A Comparison of Angiosperm Phylogenies from Nuclear 18S rDNA and rbcL Sequences

Daniel L. Nickrent, Douglas E. Soltis

Annals of the Missouri Botanical Garden, Volume 82, Issue 2 (1995), 208-234.

Stable URL:
http://links.jstor.org/sici?sici=0026-6493%281995%2982%3A2%3C208%3AACOAPF%3E2.0.CO%3B2-U

Your use of the JSTOR archive indicates your acceptance of JSTOR's Terms and Conditions of Use, available at http://www.jstor.org/about/terms.html. JSTOR's Terms and Conditions of Use provides, in part, that unless you have obtained prior permission, you may not download an entire issue of a journal or multiple copies of articles, and you may use content in the JSTOR archive only for your personal, non-commercial use.

Each copy of any part of a JSTOR transmission must contain the same copyright notice that appears on the screen or printed page of such transmission.

Annals of the Missouri Botanical Garden is published by Missouri Botanical Garden Press. Please contact the publisher for further permissions regarding the use of this work. Publisher contact information may be obtained at http://www.jstor.org/journals/mobot.html.

Annals of the Missouri Botanical Garden
©1995 Missouri Botanical Garden Press

JSTOR and the JSTOR logo are trademarks of JSTOR, and are Registered in the U.S. Patent and Trademark Office. For more information on JSTOR contact jstor-info@umich.edu.

©2002 JSTOR

http://www.jstor.org/
Fri Feb 15 16:06:49 2002

A COMPARISON OF ANGIOSPERM PHYLOGENIES FROM NUCLEAR 18S rDNA AND *rbc*L SEQUENCES[1]

Daniel L. Nickrent[2] *and Douglas E. Soltis*[3]

Abstract

To investigate the phylogenetic utility of entire, nuclear-encoded small-subunit (18S) ribosomal DNA sequences, we compared the rate of evolution and phylogenetic resolution of entire 18S sequences with those for the chloroplast gene *rbc*L using a suite of 59 angiosperms and 3 gymnosperms (*Gnetum*, *Ephedra*, and *Zamia*) as outgroups. For *rbc*L, 482 (33.6%) of the 1431 base positions were phylogenetically informative, whereas for 18S rDNA 341 (18.4%) of the 1853 positions were informative. Pairwise comparisons within the angiosperms show that *rbc*L is generally about three times more variable than 18S rDNA. However, because the 18S region is approximately 400 base pairs longer than *rbc*L, the ratio of the number of phylogenetically informative sites per molecule is only about 1.4 times greater for *rbc*L compared to 18S rDNA. Not only are sites more variable in *rbc*L than in 18S rDNA, but this variability is more evenly distributed over the length of *rbc*L. In contrast, 18S rDNA shows highly variable regions interspersed with regions of extreme conservation. Minimum-length Fitch trees were constructed for each matrix, and the results were compared to a tree derived from a previous global analysis of *rbc*L sequences based on 499 seed plants. Parsimony analyses showed that several clades are strongly supported by both data sets, such as Gnetales, monocots, paleoherbs, Santalales, and various clades within Rosidae s.l. and Asteridae s.l. Some clades (e.g., Santalales) have higher base substitution rates for 18S rDNA, permitting the assessment of inter- and intrafamilial relationships. This comparative study indicates that 18S rDNA sequences contain sufficient information to conduct phylogenetic studies at higher taxonomic levels (family and above) within angiosperms. rDNA sequences are best applied to such deep divergences, but the amount of variation differs significantly among taxonomic groups.

The major morphologically based angiosperm classifications proposed during the past 15 years show marked similarities, yet also differ in fundamental ways. Recently, systematists have explored the utility of both DNA and RNA sequencing to resolve higher-level relationships within the angiosperms, as well as seed plants in general (e.g., Hamby & Zimmer, 1992; Chase et al., 1993). The largest molecular phylogenetic study conducted to date (Chase et al., 1993) employed sequence data for the chloroplast gene *rbc*L and was based on sequences for 499 species of angiosperms and other seed plants. The gene *rbc*L is typically 1428 base pairs in length, and the advantages of using this gene in phylogenetic reconstruction have been thoroughly reviewed (e.g., Ritland & Clegg, 1987; Palmer et al., 1988; Chase et al., 1993). These advantages include easy amplification via the polymerase chain reaction, essentially no insertion-deletion events, appropriate length and base substitution rate for inferring phylogeny at higher levels, and the availability of a set of sequencing primers (provided free of charge by G. Zurawski). Although some variation in the rate of *rbc*L sequence evolution occurs from lineage to lineage (Bousquet et al., 1992; Gaut et al., 1992), unequal rates of evolution do not appear to be sufficient to obscure major phylogenetic relationships (Chase et al., 1993). Because of these numerous advantages, *rbc*L sequences now exist for over 1500 taxa (M. Chase, pers. comm.), making *rbc*L the most frequently sequenced protein-coding gene. During the past several years, the phylogenetic analysis of *rbc*L sequences has provided unprecedented insights into higher-level relationships in angiosperms and gymnosperms (e.g., Chase et al., 1993; Conti et al., 1993; Duvall et al., 1993; Michaels et al., 1993; Morgan & Soltis, 1993; Qiu et al., 1993).

Because most evolutionary studies are devoid of positive controls to prove or disprove particular events, the strongest support that can be obtained in phylogenetic reconstruction is congruence re-

[1] Laboratory assistance was provided by Robert Kuzoff, Ellen Starr, Mark Berning, Brent Beilschmidt, Joseph Agan, and Ann Marie Brown. Gerard Zurawski kindly supplied several *rbc*L primers and Mark Chase provided several unpublished *rbc*L sequences. Genomic DNA of *Saruma henryi* was donated by Y.-L. Qiu and W. C. Dickison. This work was supported by NSF grants BSR-8918385 (to DLN) and BSR-9007614 (to DES).

[2] Department of Plant Biology, Southern Illinois University, Carbondale, Illinois 62901, U.S.A.

[3] Department of Botany, Washington State University, Pullman, Washington 99164, U.S.A.

sulting from analysis of multiple independent data sets (Miyamoto & Cracraft, 1991). Hypotheses of relationships are either strongly or weakly supported based upon statistical tests involving the data themselves (e.g., consistency index, bootstrap values, decay values). The more numerous and more varied the data sets that corroborate a given relationship, the greater the support for that relationship. Although *rbc*L and, more recently, other chloroplast genes such as *mat*K (see Johnson & Soltis, 1995, this issue) and *ndh*F (see Olmstead & Reeves, 1995, this issue) have been shown to have great utility in phylogeny estimation, many workers have emphasized the need for comparison of chloroplast DNA-based phylogenetic trees with those from other sources, especially those based on sequences from nuclear-encoded genes (e.g., Palmer, 1985; Rieseberg & Soltis, 1991; Doyle, 1992; Friedlander et al., 1992; Chase et al., 1993). At lower taxonomic levels (genus and below) comparative sequencing of the nuclear internal transcribed spacer (ITS) region has shown tremendous potential for inferring phylogenies (see Baldwin et al., 1995, this issue) and has stimulated the comparison of phylogenies based on chloroplast and nuclear DNA. At higher taxonomic levels, the phylogenetic trees presented for angiosperms based on *rbc*L sequences (Chase et al., 1993) have similarly stimulated interest in conducting a comparable phylogenetic analysis based on nuclear gene sequences. In this paper we explore the utility of entire nuclear 18S rDNA sequences for inferring phylogeny at higher levels within the angiosperms.

In plants, ribosomes exist in the chloroplasts, mitochondria, and cytoplasm and are composed of a small and a large subunit, each of which contains rRNA and associated proteins. Although sedimentation coefficients vary slightly, plant nuclear small-subunit rRNA will be referred to here as the 18S rRNA. The 18S, 5.8S, and 26S nuclear rRNA genes occur as a unit (cistron) separated by spacer regions. These cistrons are repeated hundreds to thousands of times in tandem arrays within the genome (Appels & Honeycutt, 1986). Ribosomal RNA cistrons are usually located in the nucleolar organizing region of the nucleus and may reside on several different chromosomes in plants (Thompson & Flavell, 1988). Sequence similarity between the individual cistrons within a single organism is extremely high, possibly due to unequal crossing over during meiosis, gene conversion, slippage, transposition, and RNA-mediated changes (Arnheim et al., 1980; Dover, 1982; Arnheim, 1983; Dover, 1987). The homogeneity of ribosomal RNA cistrons has been referred to as concerted evolution (Brown et al., 1972; Arnheim et al., 1980; Zimmer et al., 1980). Ribosomal loci represent an extreme type of concerted evolution (with essentially complete homogenization), making them advantageous for reconstruction of deep phylogenetic events (Sanderson & Doyle, 1993). Recent summaries of ribosomal RNA structure, function, gene organization, and evolution have been presented (Jorgansen & Cluster, 1988; Hillis & Dixon, 1991; Hamby & Zimmer, 1992).

Numerous low-molecular-weight (5S and 5.8S) rRNA sequences now exist (see compilation by Specht et al., 1991), and attempts have been made to use these sequences in addressing the origin and evolution of green plants (Hori et al., 1985; Hori & Osawa, 1987). However, because these molecules are less than 200 bp in length, they provide a very limited number of phylogenetically informative sites; hence, large numbers of equally parsimonious solutions often result when conducting studies using many taxa (Bremer et al., 1987). Specifically addressing 5S rRNA sequences, Mishler et al. (1988) summarized concerns for the use of rRNA sequences for phylogenetic reconstruction that apply to the 18S and 26S as well, such as cosubstitution in stem regions of helices, transition/transversion bias, alignment problems, different evolutionary rates, and homoplasy.

Both large- and small-subunit ribosomal RNA sequences have been used to examine the very deepest branches among organisms, such as the domains Eukarya, Bacteria, and Archae (Wolters & Erdmann, 1986; Olsen, 1987; Woese, 1987). Ribosomal RNA sequence data have also been used to elucidate phylogenetic relationships in animals (e.g., Sogin et al., 1986; Field et al., 1988; Wada & Satoh, 1994), protozoa (Schlegel et al., 1991), algae (Bhattacharya & Druehl, 1988; Buchheim et al., 1990; Huss & Sogin, 1990; Kantz et al., 1990; Hendriks et al., 1991; Chapman & Buchheim, 1991), and fungi (Förster et al., 1990; Swann & Taylor, 1993). Prior to 1990, most rRNA sequences were being determined from cloned material or by using Sanger dideoxynucleotide reactions and reverse transcriptase with rRNA templates (Lane et al., 1985). During the past several years, most workers have moved to direct sequencing of ribosomal DNA (rDNA) amplified via the polymerase chain reaction (PCR; Mullis & Faloona, 1987). The major reasons for the shift to DNA sequencing are: (1) rRNA is labile to RNases, making it methodologically difficult to extract, purify, and store; (2) rRNA secondary structure causes polymerase stalling, visualized as "hard stops" on sequencing gels, resulting in ambiguous sequence;

(3) when RNA is extracted from a tissue sample only RNA genes can be sequenced, whereas, theoretically, any gene (nuclear, plastid, mitochondrial) is available from total genomic DNA extracts; (4) DNA is easier to extract and is more stable than RNA; and (5) with DNA, both strands are available for sequencing, allowing more complete coverage of the molecule (as well as an opportunity to double-check each base position) by using primers designed for both the coding and noncoding strands.

The first 18S ribosomal RNA sequences of angiosperms were of rice (Takaiwa et al., 1984), maize (Messing et al., 1984), and soybean (Eckenrode et al., 1985). Later, the complete 18S rDNA sequence of the cycad *Zamia pumila* was published; with the above three higher plants and several outgroups, a phylogenetic tree was produced (Nairn & Ferl, 1988). Seven additional (partial) 18S rRNA sequences of members of the Poaceae were later determined, and a phylogenetic analysis of this family was conducted (Hamby & Zimmer, 1988).

Phylogenetic relationships in the parasitic plant order Santalales were examined by Nickrent & Franchina (1990). This study was the first since the work by Nairn & Ferl (1988) to use essentially complete 18S rRNA sequences in a phylogenetic analysis. Sequences from 13 angiosperm species representing 10 families were analyzed, and one most parsimonious tree was obtained that supported the monophyly of the order Santalales, confirmed the basal position of Olacaceae within the order, and showed Viscaceae to be derived from Santalaceae. This study indicated that sufficient information exists in complete 18S rRNA sequences to allow phylogenetic comparisons to be made at the family level and above.

Despite the phylogenetic promise of these initial analyses, relatively few 18S rRNA sequences were determined in the years that followed, perhaps due in part to the tremendous interest in *rbc*L sequencing for inferring phylogeny at this same level. As a result, the phylogenetic potential of 18S sequence data remained unexplored. A few entire sequences were published, including *Alnus glutinosa* (Savard & Lalonde, 1991), *Arabidopsis thalliana* (Unfried et al., 1989), *Lycopersicon esculentum* (Kiss et al., 1989), and *Sinapis alba* (Rathgeber & Capesius, 1990). A sequence for *Fragaria* × *ananassa* (Simovic et al., 1992) exists, but contains a large number of base changes atypical of other Rosaceae (and was therefore not included in the present study). Several studies, however, explored the phylogenetic potential of partial 18S sequences.

For example, Troitsky et al. (1991) used five different rRNA molecules (including nuclear 18S rRNA) to examine the early evolution of seed plants. For 18S rRNA, 21 sequences representing 256 bp (from position 499 to 755 on *Glycine*) were used. Six dicots and eight monocots were included, as were *Ephedra* and *Gnetum*, two cycads, two gymnosperms (*Podocarpus* and *Taxus*), and *Lycopodium*. Two conclusions of this study were that the divergence of angiosperms from gymnosperms occurred before the early Carboniferous, i.e., at least 360 million years before present, and that the Gnetopsida are not monophyletic. Given the small number of base pairs used and that no statistical support for the clades was provided, these results must be viewed with caution. More recently, Chaw et al. (1993) used 18S rDNA sequence data to demonstrate support for the placement of *Taxus* in Coniferales; however, only four sequences (for *Taxus*, *Pinus*, *Podocarpus*, and *Ginkgo*) and an outgroup (*Zamia*) were used in the analysis.

