liverworts and the green algal outgroups was used to support the concept that liverworts are the earliest embryophytes.

As documented using delete-one and delete-two jackknife approaches, taxon inclusion is of critical importance in obtaining stable topologies for the land plant tree. The particular combination of hornwort taxa appears to strongly affect not only the position of that clade but also the level of bootstrap support for others. Previous molecular phylogenetic analyses that used a single hornwort exemplar [usually *Phaeoceros laevis* (L.) Prosk. = *Anthoceros laevis* L.] should be re-analyzed after adding additional anthocerote sequences. Branches on the 19S rDNA tree (Figs. 2 and 3) leading to *Phaeoceros* are four to seven times longer than those leading to *Notothylas* and *Megaceros,* hence rate heterogeneity must be considered along with incomplete taxon sampling. Sequences in addition to the two liverwort taxa included here are certainly required to increase taxon density. Examination of the three additional partial liverwort 19S rDNA sequences (data not shown) indicated they were very similar to *Marchantia* and *Calypogeia,* hence the rate heterogeneity observed in hornworts may not be an issue in liverworts. This result is of interest given that significant rate increases in *rbcL* were documented for leafy/simple thalloid liverworts as compared with complex thalloids (Lewis, Mishler, and Vilgalys, 1997). Although not specifically discussed, it appears from the distance matrix for *rbcL* that rate differences also exist between the three hornworts sampled (*Anthoceros punctatus, Megaceros vincentianus,* and *M. aenigmaticus*) in that study (Lewis, Mishler, and Vilgalys, 1997).

Conclusions—This study represents the first molecular phylogenetic study to employ mitochondrial-encoded 19S rDNA. This molecule contains a greater proportion of phylogenetically informative sites and lower amounts of homoplasy than either nuclear 18S or plastid 16S rDNA, hence it is well suited for examining intermediate divergences within embryophytes. Both parsimony and maximum likelihood analyses produced similar topologies that differed only in the placement of hornworts and that generally support the classification of Kenrick and Crane (1997). The majority of embryophyte clades identified and named in that classification were also recovered in this analysis and most received high bootstrap support. Monophyletic groups included hornworts, liverworts, mosses, lycopsids, leptosporangiate ferns, eusporangiate ferns, gymnosperms, and angiosperms. Derived relationships that are not reflected in the Kenrick and Crane classification (but which have support from other studies) include a sister relationship between *Equisetum* and leptosporangiate ferns and monophyletic gymnosperms sister to angiosperms. Tree topologies and bootstrap support for clades are strongly affected by taxon inclusion/exclusion as demonstrated with the hornworts. Future molecular phylogenetic studies should employ more than a single hornwort exemplar taxon to avoid anomalous relationships. Whether liverworts or hornworts are the basalmost clade in the embryophyte clade cannot be distinguished based upon 19S rDNA sequences. Although the majority of molecular analyses support the hornworts basal hypothesis, both hypotheses should at present be considered equally viable. Heterogeneous substitution rates have been reported for particular lineages and particular genes in most embryophyte phylogenetic studies, thus the effects of long-branch attraction (Felsenstein, 1978) must be considered when interpreting the resulting relationships. In regards to ribosomal genes, attempts to account for rate heterogeneity across sites have been made (Van de Peer et al., 1993; Van de Peer, Chapelle, and De Wachter, 1996; Van de Peer, Van der Auwera, and De Wachter, 1996), and methods such as likelihood that provide more realistic models of sequence evolution are preferred. Our future work will be focused on constructing combined data sets that consist of various data partitions (from ribosomal as well as other genes) that can then be analyzed to determine the effect rate heterogeneity has on resolving embryophyte relationships.

LITERATURE CITED

BOPP, M., AND I. CAPESIUS. 1996. New aspects of bryophyte taxonomy provided by a molecular approach. *Botanica Acta* 109: 341–440.

BOWE, L. M., AND C. W. DEPAMPHILIS. 1996. Effects of RNA editing and gene processing on phylogenetic reconstruction. *Molecular Biology and Evolution* 13: 1159–1166.

BREMER, K. 1985. Summary of green plant phylogeny and classification. *Cladistics* 1: 369–385.

———. 1988. The limits of amino acid sequence data in angiosperm phylogenetic reconstruction. *Evolution* 42: 795–803.

———, C. J. HUMPHRIES, B. D. MISHLER AND S. P. CHURCHILL. 1987. On cladistic relationships in green plants. *Taxon* 36: 339–349.