By far the largest analyses of 18S sequences have been undertaken by Zimmer and her collaborators, who conducted phylogenetic studies using direct sequencing of rRNA from approximately 60 vascular plant species (Zimmer et al., 1989; Hamby & Zimmer, 1992). Their efforts toward producing a molecular phylogeny of the angiosperms were based on the sequencing of a portion of the small- (18S) and large- (26S) subunit rRNA molecules. Hamby & Zimmer (1992) used a total of 1701 base positions per taxon (1097 base positions from the 18S region and 604 positions from the 26S region) in a phylogenetic analysis of seed plants that included 29 dicot and 17 monocot genera. Two shortest trees were found with a large number of equally parsimonious solutions one or several steps longer. The shortest trees had a number of features in accord with various existing classifications, such as the presence of a monophyletic Gnetales clade as sister to the angiosperms, and the basal position within the angiosperms of several Magnoliid groups, such as Nymphaeaceae and Piperaceae. Sampling within nonmagnoliid groups was sparse, however, which could explain the unusual relationships suggested among more derived angiosperms (e.g., the presence of a clade composed of *Ranunculus*, *Duchesnea*, *Spinacia*, and *Stellaria*). Because many of the interior and basal dicot nodes were poorly supported in the rRNA tree, systematists remained unsure of the utility of ribosomal RNA sequences for resolving questions of angiosperm phylogeny. More recently, relationships among the tribes of Onagraceae were examined by Bult & Zimmer (1993) using partial

sequences of 18S and 26S rRNA. Although relatively few phylogenetically informative sites were found, several relationships were in accord with those revealed by *rbc*L analysis (Conti et al., 1993).

Using the same primers as the Zimmer group, Martin & Dowd (1991) obtained partial 18S and 26S rRNA sequences for 12 angiosperm species from 7 families. Their purpose was to assess the relative merits of phylogeny estimation using ribosomal sequences with those derived from *rbc*L. The authors concluded "both phylogenetic trees gave good grouping within families but in neither case was there resolution of the branching order of major taxa. . . ." The authors further stated "that neither macromolecule alone was likely to yield a solution to the problem of angiosperm phylogeny and therefore that studies of both, at least, will be required." However, aside from two familial placeholders, only two species (maize and rice) were shared by the two data sets; furthermore, taxon density was clearly a limitation in making any statements regarding the branching order of major taxa. Additional studies using partial 18S rRNA sequences include analyses of six angiosperm families (Boulter & Gilroy, 1992), Papilionaceae (Martin & Dowd, 1993), and Byblidaeae–Roridulaceae (Conran & Dowd, 1993). The latter study examined the phylogenetic placement of two carnivorous genera, *Roridula* and *Byblis*, that have been variously classified using morphological and chemical characters. The analysis of partial 18S rRNA sequences from these genera and 26 other angiosperms supported the position of *Roridula* in the lower Rosidae and the placement of *Byblis* in the Asteridae, results in agreement with *rbc*L sequence analysis (Albert et al., 1992).

The major goal of this study was to explore in more detail the potential of entire 18S rDNA sequences for inferring phylogeny at higher levels in the angiosperms. We wanted to understand better the rate of evolution and distribution of base substitutions of 18S rDNA compared to the widely used chloroplast gene *rbc*L. To accomplish these objectives, comparable 18S and *rbc*L sequence data sets were constructed for a similar suite of 59 angiosperms and 3 gymnosperms. Minimum-length Fitch trees were constructed, and relationships as well as evolutionary rates were compared for the entire 18S gene and *rbc*L. These phylogenetic analyses were not meant to be exhaustive studies of angiosperm relationships; we recognize that taxonomic density is a limitation in this study. Rather, our goal was to assess the relative merits and attributes of each molecule through direct comparison. This study differs from previous comparative analyses of 18S and *rbc*L sequences (e.g., Martin & Dowd, 1991) in that essentially complete sequences of both genes were used, taxon sampling was more extensive, and, in the majority of cases, the same taxon was sequenced for both genes.

Materials and Methods

Taxon sampling and sequence acquisition

Given the large, taxonomically diverse array of *rbc*L sequences that is already available, taxon inclusion for this study was determined mainly by the availability of 18S rRNA or rDNA sequences. We therefore determined the 18S rDNA sequences of additional plant taxa to provide greater coverage of the major clades identified in the global *rbc*L analysis of Chase et al. (1993). The *rbc*L sequences were chosen to correspond at the specific or, secondarily, the generic level to an available sequence of 18S rRNA. Some *rbc*L sequences included in Chase et al. (1993), but not deposited in Genbank, were kindly provided by Mark Chase (*Brassica, Pachysandra, Pisum,* and *Impatiens*). An *rbc*L sequence of *Hydrocotyle* that was not included in the Chase et al. analysis was also kindly provided by G. Plunkett. The original *rbc*L sequence for *Zea* contained errors; hence the newly determined sequence (Gaut et al., 1992) was used herein. In addition to the three published *rbc*L sequences of Santalales (Morgan & Soltis, 1993), 12 other *rbc*L sequences were determined to increase sampling within this one order, thereby allowing phylogeny comparison of more closely related species using both molecules. A data set of 62 taxa was ultimately identified for which sequences of both molecules were available for use in this study (Tables 1 and 2). Of the 62 sequence pairs (18S rDNA and *rbc*L) used herein, 37 were from the same species; an additional 15 sequences were from different species within the same genus. Different genera were used for ten families to allow a broader sampling within the angiosperms (*rbc*L/18S rDNA): *Betula/Alnus, Pachysandra/Buxus, Convolvulus/Cuscuta, Polemonium/Gilia, Pisum/Glycine, Lambertia/Knightia, Reinwardtia/Linum, Byrsonima/Malpighia, Pyrola/Monotropa,* and *Mahonia/Podophyllum.* Although members of three of the above generic pairs are classified in separate families (i.e., Cuscutaceae/Convolvulaceae, Monotropaceae/Pyrolaceae, Podophyllaceae/Berberidaceae), they were deemed to be related closely enough (based upon traditional classifications) to be used as placeholders. Following Cronquist (1981), 47 angiosperm families are represented in this study.

Seven of the 62 18S rDNA sequences used in

TABLE 1. Voucher information for 18S rDNA sequences.

Species*	Family	Voucher	Reference† (source)	Sequence	GenBank number
Acorus calamus L.	Araceae	Nickrent 2942, SIU	†(Nickrent)	6–1761	L24078
Akebia quinata (Houtt.) Decne.	Lardizabalaceae	Nickrent 2945, SIU	†(Nickrent)	66–1705	L31795
Alnus glutinosa (L.) Gaertn.*	Betulaceae	unknown	Savard & Lalonde, 1991	79–1782	X54984
Antidaphne viscoidea (Griseb.) Kuijt	Eremolepidaceae	Nickrent 2730, SIU	†(Nickrent)	6–1784	L24080
Arceuthobium verticilliflorum Englem.	Viscaceae	Nickrent 2065, SIU	Nickrent & Starr, 1994	10–1808	L24042
Aristolochia tomentosa Sims.	Aristolochiaceae	Nickrent 2922, SIU	†(Nickrent)	5–1765	L24083
Asarum canadense L.	Aristolochiaceae	Nickrent 2888, SIU	Nickrent & Starr, 1994	6–1761	L24043
Brassica hirta Moench [= *Sinapis alba* L.]	Brassicaceae	unknown	Rathgeber & Capesius, 1990	1–1804	X17062
Buxus sempervirens L.*	Buxaceae	Nickrent 2890, SIU	Nickrent & Franchina, 1990	2–1747	X16599
Chrysosplenium iowense Rydb.	Saxifragaceae	Wendel *s.n.*, ISC	†(Soltis)	23–1778	L28136
Cornus florida L.	Cornaceae	Nickrent 2897, SIU	Nickrent & Franchina, 1990	3–1757	X17370
Cuscuta gronovii Willd.*	Cuscutaceae	Nickrent 3015, SIU	†(Nickrent)	6–1764	L24747
Dendrophthora clavata (Benth.) Urban	Viscaceae	Nickrent 2182, SIU	†(Nickrent)	5–1765	L24086
Drimys winteri Forster & Forster	Winteraceae	Nickrent 2951, SIU	†(Nickrent)	9–1763	L24089
Ephedra antisyphilitica BErl. ex C. A. May	Ephedraceae	Nickrent 2995, SIU	†(Nickrent)	11–1767	L24091
Eubrachion ambiguum (Hooker & Arnott) Engl.	Eremolepidaceae	Nickrent 2699, SIU	†(Nickrent)	4–1759	L24141
Euonymus alatus (Thunb.) Siebold	Celastraceae	Nickrent 2894, SIU	Nickrent & Franchina, 1990	3–1757	X16600
Francoa sonchifolia Cav.	Saxifragaceae	Soltis & Soltis 2479, WS	†(Soltis)	23–1764	L28137
Gaiadendron punctatum (R. & P.) G. Don	Loranthaceae	Nickrent 2729, SIU	†(Nickrent)	13–1761	L24143
Gilia capitata Sims.*	Polemoniaceae	Johnson 9215, WS	†(Soltis)	24–1770	L28138
Ginalloa arnottiana Korth.	Viscaceae	Beaman 9074, K	†(Nickrent)	7–1766	L24144
Glycine max (L.) Merrill*	Fabaceae	unknown	Eckenrode et al., 1985	1–1807	X02623
Gnetum leyboldii Tul.	Gnetaceae	Nickrent 2946, SIU	Nickrent & Starr, 1994	10–1812	L24045
Gossypium hirsutum L.	Malvaceae	Nickrent 2904, SIU	†(Nickrent)	13–1757	L24145
Hedera helix L.	Araliaceae	Nickrent 2887, SIU	Nickrent & Franchina, 1990	4–1759	X16604
Heuchera micrantha Dougl. ex Lindl.	Saxifragaceae	Soltis & Soltis 1949, WS	†(Soltis)	21–1778	X28139
Houttuynia cordata Thunb.	Saururaceae	Nickrent 2940, SIU	†(Nickrent)	6–1764	L24147
Hydrangea macrophylla (Thunb.) Ser.	Hydrangeaceae	Morgan 2150, WS	†(Soltis)	24–1528	L28140
Hydrocotyle sibthorpioides Lam.	Apiaceae	Nickrent 2886, SIU	†(Nickrent)	3–1735	X16605
Impatiens pallida Nutt.	Balsaminaceae	Nickrent 2938, SIU	†(Nickrent)	9–1758	L24148
Knightia excelsa R. Br.*	Proteaceae	Soltis *s.n.*	†(Nickrent)	6–1761	L24155
Korthalsella lindsayi (D. Oliv.) Engl.	Viscaceae	Nickrent 2740, SIU	†(Nickrent)	6–1710	L24150

TABLE 1. Continued.

Species*	Family	Voucher	Reference† (source)	Sequence	GenBank number
Lepuropetalon spathulatum (Muhl.) Elliott	Saxifragaceae	Thomas *s.n.*, NLU	†(Soltis)	23–1764	L28141
Linum perenne L.*	Linaceae	Nickrent 2900, SIU	†(Nickrent)	11–1746	L24401
Lycopersicon esculentum Mill.	Solanaceae	unknown	Kiss et al., 1989	1–1800	X51576
Malpighia coccigera L.*	Malpighiaceae	Nickrent 2903, SIU	Nickrent & Starr, 1994	15–1760	L24046
Monotropa uniflora L.*	Monotropaceae	Nickrent 3018, SIU	†(Nickrent)	6–1764	L26062
Morus alba L.	Moraceae	Nickrent 2924, SIU	†(Nickrent)	6–1759	L24398
Misodendron brachystachyum D.C.	Misodendraceae	Nickrent 2829, SIU	†(Nickrent)	5–1761	L24397
Notothixos subaureus Oliver	Viscaceae	Nickrent 2790, SIU	†(Nickrent)	5–1717	L24402
Nymphaea tuberosa Paine	Nymphaeaceae	Nickrent 2906, SIU	†(Nickrent)	1–1759	L24404
Nyssa sylvatica Marsh	Nyssaceae	Nickrent 2905, SIU	Nickrent & Franchina, 1990	3–1751	X16603
Opilia amentacea Roxb.	Opiliaceae	Nickrent 2816, SIU	†(Nickrent)	4–1760	L24407
Oryza sativa L.	Poaceae	unknown	Takaiwa et al., 1984	1–1812	X00755
Osyris lanceolata Hochst. & Steud.	Santalaceae	Nickrent 2731, SIU	†(Nickrent)	5–1801	L24409
Paeonia lactiflora Pall.	Paeoniaceae	Nickrent 2988, SIU	†(Nickrent)	6–1763	L24410
Peperomia serpens (Swartz) Loud.	Piperaceae	Nickrent 2907, SIU	†(Nickrent)	32–1761	L24411
Phoradendron serotinum (Raf.) M. C. Johnst.	Viscaceae	Nickrent 2077, SIU	Nickrent & Franchina, 1990	3–1769	X16607
Pittosporum japonicum Hort. ex C. Presl.	Pittosporaceae	Riesberg *s.n.*, RSA	†(Soltis)	39–1775	L28142
Podophyllum peltatum L.*	Berberidaceae s.l.	Nickrent 2891, SIU	†(Nickrent)	4–1766	L24413
Prunus persica (L.) Batsch.	Rosaceae	E. E. Dickson, Bailey Hort.	†(Soltis)	23–1778	L28749
Ranunculus sardous Crantz	Ranunculaceae	Nickrent 2932, SIU	†(Nickrent)	6–1761	L24092
Ribes aureum Pursh	Grossulariaceae	Soltis & Soltis 2220, WS	†(Soltis)	29–1770	L28143
Santalum album L.	Santalaceae	Nickrent 2734, SIU	†(Nickrent)	5–1764	L24416
Saruma henryi Oliver	Aristolochiaceae	Qiu 91018, NCU	†(Nickrent)	5–1765	L24417
Schoepfia schreberi Gemlin	Olacaceae	Nickrent 2599, SIU	†(Nickrent)	7–1758	L24418
Sparganium eurycarpum Engl.	Sparganiaceae	Nickrent 2943, SIU	†(Nickrent)	6–1749	L24419
Spinacia oleracea L.	Chenopodiaceae	Nickrent 2896, SIU	†(Nickrent)	8–1764	L24420
Tropaeolum majus L.	Tropaeolaceae	Chase 113, NCU	†(Soltis)	24–1776	L28750
Viscum album L.	Viscaceae	Nickrent 2253, SIU	†(Nickrent)	4–1762	L24426
Zea mays L.	Poaceae	unknown	Messing et al., 1984	1–1809	K02202
Zamia pumila L.	Zamiaceae	unknown	Nairn & Ferl, 1988	1–1813	M20017

† Sequence first reported here.
* Different genera used as placeholders for family for phylogenetic analyses.

TABLE 2. Voucher information for *rbc*L sequences.

Species	Family	Voucher	Reference (source)	Sequence	GenBank number
Acorus calamus L.	Araceae	French 232, CH	Duvall et al., 1993 (Duvall)	29–1270, 1319–1352	M91625
Akebia quinata (Houtt.) Decne.	Lardizabalaceae	Qiu 91020, NCU	Chase et al. 1993 (Qiu)	29–1428	L12627
Antidaphne viscoidea (Griseb.) Kuijt	Eremolepidaceae	Nickrent 2730, SIU	†(Nickrent)	4–1428	L26068
Arceuthobium verticilliflorum Engelm.	Viscaceae	Nickrent 2065, SIU	†(Nickrent)	3–1428	L26067
Aristolochia macrophylla Lam.	Aristolochiaceae	Qiu 91019, NCU	Qiu et al., 1993	29–1428	L12630
Asarum canadense L.	Aristolochiaceae	none	Chase et al., 1993 (Olmstead)	27–1428	L14290
Betula nigra L.*	Betulaceae	Chase 255, NCU	Albert et al., 1992 (Qiu)	27-1428	L01889
Brassica campestris L.	Brassicaceae	none	Olmstead et al., 1992 (Nugent & Palmer)	1–1428	—
Byrsonima crassifolia (L.) Kunth.*	Malpighiaceae	Fairchild Trop. G. 81680, MICH	Albert et al., 1992 (Chase)	29–1428	L01892
Chrysosplenium iowense Rydb.	Saxifragaceae	Wendel *s.n.*, ISC	†(Soltis)	16–1428	—
Convolvulus tricolor L.*	Convolvulaceae	none	Olmstead et al., 1992	103–1428	L11683
Cornus florida L.	Cornaceae	Xiang 250, WS	Xiang et al., 1993	27–1428	L11215
Dendrophthora clavata (Benth.) Urban	Viscaceae	Nickrent 2182, SIU	†(Nickrent)	3–1428	L26069
Drimys winteri Forster & Forster	Winteraceae	Qiu 90016, NCU	Albert et al., 1992 (Qiu)	29–1428	L01905
Ephedra tweediana C. A. Mey.	Ephedraceae	none	Chase et al., 1993 (Chase)	30–1428	L12677
Eubrachion ambiguum (Hooker & Arnott) Engl.	Eremolepidaceae	Nickrent 2699, SIU	†(Nickrent)	12–1428	L26071
Euonymus alatus (Thunb.) Siebold	Celastraceae	Chase 137, NCU	Chase et al., 1993 (Chase)	27–1428	L13184
Francoa sonchifolia Cav.	Saxifragaceae	Soltis & Soltis 2479, WS	Soltis et al., 1990	31–1407	L11184
Gaiadendron punctatum (R. & P.) G. Don	Loranthaceae	Nickrent 2729, SIU	†(Nickrent)	4–1373	L26072
Ginalloa arnottiana Korth	Viscaceae	Nickrent 2982, SIU	†(Nickrent)	6–1428	L26070
Gnetum gnemon L.	Gnetaceae	Chase 208, NCU	Chase et al., 1993 (Chase)	26–1351	L12680
Gossypium robinsonii F. Muell.	Malvaceae	Wendel *s.n.*, ISC	Chase et al., 1993 (Chase)	27–1428	L13186
Hedera helix L.	Araliaceae	Jansen *s.n.*, MICH	Olmstead et al., 1993	27–1428	L01924
Heuchera micrantha Dougl. ex Lindl.	Saxifragaceae	Soltis & Soltis 1949, WS	Soltis et al., 1990	31–1407	L01925
Houttuynia cordata Thunb.	Saururaceae	Les *s.n.*, CONN	Chase et al., 1993 (Les)	29–1183	L08762
Hydrangea macrophylla (Thunb.) Ser.	Hydrangeaceae	Morgan 2150, WS	Morgan & Soltis, 1993	31–1407	L11187
Hydrocotyle sp.	Apiaceae	unknown	Chase (pers. comm.)	16–1407	—
Impatiens capensis Meerb.	Balsaminaceae	Chase 114, NCU	Chase et al., 1993 (Chase)	29–1428	—
Korthalsella lindsayi (D. Oliv.) Engl.	Viscaceae	Nickrent 2740, SIU	†(Nickrent)	13–1428	L26073
Lambertia inermis R. Br.*	Proteaceae	Nat. Trop. BG, Hawaii	Morgan & Soltis, 1993	31–1407	L11190
Lepuropetalon spathulatum (Muhl.) Elliott	Saxifragaceae	Thomas *s.n.*, NLU	Morgan & Soltis, 1993	31–1407	L11192
Lycopersicon esculentum Mill.	Solanaceae	none	Olmstead et al., 1993	1–1428	L14403

TABLE 2. Continued.

Species	Family	Voucher	Reference (source)	Sequence	GenBank number
Mahonia bealei (Fortune) Carr.*	Berberidaceae	Qiu 74, NCU	Qiu et al., 1993	29–1428	L12657
Morus alba L.	Moraceae	none	Albert et al., 1992 (Golenberg)	1–1407	L01933
Misodendron brachystachyum D.C.	Misodendraceae	Nickrent 2829, SIU	†(Nickrent)	5–1428	L26074
Notothixos subaureus Oliver	Viscaceae	Nickrent 2790, SIU	†(Nickrent)	15–1428	L26075
Nymphaea odorata Aiton	Nymphaeaceae	Les *s.n.*, CONN	Les et al., 1991	29–1183§	M77034
Nyssa ogeche Marsh	Nyssaceae	U.S. Natl. Arb. *s.n.*	Xiang et al., 1993	31–1407	L11228
Opilia amentacea Roxb.	Opiliaceae	Nickrent 2816, SIU	†(Nickrent)	3–1428	L26076
Oryza sativa L.	Poaceae	unknown	Nishizawa & Hirai, 1987	1–1428	D00207
Osyris lanceolata Hochst. & Steud.	Santalaceae	Nickrent 2731, SIU	Morgan & Soltis, 1993	31–1407	L11196
Pachysandra procumbens Michx.*	Buxaceae	Chase 207, NCU	Chase et al., 1993 (Chase)	58–1360	—
Paeonia tenuifolia L.	Paeoniaceae	Kron 2115, NCU	Chase et al., 1993 (Kron)	29–1428	L13687
Peperomia sp.	Piperaceae	Qiu 91047, NCU	Qiu et al., 1993	27–1428	L12661
Phoradendron serotinum (Raf.) M. C. Johnst.	Viscaceae	Nickrent 2077, ILL	Morgan & Soltis, 1993	31–1407	L11199
Pisum sativum L.*	Fabaceae	unknown	Zurawski et al., 1986	1–1428	—
Pittosporum japonicum Hort. ex C. Presl.	Pittosporaceae	Rieseberg *s.n.*, RSA	Morgan & Soltis, 1993	31–1407	L11202
Polemonium reptans L.*	Polemoniaceae	none	Olmstead et al., 1992	103–1428	L11687
Prunus virginiana L.	Rosaceae	Soltis & Soltis 2475, WS	Albert et al., 1992 (Morgan & Soltis)	31–1407	L01947
Pyrola rotundifolia L.*	Pyrolaceae	Kron 1906, NCU	Kron & Chase, 1993	29–1428	L12622
Ranunculus trichophyllus Chaix	Ranunculaceae	Les *s.n.*, CONN	Chase et al., 1993 (Les)	29–1183	L08766
Reinwardtia indica Dumort.*	Linaceae	Chase 230, NCU	Chase et al., 1993 (Chase)	27–1366	L13188
Ribes aureum Pursh	Grossulariaceae	Soltis & Soltis 2220, WS	Morgan & Soltis, 1993	31–1407	L11204
Santalum album L.	Santalaceae	Nickrent 2734, SIU	†(Nickrent)	7–1428	L26077
Saruma henryi Oliver	Aristolochiaceae	Qiu 91018, NCU	Qiu et al., 1993	29–1428	L12664
Schoepfia schreberi Gmelin	Olacaceae	Nickrent 2599, SIU	Morgan & Soltis, 1993	30–1428	L11205
Sparganium americanum Engl.	Sparganiaceae	Chase 257, NCU	Duvall et al., 1993 (Duvall)	20–1351	M91633
Spinacia oleracea L.	Chenopodiaceae	unknown	Zurawski et al., 1981	1–1428	J01443
Tropaeolum majus L.	Tropaeolaceae	Chase 113, NCU	Price & Palmer 1993	27–1428	L14706
Viscum album L.	Viscaceae	Nickrent 2253, SIU	†(Nickrent)	13–1428	L26078
Zea mays L.	Poaceae	unknown	Gaut et al., 1992	1–1428	Z11973
Zamia inermis Vovides, Reese & Vásquez-Torres	Zamiaceae	Schutzman *s.n.*, FLAS	Chase et al., 1993 (Hills & Chase)	31–1428	L12683

† Sequence first reported here.
* Different genera used as placeholders for family for phylogenetic analyses.
§ The remaining sequence (1184–1428) was not available through Genbank, but was determined by Qiu (Chase et al., 1993) and sent by Mark Chase.

this study were previously published by workers other than the authors (Table 1). With the exception of six sequences that were obtained via direct sequencing of rRNA using reverse transcriptase (Nickrent & Franchina, 1990), all of the remaining rDNA sequences were derived from PCR products.

AMPLIFICATION AND SEQUENCING

The genomic DNAs used for amplification and sequencing of 18S rDNA and *rbc*L were extracted using a modification of the hot CTAB method (Doyle & Doyle, 1987; Nickrent, 1994). Plant samples were derived from either fresh, silica gel-dried, or herbarium material. The PCR protocols employed, as well as the oligonucleotide primers used for the amplification and subsequent sequencing of rDNA, are provided in Nickrent (1994) and Nickrent & Starr (1994). The general PCR strategy and amplification primers used for *rbc*L are provided in Morgan & Soltis (1993). For both *rbc*L and 18S rDNA, the first author employed direct sequencing of double-stranded PCR products. These products were prepared for sequencing by gel purification whereby the PCR bands are bound to DEAE membranes, eluted, and precipitated in ethanol. DNA so prepared is denatured at 100°C and snap-cooled for primer annealing. In contrast, the second author used each of the two PCR primers individually to generate single-stranded DNA from the double-stranded PCR products. Single-stranded 18S and *rbc*L DNAs were then purified by precipitation with 20% PEG/2.5 M NaCl, as described by Morgan & Soltis (1993). In all instances, the chain-termination method of sequencing was employed (Sanger et al., 1977) using [^{35}S] dATP and the Sequenase® version 2.0 kit. DNA fragments were separated in 6% acrylamide gels; gels were subsequently fixed and used to expose film using standard techniques. Compressions and other structure-related artifacts were resolved either through the use of alternative nucleotides (deaza-dGTP, dITP) or by sequencing the same region on the complementary strand.

SEQUENCE FEATURES AND MULTIPLE ALIGNMENTS

All alignments were initially conducted on a SUN Spark Station running the Genetic Data Environment (GDE, version 2.2; Smith, 1992). These alignments were downloaded to a Macintosh computer and directly imported into MacClade (version 3.01; Maddison & Maddison, 1992). For each molecule, the chart features of MacClade were used to examine patterns of variability and conservation, transition/transversion bias, and (for *rbc*L) the number of changes per codon position. This program was also used to determine the number of phylogenetically informative sites for each alignment. MacClade files were saved in Nexus format and then imported into PAUP (version 3.1; Swofford, 1993) for parsimony analyses. Files containing the complete alignments of both molecules are available from both authors by sending a formatted 3.5-inch diskette.

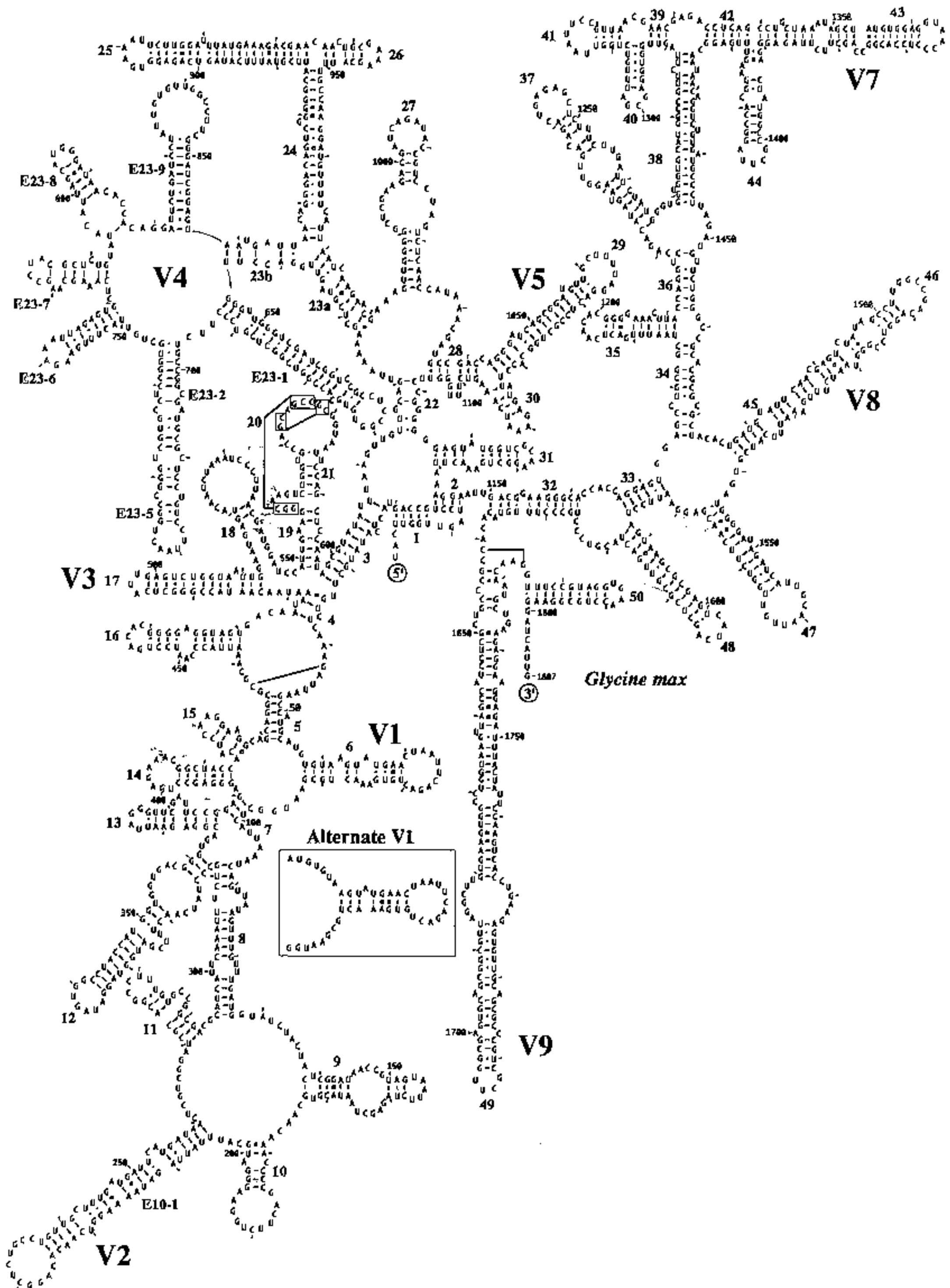
The alignment of *rbc*L sequences is straightforward and can be accomplished easily by sight because very few length mutations exist. In contrast, rDNA sequence alignment was performed as an iterative process that simultaneously dealt with phylogenetic relationships, compensatory mutations, and higher-order structure. The higher-order rRNA structure of *Glycine max* (Fig. 1), like the recently proposed structure for *Rafflesia keithii* (Nickrent & Starr, 1994), is similar to the one given for maize by Gutell et al. (1985) but includes the structural changes proposed for yeast (Gutell, 1993) and eukaryotes (Neefs et al., 1993). This structure differs somewhat from that proposed by Senecoff & Meagher (1992), which was based largely on a mammalian model. The soybean structure given here was used as a reference for examining structural variation in the other plant species examined and as a guide during the alignment process.