———, AND H.-E. WANNTORP. 1981. A cladistic classification of green plants. *Nordic Journal of Botany* 1:1–3

CAPESIUS, I., AND M. BOPP. 1997. New classification of liverworts based on molecular and morphological data. *Plant Systematics and Evolution* 207: 87–97.

CHASE, M., AND V. A. ALBERT. 1998. A perspective on the contribution of plastid *rbcL* DNA sequences to angiosperm phylogenetics. *In* D. E. Soltis, P. S. Soltis, and J. J. Doyle [eds.], Molecular systematics of plants II. DNA sequencing. 488–507. Kluwer Academic Publishers, Boston, MA.

CHAW, S.-M., A. ZHARKIKH, H. M. SUNG, T. C. LAU AND W. H. LI. 1997. Molecular phylogeny of extant gymnosperms and seed plant evolution: analysis of nuclear 18S rRNA sequences. *Molecular Biology and Evolution* 14: 56–68.

CROWE, C., L. M. PIKE, K. S. CROSS, AND K. S. RENZAGLIA. 1997. The *psba* gene sequence can be used to infer phylogenetic relationships among the major lineages of bryophytes. *American Journal of Botany* 83: 14–15 (Abstract).

DUFF, R. J., AND D. L. NICKRENT. 1997. Characterization of mitochondrial small-subunit ribosomal RNAs from holoparasitic plants. *Journal of Molecular Evolution* 45: 631–639.

FELSENSTEIN, J. 1978. Cases in which parsimony or compatibility will be positively misleading. *Systematic Zoology* 27: 401–410.

———. 1985. Confidence limits on phylogenies: an approach using the bootstrap. *Evolution* 39: 783–791.

GARBARY, D. J., AND K. S. RENZAGLIA. 1998. Bryophyte phylogeny and the evolution of land plants: evidence from development and ultrastructure. *In* J. W. Bates, N. W. Ashton, and J. G. Duckett [eds.], Bryology for the twenty-first century. W. S. Mancy and Sons, Ltd., Leeds, UK.

———, ———, AND J. G. DUCKETT. 1993. The phylogeny of land

plants: a cladistic analysis based on male gametogenesis. *Plant Systematics and Evolution* 188: 237–269.

GILBERT, D. G. 1993. SeqApp, version 1.9a157. Biocomputing Office, Biology Department, Indiana University, Bloomington, IN.

GOREMYKIN, V., V. BOBROVA, J. PAHNKE, A. TROITSKY, A. ANTONOV, AND W. MARTIN. 1996. Noncoding sequences from the slowly evolving chloroplast inverted repeat in addition to *rbcL* data do not support gnetalean affinities of angiosperms. *Molecular Biology and Evolution* 13: 383–396.

GRAHAM, L. E. 1993. Origin of land plants. John Wiley & Sons, New York, NY.

———, C. F. DELWICHE, AND B. D. MISHLER. 1991. Phylogenetic connections between the ''green algae'' and ''bryophytes.'' *Advances in Bryology* 4: 213–244.

GRAYBEAL, A. 1994. Evaluating the phylogenetic utility of genes—a search for genes informative about deep divergences among vertebrates. *Systematic Biology* 43: 174–193.

GUTELL, R. R. 1994. Collection of small subunit (16S- and 16S-like) ribosomal RNA structures: 1994. *Nucleic Acids Research* 22: 3502–3507.

HASEBE, M., M. ITO, R. KOFUJI, K. IWATASUKI, AND K. UEDA. 1992. Phylogenetic relationships in Gnetophyta deduced from *rbcL* gene sequences. *Botanical Magazine, Tokyo* 105: 385–391.

———, P. G. WOLF, K. M. PRYER, K. UEDA, M. ITO, R. SANO, G. J. GASTONY, J. YOKOYAMA, J. R. MANHART, N. M. MURAKAMI, E. H. CRANE, C. H. HAUFLER, AND W. D. HAUK. 1995. Fern phylogeny based on *rbcL* nucleotide sequences. *American Fern Journal* 85: 134–181.

HASEGAWA, M., H. KISHINO, AND T. YANO. 1985. Dating of the human-ape splitting by a molecular clock of mitochondrial DNA. *Journal of Molecular Evolution* 21: 160–174.

HEDDERSON, T. A., R. CHAPMAN, AND C. J. COX. 1998. Bryophytes and the origins and diversification of land plants: new evidence from molecules. *In* J. W. Bates, N. W. Ashton, and J. G. Duckett [eds.], Bryology for the twenty-first century, 65–77. British Bryological Society.

———, ———, AND W. L. ROOTES. 1996. Phylogenetic relationships of bryophytes inferred from nuclear-encoded rRNA gene sequences. *Plant Systematics and Evolution* 200: 213–224.