Until recently, the secondary structure of rRNA made direct sequencing of this molecule very difficult and aroused some criticism over the utility of rRNA for phylogenetic reconstruction in plants. With the advent of PCR, the amplification and sequencing of rDNA is no more difficult than for other genes (Nickrent & Starr, 1994). Furthermore, secondary structure provides much-needed corroboratory information regarding base pairing and compensatory base changes that is essential in producing alignments that reflect proper base homology.

The 18S rDNA sequences obtained were approximately 1770 base pairs in length. Published

→

FIGURE 1. Proposed secondary structural model for the small-subunit (18S) ribosomal RNA of *Glycine max*. The primary sequence of soybean was determined by Eckenrode et al. (1985). This structural model follows the general models proposed by Gutell et al. (1985) for *Zea*, Gutell (1993) for *Saccharomyces*, and Neefs et al. (1993) for



eukaryotes in general. Helix numbering corresponds to Neefs et al. (1993). The structure for helix 6 (V1 region) follows Gutell (1993), with the alternative interpretation according to Neefs et al. (1993). The structure for the V4 region follows Nickrent & Sargent (1991). For an alternative model of the V4 involving a pseudoknot between helices E23-8 and E23-9, see Neefs et al. (1993). The V6 region is absent in eukaryotes. Tertiary interactions are indicated by thick lines.

complete 18S rDNA sequences of higher plants vary in length from 1800 to 1813 base pairs (mean of 1807 base pairs); thus, in this study, 97.9% of the total length of the molecule was obtained for most taxa. Owing to alignment spacers ("-"), the total length of the matrix was 1853. Insertion/deletion events (indels) were treated as missing data. Certain regions of 18S rDNA are variable in primary sequence and length, such as the termini of helices E10-1, E23-1, and 49 (Fig. 1). These regions confound unambiguous alignment; hence positions 227–239 and 676–685 on the alignment (equivalent to sites 224–232 and 664–673 on the *Glycine* molecule) were eliminated from analysis as suggested by Swofford & Olsen (1990). Those base pairs corresponding to the 25e forward 18S rDNA PCR primer (positions 1–20) were removed from the analysis. Similarly, those base pairs corresponding to the 1769 reverse PCR primer (sites 1810–1853 on the alignment 1764–1807 on *Glycine*) were eliminated. With the exception of the excluded base pairs, alignment of the 18S rDNA sequences was straightforward because most length mutations involve single base insertions or deletions.

The total length of the *rbc*L data matrix was 1431. However, the first 30 base pairs were not used, because, after amplification, this portion of the gene is identical to the Z1 forward amplification primer. Sequences of *Nymphaea*, *Houttuynia*, and *Ranunculus* were incomplete (see Table 2); "?" was used to indicate missing sites.

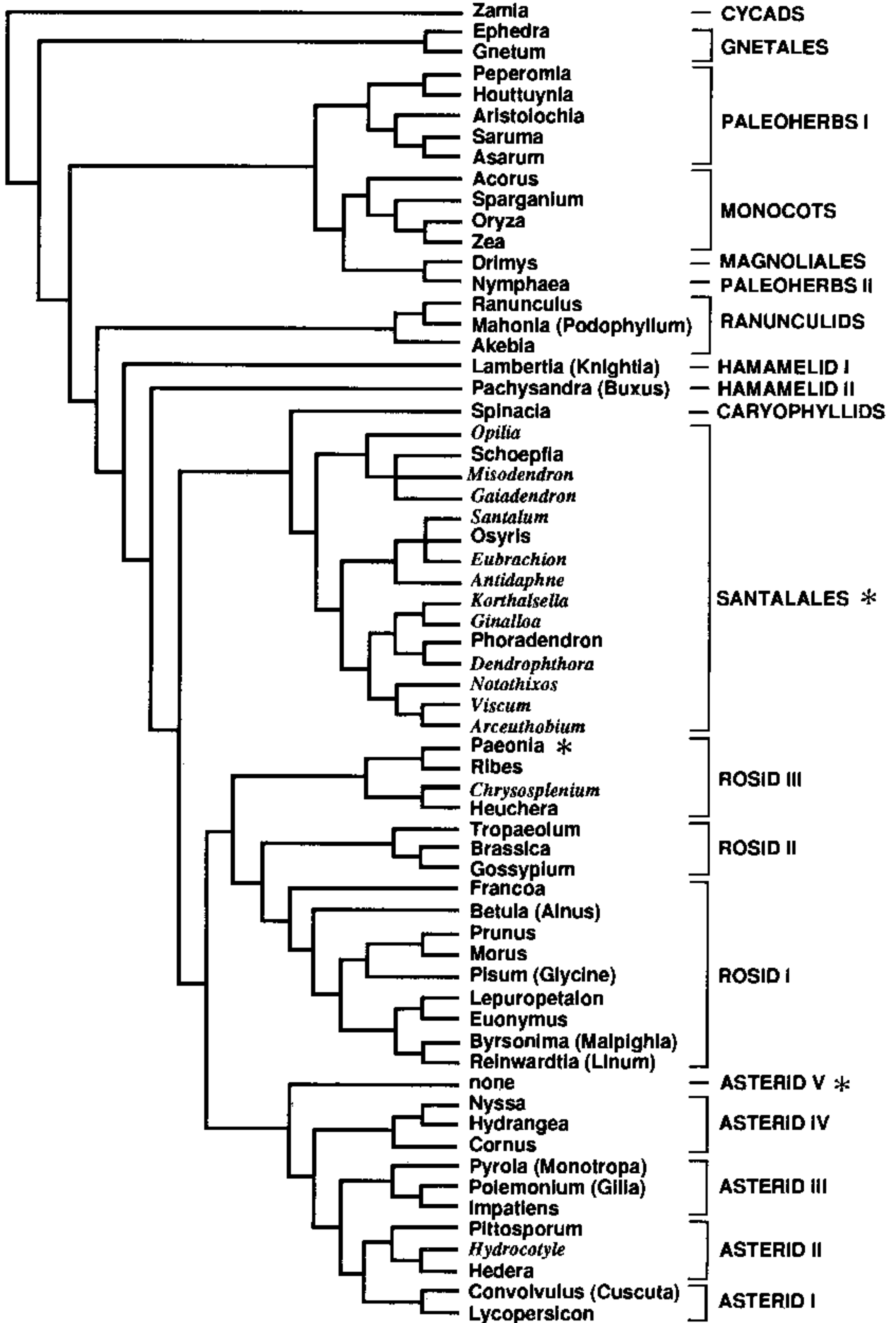
PHYLOGENETIC ANALYSES

Minimum-length Fitch parsimony trees were constructed using PAUP version 3.1.1 (Swofford, 1993) with MULPARS and TBR branch swapping. Given the number of taxa (62), only heuristic search strategies could be employed. Both data sets were analyzed giving all changes equal weight. Trials using the character-state transformation weighting model of Albert et al. (1993) for the *rbc*L matrix and a transformation matrix encompassing a 10:1 bias favoring transitions over transversions for the 18S rDNA data gave similar results as trials with equal weighting. To determine whether multiple, equally parsimonious "islands" of most parsimonious trees exist (Maddison, 1991), 100 replicate searches with random taxon addition were conducted. To obtain estimates of reliability for monophyletic groups, bootstrap (Felsenstein, 1985) analyses (100 replicates) were conducted. For both data sets, the bootstrap analysis was performed using simple taxon addition, TBR branch swapping, ACCTRAN character-state optimization, and unweighted characters.

The phylogenetic trees derived from the global *rbc*L analyses of Chase et al. (1993) were used to construct a "reference tree" for the subset of taxa used in this study (Fig. 2). This reference tree was also constructed to assess the effect of taxon sampling and density on the stability and composition of various clades as determined by the present analyses. The Search II strategy employed by Chase et al. (1993) resulted in 3900 shortest trees, one of which was chosen at random and depicted in their figure 2B. Search II was preferred by the authors (over their Search I) because it included a greater diversity of taxa, was able to save more trees of shortest length, and did not use relative weighting of character-state transformations. In figure 2B of Chase et al. (1993), the angiosperms were divided into 19 major groups, the composition of which varied from single genera to groups of many families. Owing to lack of an 18S rDNA sequence, the reference tree constructed here does not include four of the major *rbc*L clades: *Ceratophyllum*, *Gunnera*, Laurales, and Asterid V. The Asterid V clade was not in the Search II topology of Chase et al. (1993), but Asterid V included Santalales in the tree resulting from Search I. It is well-known that sampling affects tree topologies (Felsenstein, 1985); the reference tree of Figure

→

FIGURE 2. The "reference tree" constructed from the topology found in tree 2B in the global *rbc*L analysis that included 499 taxa (Chase et al., 1993). This tree represents a null hypothesis that assumes *rbc*L is insensitive to taxon sampling, i.e., all topologies using fewer taxa are fully concordant with the global topology. Instances where two generic names are given represent cases where different generic representatives of a family were used (*rbc*L taxon first followed by 18S rDNA taxon in parentheses). Those taxa in italics were not included in the study by Chase et al. (1993); their placement on the tree is derived from the analysis reported here (e.g., Santalales) or based upon traditional familial classifications (e.g., *Hydrocotyle*, Apiaceae). The names of the major angiosperm clades (right side) correspond to Chase et al. (1993). Taxa marked with an asterisk (*, Santalales and *Paeonia*) were located on the Asterid V clade in Search I of Chase et al. (1993).


Zamia
Ephedra
Gnetum
Peperomia
Houttuynia
Aristolochia
Saruma
Asarum
Acorus
Sparganium
Oryza
Zea
Drimys
Nymphaea
Ranunculus
Mahonia (Podophyllum)
Akebia
Lambertia (Knightia)
Pachysandra (Buxus)
Spinacia
Opilia
Schoepfia
Misodendron
Gaiadendron
Santalum
Osyris
Eubrachion
Antidaphne
Korthalsella
Ginalloa
Phoradendron
Dendrophthora
Notothixos
Viscum
Arceuthobium
Paeonia *
Ribes
Chrysosplenium
Heuchera
Tropaeolum
Brassica
Gossypium
Francoa
Betula (Alnus)
Prunus
Morus
Pisum (Glycine)
Lepuropetalon
Euonymus
Byrsonima (Malpighia)
Reinwardtia (Linum)
none
Nyssa
Hydrangea
Cornus
Pyrola (Monotropa)
Polemonium (Gilia)
Impatiens
Pittosporum
Hydrocotyle
Hedera
Convolvulus (Cuscuta)
Lycopersicon
CYCADS
GNETALES
PALEOHERBS I
MONOCOTS
MAGNOLIALES
PALEOHERBS II
RANUNCULIDS
HAMAMELID I
HAMAMELID II
CARYOPHYLLIDS
SANTALALES *
ROSID III
ROSID II
ROSID I
ASTERID V *
ASTERID IV
ASTERID III
ASTERID II
ASTERID I

2 should therefore be interpreted as a null hypothesis that assumes *rbc*L is insensitive to taxon inclusion and that the topology of a restricted analysis is congruent with that of a global analysis.

RESULTS

GENERAL FEATURES OF *rbc*L AND 18S rDNA

The length of *rbc*L is highly conserved in higher plants with few insertion/deletion events reported (Chase et al., 1993). Positions 1426–1428 form the most common stop codon, although longer reading frames up to 1458 bp have been reported in Asteraceae (Kim et al., 1992). Among the taxa analyzed herein, a single insertion of three bases occurs in *Zea* beginning at position 1404, whereas all other full length *rbc*L sequences used herein are of length 1428. For *rbc*L, 482 (33.6%) of the 1431 base positions are potentially phylogenetically informative. The length of complete 18S rDNA also varies: 1800 bp (*Lycopersicon*), 1804 bp (*Brassica*), 1807 bp (*Glycine*), 1809 bp (*Zea*), 1812 bp (*Oryza*), and 1813 bp (*Zamia*). Of the 1853 positions for the 18S rDNA alignment, 341 (18.4%) are potentially phylogenetically informative.

When one compares sequences of two distantly related taxa, for example, *Zamia* and *Pisum* (Fabaceae) for *rbc*L and *Zamia* and *Glycine* (Fabaceae) for 18S rDNA, the *rbc*L sequence comparison yields 191 mutational differences (13.3% of the 1431 sites), whereas comparison of 18S sequences yields 138 differences (7.6% of the ca. 1810 sites). Similar comparisons using angiosperms show that *rbc*L is generally about three times more variable than 18S rDNA. For example, comparison of *Pisum* and *Spinacia rbc*L sequences demonstrates that 139 of the 1428 sites (9.7%) are different. In contrast, a similar comparison of *Glycine* (Fabaceae) and *Spinacia* 18S rDNA sequences indicates that only 62 of the ca. 1808 (3.4%) sites are different. Not only is the rate of evolution of *rbc*L considerably higher than that of the 18S gene (about 3 times faster), but even when the greater length of the 18S gene is taken into consideration, *rbc*L still exhibits approximately 1.4 times as many variable sites as 18S rDNA. The distribution of variable sites for the two molecules is also quite different. When the number of steps from one of the equally most parsimonious *rbc*L and 18S rDNA cladograms (e.g., Figs. 6 and 8, respectively) is plotted against site, the different variability patterns are graphically illustrated (Figs. 3 and 4). For *rbc*L, sites in general are more variable, and, although certain regions clearly are more variable than others, this variability appears more evenly distributed over the entire length of the molecule (Fig. 3) than for 18S (Fig. 4). That is, 18S rDNA shows highly variable regions interspersed with regions of extreme conservation (Fig. 4). Significantly, the variable domains indicated on the secondary structure of *Glycine* (V1–V9, Fig. 1) can be readily identified in Figure 4. The secondary structural study conducted by Senecoff & Meagher (1992) used dimethyl sulfate to modify (and thereby identify) adenine and cytosine residues of single-stranded portions of the soybean 18S rRNA molecule. Their data largely confirm the higher-order structure shown in Figure 1, especially for variable regions 1 and 4.