HIESEL, R., A. VON HAESELER, AND A. BRENNICKE. 1994. Plant mitochondrial nucleic acid sequences as a tool for phylogenetic analysis. *Proceedings of the National Academy of Sciences, USA* 91: 634–638.

HILLIS, D. M., AND J. P. HUELSENBECK. 1992. Signal, noise, and reliability in molecular phylogenetic analyses. *Journal of Heredity* 83.

———, B. K. MABLE, AND C. MORITZ. 1996. Application of molecular systematics: the state of the field and a look to the future. *In* D. M. Hillis, C. Moritz, and B. K. Mable [eds.], Molecular systematics, 515–543. Sinauer, Sunderland, MA.

KENRICK, P., AND P. R. CRANE. 1991. Water-conducting cells in early fossil land plants: implications for the early evolution of tracheophytes. *Botanical Gazette* 152: 335–356.

———, AND ———. 1997. The origin and early diversification of land plants: a cladistic study. Smithsonian Institute Press, Washington, DC.

KISHINO, H., AND M. HASEGAWA. 1989. Evaluation of the maximum likelihood estimate of the evolutionary tree topologies from DNA sequence data, and the branching order of Hominoidea. *Journal of Molecular Evolution* 29: 170–179.

KRANZ, H. D., D. MIKS, M.-L. SIEGLER, I. CAPESIUS, AND C. W. SENSEN. 1995. The origin of land plants: phylogenetic relationships among charophytes, bryophytes, and vascular plants inferred from complete small-subunit ribosomal RNA gene sequences. *Journal of Molecular Evolution* 41: 74–84.

LANYON, S. M. 1985. Detecting internal inconsistencies in distance data. *Systematic Zoology* 34: 397–403.

LEWIS, L. A., B. D. MISHLER, AND R. VILGALYS. 1997. Phylogenetic relationships of the liverworts (Hepaticae), a basal embryophyte lineage, inferred from nucleotide sequence data of the chloroplast gene *rbcL*. *Molecular Phylogenetics and Evolution* 7: 377–393.

MADDISON, D. R. 1991. The discovery and importance of multiple islands of most-parsimonious trees. *Systematic Zoology* 40: 315–328.

MADEN, A. R., D. P. WHITTIER, D. J. GARBARY, AND K. S. RENZAGLIA. 1997. Ultrastructure of the spermatozoid of *Lycopodiella lateralis* (Lycopodiaceae). *Canadian Journal of Botany* 75:1728–1738.

MALEK, O., K. LÄTTIG, R. HIESEL, A. BRENNICKE, AND V. KNOOP. 1996. RNA editing in bryophytes and a molecular phylogeny of land plants. *EMBO Journal* 15: 1403–1411.

MANHART, J. R. 1994. Phylogenetic analysis of green plant *rbcL* sequences. *Molecular Phylogenetics and Evolution* 3: 114–127.

———. 1995. Chloroplast 16S rDNA sequences and phylogenetic relationships of fern allies and ferns. *American Fern Journal* 85: 182–192.

MISHLER, B. D., AND S. P. CHURCHILL. 1984. A cladistic approach to the phylogeny of the bryophytes. *Brittonia* 36: 406–424.

———, AND ———. 1985. Transition to a land flora: phylogenetic relationships of the green algae and bryophytes. *Cladistics* 1: 305–328.

———, L. A. LEWIS, M. A. BUCHHEIM, K. S. RENZAGLIA, D. J. GARBARY, C. F. DELWICHE, F. W. ZECHMAN, T. S. KANTZ, AND R. L. CHAPMAN. 1994. Phylogenetic relationships of the ''green algae'' and ''bryophytes.'' *Annals of the Missouri Botanical Garden* 81: 451–483.

MIYAMOTO, M. M., AND J. CRACRAFT. 1991. Phylogenetic inference, DNA sequence analysis, and the future of molecular systematics. *In* M. M. Miyamoto and J. Cracraft [eds.], Phylogenetic analysis of DNA sequences, 1–17. Oxford University Press, New York, NY.

NICKRENT, D. L. 1994. From field to film: rapid sequencing methods for field collected plant species. *BioTechniques* 16: 470–475.

———, R. J. DUFF, A. E. COLWELL, A. D. WOLFE, N. D. YOUNG, K. E. STEINER, AND C. W. DEPAMPHILIS. 1998. Molecular phylogenetic and evolutionary studies of parasitic plants. *In* D. E. Soltis, P. S. Soltis, and J. J. Doyle [eds.], Molecular systematics of plants II. DNA sequencing. 211–241. Kluwer Academic Publishers, Boston, MA.

ODO, K., K. YAMATO, E. OHTA, Y. NAKAMURA, M. TAKEMURA, N. NOZATO, K. AKASHI, T. KANEGAE, Y. OGURA, T. KOHCHI, AND K. OHYAMA. 1992. Gene organization deduced from the complete sequence of liverwort *Marchantia polymorpha* mitochondrial DNA. *Journal of Molecular Biology* 223: 1–7.

OLSEN, G. J. 1987. Earliest phylogenetic branchings: comparing rRNA-based evolutionary trees inferred with various techniques. *Cold Spring Harbor Symposium on Quantitative Biology* 52: 825–837.

PALMER, J. D. 1992. Mitochondrial DNA in plant systematics: applications and limitations. *In* P. S. Soltis, D. E. Soltis, and J. J. Doyle [eds.], Molecular systematics of plants, 36–49. Chapman Hall, New York, NY.

———, C. A. MAKAROFF, I. J. APEL, AND M. SHIRZADEGAN. 1990. Fluid structure of plant mitochondrial genomes: evolutionary and functional implications. *In* M. T. Clegg and S. J. O'Brien [eds.], Molecular evolution, 85–96. Alan R. Liss, New York, NY.

PARENTI, L. R. 1980. A phylogenetic analysis of the land plants. *Biological Journal of the Linnaean Society* 13: 225–242.

PENNY, D., M. D. HENDY, AND M. A. STEEL. 1991. Testing the theory of descent. *In* M. M. Miyamoto and J. Cracraft [eds.], Phylogenetic analysis of DNA sequences, 155–183. Oxford University Press, New York, NY.

PRICE, R. A. 1996. Systematics of the Gnetales: a review of morphological and molecular evidence. *International Journal of Plant Sciences* 157 (6, Supplement): 40–49.

PRYER, K. M., A. R. SMITH, AND J. E. SKOG. 1995. Phylogenetic relationships of extant ferns based on evidence from morphology and *rbcL* sequences. *American Fern Journal* 85: 205–282.

QIU, Y.-L., Y. CHO, J. C. COX, AND J. D. PALMER. 1998. The gain of three mitochondrial introns identifies liverworts as the earliest land plants. *Nature (London)* 394: 671–674.

RAUBESON, L. A., AND R. K. JENSEN. 1992. Chloroplast DNA evidence on the ancient evolutionary split in vascular land plants. *Science* 255: 1697–1699.

ROBINSON, H. 1985. Comments on the cladistic approach to the phylogeny of the ''bryophytes'' by Mishler and Churchill. *Brittonia* 37: 279–281.

VAN DE PEER, Y., S. CHAPELLE, AND R. DE WACHTER. 1996. A quan-

titative map of nucleotide substitution rates in bacterial rRNA. *Nucleic Acids Research* 24: 3381–3391.

———, J.-M. NEEFS, P. D. RIJK, AND R. D. WACHTER. 1993. Reconstructing evolution from eukaryotic small-ribosomal-subunit RNA sequences: calibration of the molecular clock. *Journal of Molecular Evolution* 37: 221–232.

———, G. VAN DER AUWERA, AND R. DE WACHTER. 1996. The evolution of stramenopiles and alveolates as derived by ''substitution rate calibration'' of small ribosomal subunit RNA. *Journal of Molecular Evolution* 42: 201–210.

WATERS, D. A., M. A. BUCHHEIM, R. A. DEWEY, AND R. L. CHAPMAN. 1992. Preliminary inferences of the phylogeny of bryophytes from nuclear-encoded ribosomal RNA sequences. *American Journal of Botany* 79: 459–466.

WHITTEMORE, A. T. 1987. Transition to a land flora: a critique. *Cladistics* 3: 60–65.

WOESE, C. R. 1987. Bacterial evolution. *Microbiological Review* 51: 221–271.

WOLF, P. G. 1997. Evaluation of *atpB* nucleotide sequences for phylogenetic studies of ferns and other pteridophytes. *American Journal of Botany* 84: 1429–1440.

———, P. S. SOLTIS, AND D. E. SOLTIS. 1994. Phylogenetic relationships of dennstaedtioid ferns: evidence from *rbcL* sequences. *Molecular Phylogenetics and Evolution* 3: 383–392.

WOLFF, G., AND U. KÜCK. 1990. The structural analysis of the mitochondrial SSU rRNA implies a close phylogenetic relationship between mitochondria from plants and from the heterotrophic alga *Prototheca wickerhamii*. *Current Genetics* 17: 347–351.