Pairwise 18S rDNA sequence comparisons within the flowering plants examined here indicated that most angiosperms differ from the above noted *Glycine* sequence at only 1–5% of the sites. Higher than average rates (numbers) of nucleotide substitution in 18S rDNA can be seen, however, in certain parasitic plants such as members of Viscaceae and Cuscutaceae (included in the present study) and Balanophoraceae, Hydnoraceae, and Rafflesiaceae (Nickrent & Starr, 1994). Representatives of the latter three families were not included herein because they apparently lack an *rbc*L gene (Nickrent & dePamphilis, unpublished data). The causes of such elevated rates of 18S sequence evolution are currently under investigation by the first author.

For both data sets, transitions outnumber transversions by approximately a factor of two. For the *rbc*L data set, there were 1520 unambiguous transitions and 862 unambiguous transversions; for 18S rDNA, there were 1099 and 574 unambiguous transitions and transversions, respectively. The specific types of mutational events for the two molecules are also very similar, differing mainly in the frequency of the A to G transition (14.6% of the total changes for *rbc*L, 5.3% of the total for 18S rDNA). Steps calculated over the *rbc*L tree by codon position demonstrate that most changes occur, as expected, in the third position followed by first and second position.

PHYLOGENETIC ANALYSES

rbcL. The heuristic search of the *rbc*L data matrix yielded 12 most parsimonious trees, all in one island, of length 3090 with a consistency index excluding uninformative substitutions (C.I.–) of 0.284 and retention index (R.I.) of 0.467. The strict consensus tree is illustrated in Figure 5. The main differences among the 12 most parsimonious trees

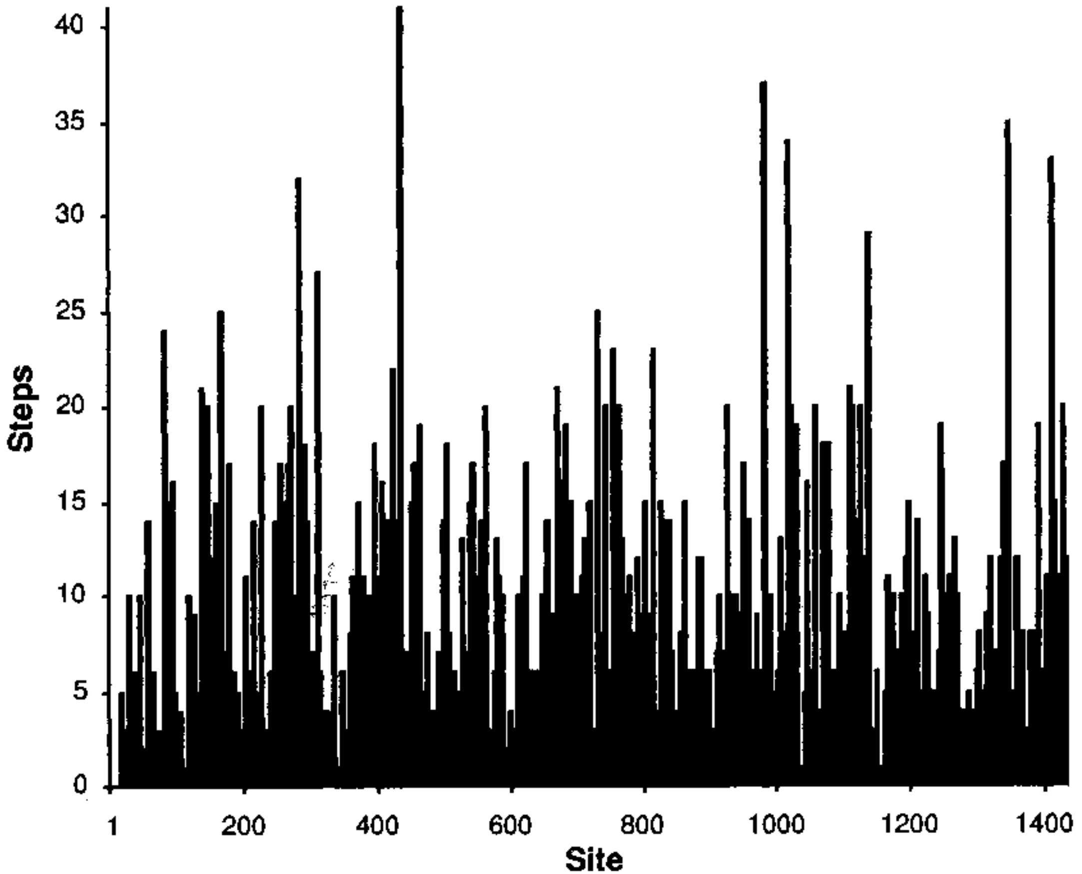


FIGURE 3. Variability histogram for *rbc*L. The number of steps (y axis) were determined from *rbc*L tree number 3 (of 12 equally parsimonious trees) as shown in Figure 6 with the interval widths (x axis) set to four base pairs. Variability is distributed relatively evenly over the 1431 sites of the molecule.

were the relationships among members of Santalales and in the relative position of *Lambertia* to nonpaleoherb dicots. The strict consensus tree shares a number of features with the *rbc*L reference tree (Fig. 2). Using *Zamia* as the outgroup, the two representatives of Gnetales (*Ephedra* and *Gnetum*) form a monophyletic group strongly supported by 66 synapomorphies that is sister to the angiosperms (Fig. 6). The angiosperms form a monophyletic group united by 39 base substitutions (bootstrap value of 93%). Within the angiosperms, the monocots examined (*Zea*, *Oryza*, and *Sparganium*), with the exception of *Acorus*, are the sister group to all other angiosperms. This relationship differs from the reference tree where the monocots, including *Acorus*, are monophyletic and are sister to the Magnoliales/Paleoherbs II group. In the present analysis, the Paleoherbs I group (*Houttuynia*, *Peperomia*, *Asarum*, *Saruma*, and *Aristolochia*) is disrupted by the inclusion of *Acorus* and *Drimys* (Magnoliales). The Ranunculids (*Akebia*, *Mahonia*, and *Ranunculus*) have the same composition and general topology as seen in the reference tree. This latter clade is part of a trichotomy in the strict consensus tree of Figure 5 that also comprises *Lambertia* (the single representative of the Hamamelid I group) and the remaining dicots. The present analysis of *rbc*L sequences does not include the closest relatives of *Lambertia* (i.e., *Sabia*, *Nelumbo*, *Platanus*) as determined in the Chase et al. analysis. *Pachysandra*, representing Hamamelid II, occupies a similar position on the strict consensus and reference trees. Several of the remaining clades have taxon compositions identical to those of the reference tree, although the topologies within these clades are not necessarily identical to those of the reference tree. These clades of identical composition include Asterid III (*Pyrola*, *Polemonium*, and *Impatiens*), Rosid II (*Brassica*, *Tropaeolum*,

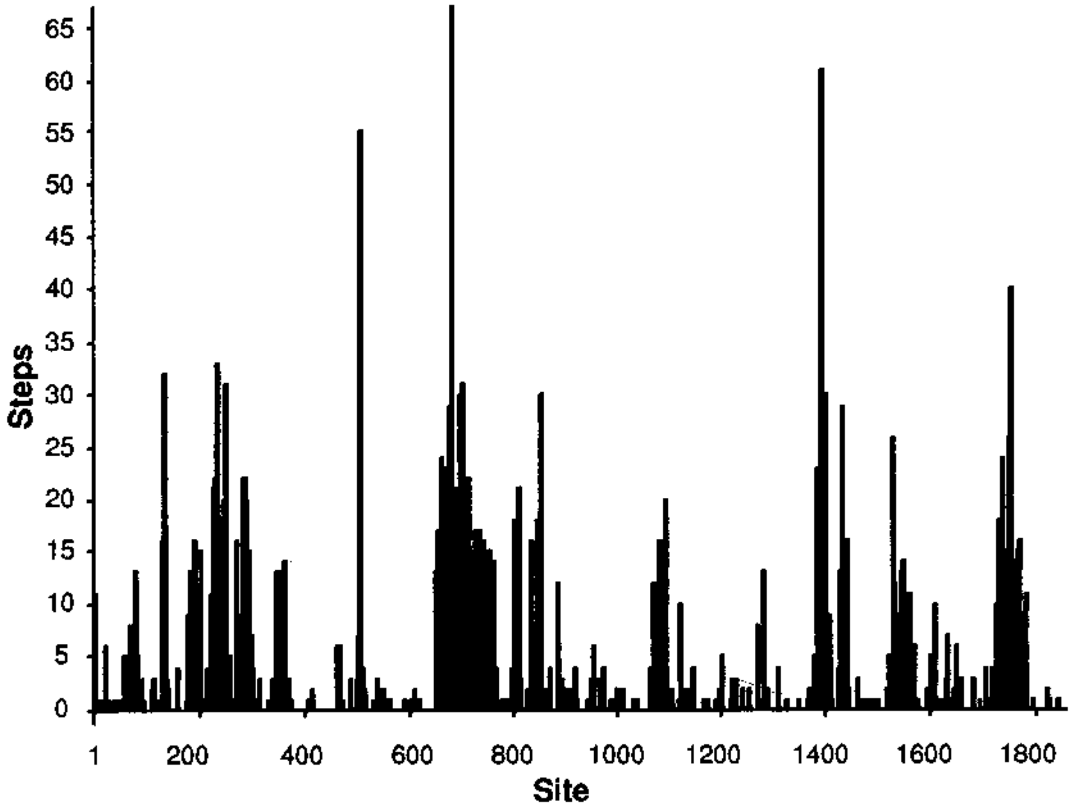


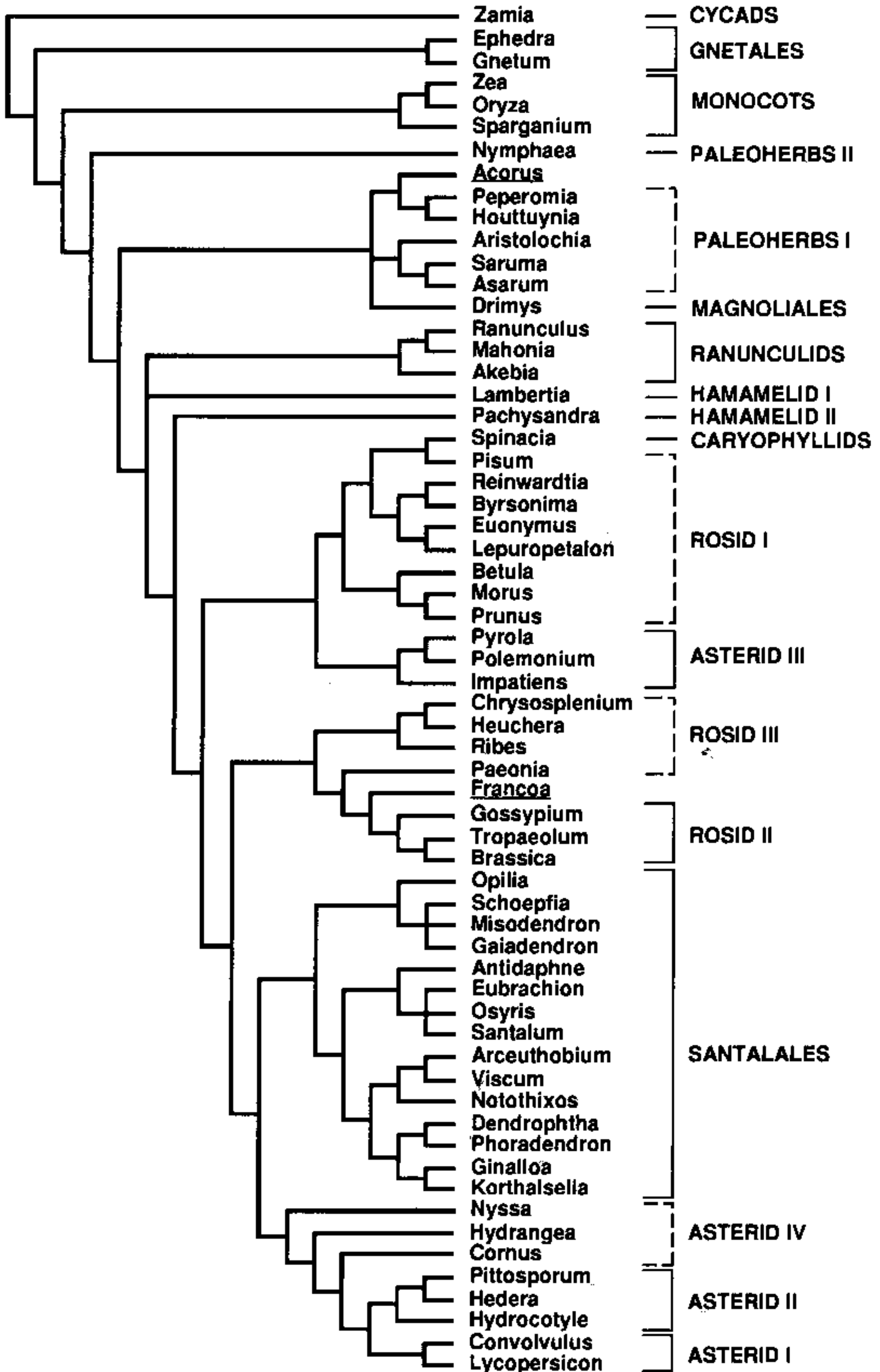
FIGURE 4. Variability histogram for 18S rDNA. The number of steps were determined from 18S rDNA tree number 5 (of 26 equally parsimonious trees) as shown in Figure 8 with the interval widths set to four base pairs. Regions of variability, generally concentrated in the variable domains (Fig. 1), are interspersed with extremely conserved sites over the 1853 total sites for the molecule.

and *Gossypium*), Santalales, Asterid II (*Pittosporum*, *Hedera*, and *Hydrocotyle*), and Asterid I (*Lycoperiscon* and *Convolvulus*). The members of Asterid IV (*Hydrangea*, *Cornus*, and *Nyssa*) do not form a monophyletic group herein, but do appear near each other at the base of the Asterid I and II groups. The sole representative of the Caryophyllids, *Spinacia*, was positioned within the Rosid I clade as opposed to sister to Santalales on the reference tree. The position of Caryophyllids differed also in the two searches conducted by Chase et al. (1993).

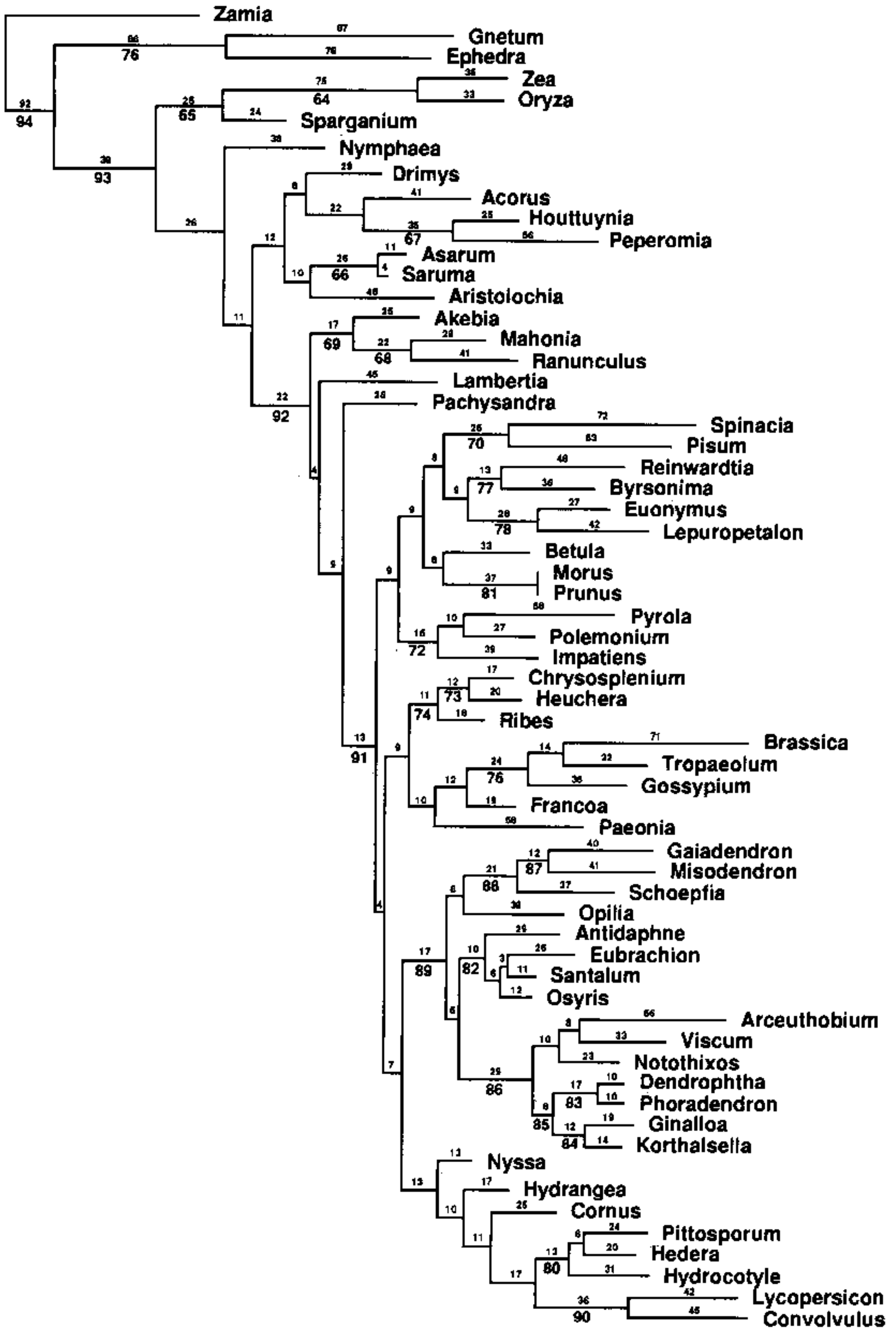
In our 62-taxa *rbc*L analysis, the Rosid I and Rosid III clades are nearly identical in composition to those of Chase et al. (1993). In our strict consensus tree, all taxa of the Rosid I clade, with the exception of *Francoa* and the addition of *Spinacia*, form a monophyletic group (compare Figs. 2 and 5). However, the omission of *Francoa* from this clade again likely reflects taxon density. The closest relatives of *Francoa* based on the Chase et al. (1993) analysis (*Greyia*, *Viviana*, *Wendtia*) were not included herein. With the exception of *Paeonia*, Rosid III also appears as a monophyletic group in our analysis of *rbc*L sequences. One should regard this difference with great caution given that the position of *Paeonia* shifts dramatically between the two searches of Chase et al. (1993). The San-

→

FIGURE 5. The strict consensus tree of 12 equally most parsimonious cladograms derived from a heuristic search of the 62-taxon *rbc*L matrix; tree length = 3104, C.I. minus uninformative sites = 0.283, R.I. = 0.463. Groups whose compositions are identical to those of the reference tree (Fig. 2) are indicated by solid braces. Groups that appear para- and polyphyletic (relative to the reference tree) are indicated by dashed braces. The positions of underlined taxa (*Acorus* and *Francoa*) differ significantly from those expected from the reference tree (monocots and Rosid I, respectively).



Zamia
Ephedra
Gnetum
Zea
Oryza
Sparganium
Nymphaea
Acorus
Peperomia
Houttuynia
Aristolochia
Saruma
Asarum
Drimys
Ranunculus
Mahonia
Akebia
Lambertia
Pachysandra
Spinacia
Pisum
Reinwardtia
Byrsonima
Euonymus
Lepuropetalon
Betula
Morus
Prunus
Pyrola
Polemonium
Impatiens
Chrysosplenium
Heuchera
Ribes
Paeonia
Francoa
Gossypium
Tropaeolum
Brassica
Opilia
Schoepfia
Misodendron
Gaiadendron
Antidaphne
Eubrachion
Osyris
Santalum
Arceuthobium
Viscum
Notothixos
Dendrophtha
Phoradendron
Ginalloa
Korthalsella
Nyssa
Hydrangea
Cornus
Pittosporum
Hedera
Hydrocotyle
Convolvulus
Lycopersicon
CYCADS
GNETALES
MONOCOTS
PALEOHERBS II
PALEOHERBS I
MAGNOLIALES
RANUNCULIDS
HAMAMELID I
HAMAMELID II
CARYOPHYLLIDS
ROSID I
ASTERID III
ROSID III
ROSID II
SANTALALES
ASTERID IV
ASTERID II
ASTERID I


Zamia
Gnetum
Ephedra
Zea
Oryza
Sparganium
Nymphaea
Drimys
Acorus
Houttuynia
Peperomia
Asarum
Saruma
Aristolochia
Akebia
Mahonia
Ranunculus
Lambertia
Pachysandra
Spinacia
Pisum
Reinwardtia
Byrsonima
Euonymus
Lepuropetalon
Betula
Morus
Prunus
Pyrola
Polemonium
Impatiens
Chrysosplenium
Heuchera
Ribes
Brassica
Tropaeolum
Gossypium
Francoa
Paeonia
Gaiadendron
Misodendron
Schoepfia
Opilia
Antidaphne
Eubrachion
Santalum
Osyris
Arceuthobium
Viscum
Notothixos
Dendrophtha
Phoradendron
Ginalloa
Korthalsella
Nyssa
Hydrangea
Cornus
Pittosporum
Hedera
Hydrocotyle
Lycopersicon
Convolvulus

talales form a monophyletic group in the present analysis, and their position here as sister to the Asterid clades is similar to the results of Search I of Chase et al., where the Santalales (represented in Chase et al. only by *Phoradendron*, *Schoepfia*, and *Osyris*), along with *Gunnera*, appear as the sister to all other asterids. In contrast, in Search II of Chase et al., the Santalales appear as sister to a clade containing the Caryophyllids. In the detailed analysis of *rbc*L sequences of asterids conducted by Olmstead et al. (1993), Santalales are not a component of Asteridae s.l. and they are likely members of a broadly defined rosid clade.

Bootstrap values and branch lengths of the 62-taxa *rbc*L tree (Fig. 6) suggest the presence of several strongly supported major clades within the angiosperms. The monocots, Paleoherbs, and Magnoliales (represented by *Drimys*) appear as the sister to the remainder of the angiosperms, which are supported as a monophyletic group by a high bootstrap value (92%). This large clade comprises most of the taxa included in the corresponding clade of Figure 2 and represents the "eudicots," which have triaperturate or triaperturate-derived pollen (Donoghue & Doyle, 1989; Chase et al., 1993; Qiu et al., 1993). Within this large eudicot clade, the Ranunculids, *Lambertia*, and *Pachysandra* appear as the sister to another large, strongly supported clade (bootstrap value of 91%). There are, however, few strongly supported subclades within this large clade. Subclades that received moderate to strong support (bootstraps of 70–80%) include the Asterid I, II, III, and IV, Rosid II, Rosid III minus *Paeonia*, and Santalales. Within the Santalales, the monophyly of the mistletoe family, Viscaceae, is supported by a bootstrap value of 86%.

18S-rDNA. Cladistic analysis of the 18S rDNA matrix yielded 26 equally parsimonious trees of length 2021, all in one island. Each of these trees had a C.I.– of 0.301 and an R.I. of 0.440. The strict consensus tree (Fig. 7) reveals that the Gnetales again appear as the sister to the angiosperms and that the angiosperms form a well-supported monophyletic group (bootstrap value of 100%, Fig. 8). Most clades in the 18S consensus tree are derived from a large polytomy, whereas the *rbc*L consensus tree (Fig. 5) displays considerable resolution. In the 18S analysis, the monocots *Zea*, *Oryza*, and *Sparganium* (minus *Acorus*) form a monophyletic group (bootstrap value of 91%) as does each of the following: Ranunculids, Rosid III, Santalales, Asterid I, and Asterid II.

A bootstrap analysis of the 18S rDNA data set (Fig. 8) indicates that *Acorus* and then *Nymphaea* are the sisters to a large clade containing the remaining angiosperms; however, these relationships received only weak support (bootstrap values less than 50%). A group of paleoherbs (*Asarum*, *Saruma*, *Aristolochia*, *Houttuynia*, and *Peperomia*) then appears as the sister to all remaining taxa. The large remaining clade corresponds, with one exception, to the "eudicots," as defined by Chase et al. (1993). The main discrepancy in the composition of eudicots between the *rbc*L tree and the 18S rDNA tree pertains to the relationships of *Drimys*: in contrast to the *rbc*L tree, the 18S tree places *Drimys* in the eudicot clade as sister to *Glycine*. The 18S eudicot clade is defined by only six base substitutions (Fig. 8) and is not present in the strict consensus tree (Fig. 7), whereas in the *rbc*L analysis, 22 base substitutions support this clade, and the bootstrap value is high (92%) (Fig. 6).

Despite the poorer resolution of the 18S than the *rbc*L tree, several subclades appear in both analyses. For example, the three genera of Aristolochiaceae (*Asarum*, *Saruma*, *Aristolochia*) form a monophyletic group, but the association with *Houttuynia* and *Peperomia*, the other subclade of the Paleoherb I group of Chase et al. (1993), is only weakly supported (bootstrap value less than 50%). The genera of Ranunculids (*Akebia*, *Mahonia/Podophyllum*, and *Ranunculus*) also form a monophyletic group in both trees, although the position of this clade is different.

Relationships within and among the several rosid clades show similarities in the 18S rDNA and *rbc*L trees, as well as several marked differences. 18S rDNA sequence data corroborate the results of *rbc*L sequence analysis in suggesting close relationships between some members of the Rosid I clade (Fig. 8, *Alnus* (representing Betulaceae), *Morus*, *Prunus*, *Francoa*, and *Malpighia* (representing Malpighiaceae)). In both the 18S and *rbc*L

←

FIGURE 6. One of the 12 equally most parsimonious phylograms (cladograms that show branch lengths) derived from the heuristic search of the 62-taxon *rbc*L matrix; tree length 3090, C.I. minus uninformative sites = 0.284, R.I. = 0.467. Numbers above the branches indicate branch lengths (i.e., number of steps or nucleotide substitutions). Numbers below the branches indicate the percentage of trees (from 100 bootstrap replications) that support that node. Branches without bootstrap percentages were found in less than 50% of the trees.

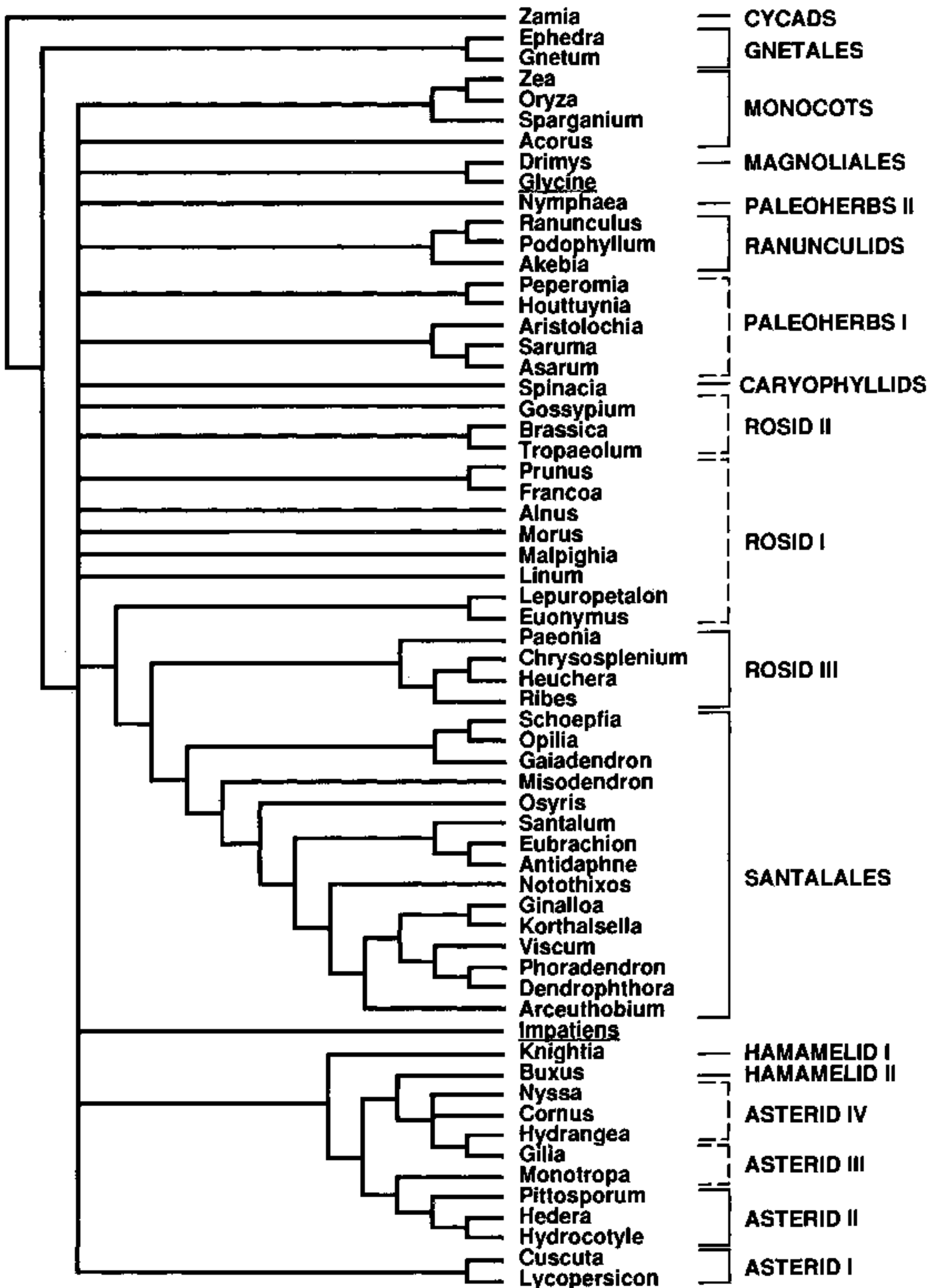


FIGURE 7. The strict consensus tree of 26 equally most parsimonious cladograms derived from a heuristic search of the 62-taxon 18S rDNA matrix; tree length = 2118, C.I. minus uninformative sites = 0.285, R.I. = 0.396. Groups whose compositions are identical to those of the reference tree (Fig. 2) are indicated by solid braces. Groups that appear para- and polyphyletic (relative to the reference tree) are indicated by dashed braces. The positions of underlined taxa (*Glycine* and *Impatiens*) differ significantly from those expected from the reference tree (Rosid I and Asterid III, respectively).

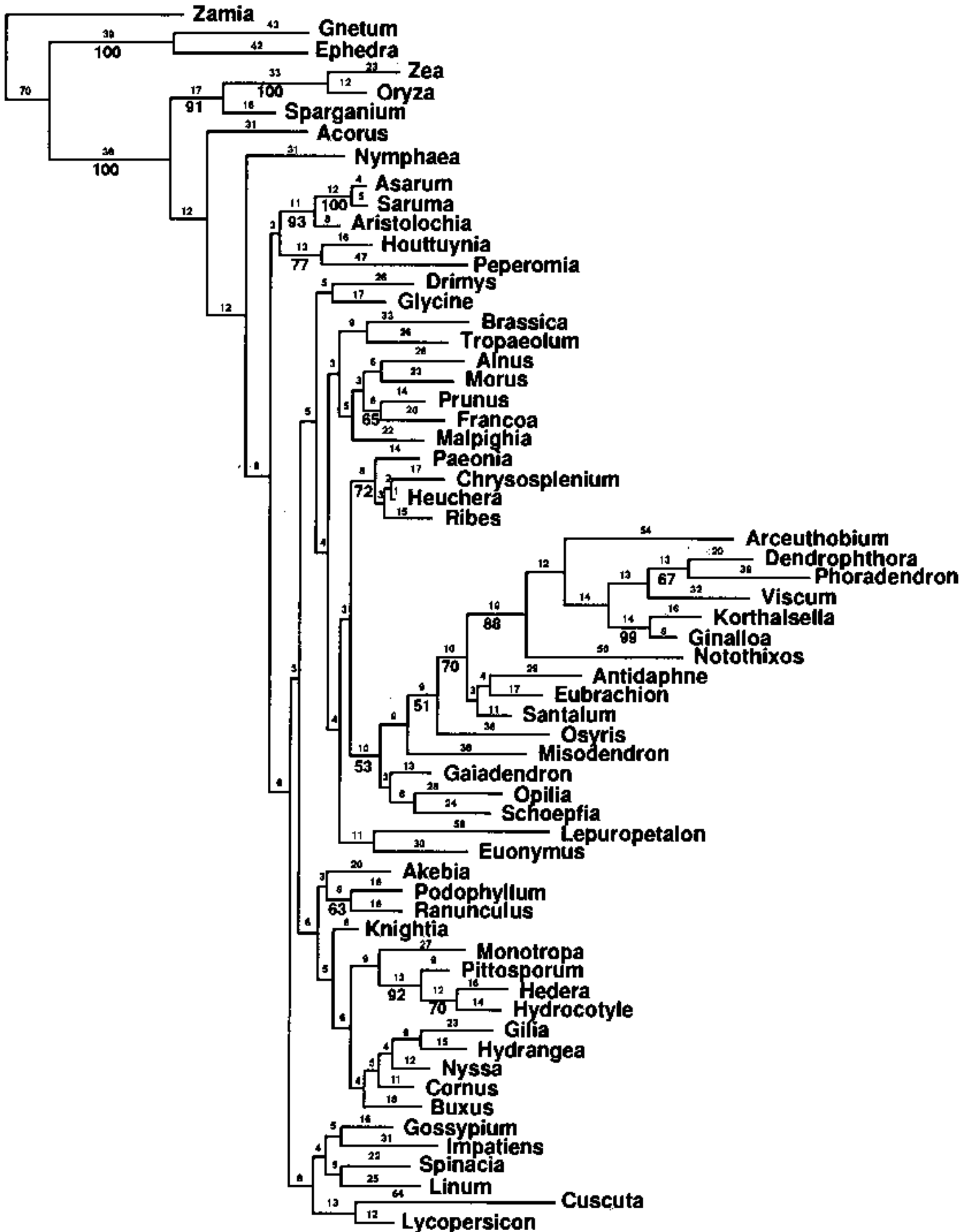


FIGURE 8. One of the 26 equally most parsimonious phylograms derived from the heuristic search of the 62-taxon 18S rDNA matrix; tree length 2021, C.I. minus uninformative sites = 0.301, R.I. = 0.440. Numbers above the branches indicate branch lengths (i.e., number of steps or nucleotide substitutions). Numbers below the branches indicate the percentage of trees (from 100 bootstrap replications) that support that node. Branches without bootstrap percentages were found in less than 50% of the trees.

sequence analysis, a close relationship is apparent between *Lepuropetalon* and *Euonymus*, and also *Morus* and *Betulaceae*. However, the position of Linaceae (represented by *Reinwardtia* in the *rbc*L analysis and by *Linum* in the 18S analysis) differs markedly between the 18S and *rbc*L trees. *Reinwardtia* is part of the Rosid I clade in the *rbc*L tree, whereas *Linum* is only distantly related to other Rosid I taxa in the 18S trees. Similarly, the placement of Fabaceae differs in the two analyses with *Pisum* appearing with other Rosid I taxa in the *rbc*L tree, but with *Glycine* appearing as the sister of *Drimys* in the 18S tree.

Both 18S and *rbc*L sequence data suggest a close relationship between *Brassica* and *Tropaeolum* (Rosid II; see Rodman et al., 1993). However, the placement of *Gossypium* differs markedly in the two trees. In the *rbc*L tree (Fig. 5), *Gossypium* is the sister of *Brassica* and *Tropaeolum*; all three genera are part of the Rosid II clade of Chase et al. (1993). In contrast, *Gossypium* is the sister of *Impatiens* in the 18S tree (Fig. 8) and is well removed phylogenetically from other Rosid II taxa.

Phylogenetic analyses of 18S and *rbc*L sequences also agree in suggesting a close relationship among *Chrysosplenium*, *Heuchera*, and *Ribes*, members of the Rosid III clade of Chase et al. (1993). The 18S analysis also places *Paeonia* in this clade, as does one of the two searches of Chase et al. (1993). Both analyses also concur in recognizing a well-supported monophyletic Santalales, although the position of this large clade differs between the two analyses. In the *rbc*L tree, Santalales appear as the sister to members of Asteridae sensu lato (Olmstead et al., 1993), whereas in the 18S analysis (Fig. 8) Santalales form the sister group of the Rosid III clade.

Several of the relationships among Asterid taxa seen in the *rbc*L analysis are also found in the shortest 18S trees. For example, *Lycopersicon* and Convolvulaceae (represented by *Convolvulus* and *Cuscuta* in the *rbc*L and 18S analyses, respectively) are sister taxa in both analyses. These taxa represent the Asterid I group of Chase et al. (1993). Similarly, the Asterid II group of *Pittosporum*, *Hedera*, and *Hydrocotyle* form a monophyletic group in both analyses. The Asterid IV subclade (Chase et al., 1993) that includes *Nyssa*, *Cornus*, and *Hydrangea* (along with *Gilia*, Polemoniaceae) also forms a subclade in the 18S analysis. In the *rbc*L consensus tree, these three genera do not form a monophyletic clade but are closely allied basal members of an Asterid assemblage.

In contrast to these similarities between the 18S and *rbc*L trees, the placement of those taxa representing the Asterid III clade (Chase et al., 1993) differs between the most parsimonious 18S and *rbc*L trees. In the 18S analysis (Fig. 8), Polemoniaceae (represented by *Gilia*) and *Monotropa* appear in clades with other Asterids. As previously mentioned, *Impatiens* emerges as sister to *Gossypium*. In contrast, *Polemonium*, *Pyrola* (a genus closely allied with *Monotropa*, Kron & Chase, 1993), and *Impatiens* form a subclade allied with the Rosid I clade in the shortest *rbc*L trees. This difference, however, may well reflect taxon density given that *Polemonium* and *Pyrola* are part of the Asteridae sensu lato when larger numbers of *rbc*L sequences are analyzed (Chase et al., 1993; Olmstead et al., 1993).

Several other placements and relationships differ dramatically between the 18S and *rbc*L trees. These include the phylogenetic positions of *Buxus*, *Gossypium*, *Spinacia*, and the sister-group relationship of *Drimys* and *Glycine* suggested by the 18S analysis. The same close relationship between *Drimys* and *Glycine* was also seen in the ribosomal RNA phylogenetic analysis of Hamby & Zimmer (1992).

Relationships within Santalales. As noted above, we analyzed Santalales in more detail to compare the resolution of 18S and *rbc*L sequence data at lower taxonomic levels. The *rbc*L sequence data reveal the presence of three clades within the order: (1) *Gaiadendron*, *Misodendron*, *Schoepfia*, *Opilia*; (2) *Antidaphne*, *Eubrachion*, *Osyris*, *Santalum*; and (3) *Arceuthobium*, *Dendrophthora*, *Phoradendron*, *Ginalloa*, *Korthalsella*, *Nothothixos*, and *Viscum*. The first group, minus *Opilia*, is strongly supported (bootstrap value of 88%) as are the second (82%) and third groups (86%). Analysis of 18S sequences reveals a very similar pattern of relationship. The Viscaceae form a monophyletic group (88% bootstrap value). Considering group 2, *Antidaphne*, *Eubrachion*, and *Santalum* form a monophyletic clade, with *Osyris* as their sister. Lastly, *Gaiadendron*, *Opilia*, and *Schoepfia* form a subclade (minus *Misodendron*) that closely corresponds to the *rbc*L group 1.

Discussion

The goal of this project was not to resolve higher-level relationships among the angiosperms, but rather to evaluate the phylogenetic potential of complete 18S rDNA sequences through a comparison of molecular phylogenies derived from both *rbc*L and 18S rDNA using similar taxon sampling and identical density and familial representation. The enormous phylogenetic potential of *rbc*L se-

quences is now well documented by numerous studies. In contrast, the phylogenetic utility of entire plant small-subunit ribosomal RNA sequences may have been underestimated. The recent study of Hamby & Zimmer (1992) certainly suggested that partial 18S, as well as 26S, sequences might help resolve the deepest branches of angiosperm phylogeny. Nickrent & Franchina (1990) had previously demonstrated that complete sequences of the 18S region held considerable phylogenetic potential. The present study further illustrates the phylogenetic potential of entire 18S rDNA sequences.

The present study indicates clearly that the rate of evolution of 18S rDNA is lower than that of *rbc*L. The percentage of sites that are potentially phylogenetically informative is almost twice as high for *rbc*L as for 18S rDNA (33.6% vs. 18.4%). However, because the 18S region is almost 400 bp longer than *rbc*L, the ratio of the number of phylogenetically informative sites per molecule is only about 1.4 times greater for *rbc*L compared to 18S rDNA. Thus, the amount of variation afforded per molecule is more comparable than suggested by rate of evolution alone. Because the number of variable sites in 18S rDNA is lower than for *rbc*L, complete sequencing of the entire 18S region becomes more critical for phylogenetic inference. Not only does this approach maximize the number of variable sites, but complete sequencing concomitantly facilitates proper alignment of 18S sequences. The two molecules also differ greatly in terms of the distribution of variation along each respective DNA region (Figs. 3 and 4). That is, base substitutions are spread much more evenly across the entire length of *rbc*L than for 18S rDNA.

Considerable variation in the evolutionary rate of *rbc*L has been shown within the angiosperms (Wilson et al., 1990; Bousquet et al., 1992; Chase et al., 1993). Although lineage rate asymmetry can contribute to spurious branch attractions (Hendy & Penny, 1989; Albert et al., 1993), it may not be extensive enough between angiosperm lineages to be problematic in terms of phylogenetic reconstruction given sufficient taxon density (Chase et al., 1993). The extent of heterogeneity of evolutionary rates among most plant lineages for 18S rDNA is not yet known. Unequal rates of 18S rDNA sequence evolution are suggested, however, for some Santalales and other parasitic plants, which exhibit an accelerated rate of evolution compared to other angiosperms (Nickrent & Franchina, 1990; Nickrent & Starr, 1994). The number of nucleotide substitutions per site (K) in pairwise comparisons among five nonparasitic angiosperms averaged 0.036 for 18S rDNA (Nickrent & Starr, 1994). In contrast, pairwise comparisons using an obligate hemiparasite (*Arceuthobium*) and several holoparasites (*Prosopanche, Balanophora, Rafflesia*, and *Rhizanthes*) result in a mean K value of 0.115 (Nickrent & Starr, 1994). Investigations of other heterotrophic angiosperms such as *Pholisma* (Lennoaceae) and *Cuscuta* (Convolvulaceae) have revealed similarly high substitution rates (Nickrent & Colwell, 1994). Accelerated substitution rates may also be present in *Lepuropetalon* and *Peperomia* based on the very long branch lengths these taxa exhibit (56 and 46, respectively). These long branch lengths could, however, simply be an artifact of the low taxon density of this analysis. It is noteworthy that *Peperomia* and *Lepuropetalon* also have much longer branch lengths than do their sister taxa in the *rbc*L tree depicted herein (Fig. 6), but this was not the case in the larger analysis of Chase et al. (1993) in which closer relatives of these taxa were included. Regardless of the cause of the long branch lengths in *Peperomia* and *Lepuropetalon*, the phylogenetic position of these taxa is similar in both the 18S and *rbc*L trees shown herein.

To evaluate the phylogenetic potential of 18S rDNA sequences, it is also important to elucidate the impact of secondary structure of the 18S rRNA transcript on phylogenetic reconstruction. As reviewed recently (Dixon & Hillis, 1993), major questions remain regarding phylogenetic analysis of rRNA or rDNA data. These questions include: should loop bases (non-pairing bases) and stem bases (pairing bases) both be used in phylogenetic reconstruction and, if so, should bases from each class (stems and loops) be considered equally informative and independent? Wheeler & Honeycutt (1988) recommended that stem base nucleotides be eliminated from phylogenetic analyses, or weighted by one-half. In contrast, in a detailed analysis of 28S rRNA genes from vertebrates, Dixon & Hillis (1993) found that characters from both stems and loops contain phylogenetic information. In addition, they found that stem bases sustain a greater number of compensatory mutations than would be expected at random, but the number of such mutations was less than 40% of that expected under a hypothesis of perfect compensation to maintain secondary structure. Dixon and Hillis therefore suggested that the weighting of stem characters be reduced by no more than 20% relative to loop characters in phylogenetic analyses. In an analysis of 18S rRNA sequences from echinoderms, Smith (1989) similarly reported that paired nucleotides were phylogenetically informa-

tive. Although the methods are at present not fully developed, incorporation of information from rRNA secondary (and tertiary) structure in phylogeny reconstruction algorithms is taking place (Van de Peer et al., 1993). These issues will require more attention in future phylogenetic studies of plants that use rDNA.

Our comparison of a similar suite of 62 taxa for both *rbc*L and 18S rDNA sequences yielded phylogenetic trees with a number of similar features, although we emphasize again that these trees should not be viewed as rigorous phylogenetic hypotheses for angiosperms. Both analyses revealed a well-supported monophyletic Gnetales as sister to a monophyletic Magnoliophyta, a result not too surprising given the sampling of taxa used. Within the angiosperms, both analyses revealed a monophyletic group of monocots (*Zea, Oryza, Sparganium*) that did not include *Acorus* as sister to other angiosperms. The distinctiveness of *Acorus* within the monocots was recently emphasized by Duvall et al. (1993). In both analyses, *Nymphaea* occurred in a similar position as sister to all other dicots. Both 18S rDNA and *rbc*L analyses recognized several identical clades, including two groups of Paleoherbs (*Houttuynia, Peperomia*; and *Asarum, Saruma, Aristolochia*), Ranunculids (*Akebia*, Berberidaceae, *Ranunculus*), several groups of Rosids (*Brassica, Tropaeolum*; *Morus*, Betulaceae; *Lepuropetalon, Euonymus*; *Chrysosplenium, Heuchera, Ribes*), Santalales, and several groups of Asterids (*Pittosporum, Hedera, Hydrocotyle*; *Lycopersicon*, Convolvulaceae). On a broader scale, very similar patterns of relationship are suggested among many of the Rosids and Asterids. Furthermore, both analyses suggest the presence of a large eudicot clade. At a lower taxonomic level, nearly identical subclades were revealed within the Santalales by both 18S rDNA and *rbc*L sequences. The degree of resolution achieved within Santalales using 18S rDNA sequences may, in part, reflect the accelerated rate of evolution of this region in this group of plants (Nickrent & Franchina, 1990; Nickrent & Starr, 1994).

The fact that phylogenetic analysis of 18S rDNA sequences for 62 taxa reveals relationships within angiosperms very similar to those obtained for a similar suite of taxa using *rbc*L sequences strongly suggests that questions of higher-level phylogeny in the angiosperms, as well as in seed plants in general, can be addressed with entire 18S rDNA sequences. The differences between the 18S and *rbc*L trees compared herein could, in large part, reflect taxon density, and also the use of different genera to represent some families (e.g., Polemoniaceae, Buxaceae, Convolvulaceae). Furthermore, the differences between the phylogenetic relationships gleaned from *rbc*L and 18S rDNA data may derive from their being, respectively, plastid and nuclear-encoded gene trees, neither of which perfectly represents the true species tree. Our analyses reinforce the findings of others (e.g., Nickrent & Franchina, 1990; Martin & Dowd, 1991; Hamby & Zimmer, 1992; Conran & Dowd, 1993; Hoot et al., 1995, this issue) in suggesting that sequencing of the 18S rDNA region holds considerable phylogenetic potential.

Although comparative sequencing of the entire 18S rDNA region holds potential for inferring phylogeny, we stress that this nuclear region will almost certainly not elucidate familial and generic level relationships to the extent possible with *rbc*L sequences simply because of the slower rate of evolution and lower overall number of base substitutions of 18S rDNA compared to *rbc*L. Whereas comparative *rbc*L sequencing has been used to resolve relationships within some angiosperm and gymnosperm families, including Onagraceae (Conti et al., 1993), Rosaceae (Morgan et al., 1994), Saxifragaceae s.s. (Morgan & Soltis, 1993), Taxodiaceae (Brunsfeld et al., 1994), Cupressaceae (Gadek & Quin, 1993), and Ericaceae (Kron & Chase, 1993), similar resolution with 18S sequences seems unlikely. In some santalalean families such as Viscaceae, 18S rDNA sequences have resolved generic-level relationships in a fashion comparable to that achieved via comparative *rbc*L sequencing. However, it is likely that the ability to resolve subfamilial and generic relationships in Santalales with 18S sequence data was facilitated by the higher substitution rate for this region in these taxa. Nonetheless, these results for Santalales illustrate that, in some instances, 18S sequence variation can be useful within families. Concomitantly, these findings also indicate that the ability of 18S rDNA sequences to provide sufficient resolution within any particular order or family must be determined empirically, just as for *rbc*L.

This study suggests that comparative sequencing of the 18S region should prove most useful for addressing phylogenetic relationships at the family level and above. Our results parallel those of Hoot et al. (1995) who showed in an analysis of Lardizabalaceae and other ranunculids that 18S rDNA sequences are more conserved than the two chloroplast genes employed (*rbc*L and *atp*B), but were useful in resolving relationships above the level of family. 18S rDNA sequence variation may be particularly well suited for addressing deeper phylo-

genetic branches within the angiosperms and in seed plants in general. The present study certainly indicates that additional sequencing of the entire 18S rDNA region is justified to obtain a broad sampling of angiosperms and other seed plants for eventual comparison with *rbc*L-based tree topologies (e.g., Chase et al., 1993). At present, only approximately 150 complete angiosperm 18S rDNA sequences exist (compared to over 1500 *rbc*L sequences for angiosperms). Further 18S rDNA sequencing within the monocots, Magnoliidae, Caryophyllidae, Hamamelidae, and Dilleniidae (sensu Cronquist) is especially needed to achieve greater taxon density for the angiosperms.

Literature Cited

ALBERT, V. A., M. W. CHASE & B. D. MISHLER. 1993. Character-state weighting for cladistic analysis of protein-coding DNA sequences. Ann. Missouri Bot. Gard. 80: 752–766.

———, S. E. WILLIAMS & M. W. CHASE. 1992. Carnivorous plants: Phylogeny and structural evolution. Science 257: 1491–1495.

APPELS, R. & R. L. HONEYCUTT. 1986. rDNA: Evolution over a billion years. Pp. 81–135 *in* S. K. Dutta (editor), DNA Systematics, Vol. II, Plants. CRC Press, Boca Raton, Florida.

ARNHEIM, N. 1983. Concerted evolution of multigene families. Pp. 38–61 *in* M. Nei & R. K. Koehn (editors), Evolution of Genes and Gene Proteins. Sinauer Associates, Sunderland, Massachusetts.

———, M. KRYSTAL, R. SCHMICKEL, G. WILSON, O. RYDER & E. ZIMMER. 1980. Molecular evidence for genetic exchanges among ribosomal genes on nonhomologous chromosomes in man and apes. Proc. Natl. Acad. U.S.A. 77: 7323–7327.

BALDWIN, B. G., M. J. SANDERSON, J. M. PORTER, M. F. WOJCIECHOWSKI, C. S. CAMPBELL & M. J. DONOGHUE. 1995. The ITS region of nuclear ribosomal DNA: A valuable source of evidence on angiosperm phylogeny. Ann. Missouri Bot. Gard. 82: 247–277.

BHATTACHARYA, D. & L. D. DRUEHL. 1988. Phylogenetic comparison of the small-subunit ribosomal DNA sequence of *Costaria costata* (Phaeophyta) with those of other algae, vascular plants and oomycetes. J. Phycol. 24: 539–543.

BOULTER, D. & J. S. GILROY. 1992. Partial sequences of 18S ribosomal RNA of two genera from each of six flowering plant families. Phytochemistry 31: 1243–1246.

BOUSQUET, J., S. H. STRAUSS, A. H. DOERKSEN & R. A. PRICE. 1992. Extensive variation in evolutionary rate of *rbc*L gene sequences among seed plants. Proc. Natl. Acad. U.S.A. 89: 7844–7848.

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

BROWN, D. D., P. C. WENSINK & E. JORDAN. 1972. A comparison of the ribosomal DNA's of *Xenopus laevis* and *Xenopus mulleri*: The evolution of tandem genes. J. Molec. Biol. 63: 57–73.

BRUNSFELD, S. J., P. S. SOLTIS, D. E. SOLTIS, P. A. GADEK, C. J. QUINN, D. D. STRENGE & T. A. RANKER. 1994. Phylogenetic relationships among genera of Taxodiaceae and Cupressaceae: Evidence from *rbc*L sequences. Syst. Bot. 19: 253–262.

BUCHHEIM, M. A., M. TURMEL, E. A. ZIMMER & R. L. CHAPMAN. 1990. Phylogeny of *Chlamydomonas* (Chlorophyta) based on cladistic analysis of nuclear 18S rRNA sequence data. J. Phycol. 26: 689–699.

BULT, C. J. & E. A. ZIMMER. 1993. Nuclear ribosomal RNA sequences for inferring tribal relationships within Onagraceae. Syst. Bot. 18: 48–63.

CHAPMAN, R. L. & M. BUCHHEIM. 1991. Ribosomal RNA gene sequences: Analysis and significance in the phylogeny and taxonomy of green algae. C.R.C. Crit. Rev. Pl. Sci. 10: 343–368.

CHASE, M. W., D. E. SOLTIS, R. G. OLMSTEAD, D. MORGAN, D. H. LES, B. D. MISHLER, M. R. DUVALL, R. A. PRICE, H. G. HILLS, Y.-L. QIU, K. A. KRON, J. H. RETTIG, E. CONTI, J. D. PALMER, J. R. MANHART, K. J. SYTSMA, H. J. MICHAELS, W. J. KRESS, K. G. KAROL, W. D. CLARK, M. HEDREN, B. S. GAUT, R. K. JANSEN, K.-J. KIM, C. F. WIMPEE, J. F. SMITH, G. R. FURNIER, S. H. STRAUSS, Q.-Y. XIANG, G. M. PLUNKETT, P. S. SOLTIS, S. M. SWENSEN, S. E. WILLIAMS, P. A. GADEK, C. J. QUINN, L. E. EGUIARTE, E. GOLENBERG, G. H. LEARN, JR., S. W. GRAHAM, S. C. H. BARRETT, S. DAYANANDAN & V. A. ALBERT. 1993. Phylogenetics of seed plants: An analysis of nucleotide sequences from the plastid gene *rbc*L. Ann. Missouri Bot. Gard. 80: 528–580.

CHAW, S.-M., H. LONG, B.-S. WANG, A. ZHARKIKH & W.-H. LI. 1993. The phylogenetic position of Taxaceae based on 18S rRNA sequences. J. Molec. Evol. 37: 624–630.

CONRAN, J. G. & J. M. DOWD. 1993. The phylogenetic relationships of *Byblis* and *Roridula* (Byblidaceae-Roridulaceae) inferred from partial 18S ribosomal RNA sequences. Pl. Syst. Evol. 188: 73–86.

CONTI, E., A. FISCHBACH & K. J. SYTSMA. 1993. Tribal relationships in Onagraceae: Implications from *rbc*L sequence data. Ann. Missouri Bot. Gard. 80: 672–685.

CRONQUIST, A. 1981. An Integrated System of Classification of Flowering Plants. Columbia Univ. Press, New York.

DIXON, M. T. & D. M. HILLIS. 1993. Ribosomal RNA secondary structure—Compensatory mutations and implications for phylogenetic analysis. Molec. Biol. Evol. 10: 256–267.

DONOGHUE, M. J. & J. A. DOYLE. 1989. Phylogenetic analysis of angiosperms and the relationships of Hamamelidae. Pp. 17–45 *in* P. R. Crane & S. Blackmore (editors), Evolution, Systematics, and Fossil History of the Hamamelidae, Vol. 1: Introduction and 'Lower' Hamamelidae. Clarendon Press, Oxford.

DOVER, G. A. 1982. Molecular drive: A cohesive mode of species evolution. Nature 299: 111–117.

———. 1987. DNA turnover and the molecular clock. J. Molec. Evol. 26: 47–58.

DOYLE, J. J. 1992. Gene trees and species trees: Molecular systematics as one-character taxonomy. Syst. Bot. 17: 144–163.

——— & J. L. DOYLE. 1987. A rapid DNA isolation procedure for small quantities of fresh leaf tissue. Phytochem. Bull. 19: 11–15.

DUVALL, M. R., M. T. CLEGG, M. W. CHASE, W. D. CLARK, W. J. KRESS, H. G. HILLS, L. E. EGUIARTE, J. F. SMITH, B. S. GAUT, E. A. ZIMMER & G. H.

LEARN, JR. 1993. Phylogenetic hypotheses for the monocotyledons constructed from *rbc*L sequence data. Ann. Missouri Bot. Gard. 80: 607–619.

ECKENRODE, V. K., J. ARNOLD & R. B. MEAGHER. 1985. Comparison of the nucleotide sequence of soybean 18S rRNA with the sequences of other small-subunit rRNAs. J. Molec. Evol. 21: 259–269.

FELSENSTEIN, J. 1985. Confidence limits on phylogenies: An approach using the bootstrap. Evolution 39: 783–791.

FIELD, K. G., G. J. OLSEN, D. J. LANE, S. J. GIOVANNONI, M. T. GHISELIN, E. C. RAFF, N. R. PACE & R. A. RAFF. 1988. Molecular phylogeny of the animal kingdom. Science 239: 748–753.

FÖRSTER, H., M. D. COFFEY, H. ELWOOD & M. L. SOGIN. 1990. Sequence analysis of the small subunit ribosomal RNAs of three zoosporic fungi and implications for fungal evolution. Mycologia 82: 306–312.

FRIEDLANDER, T. P., J. C. REGIER & C. MITTER. 1992. Nuclear gene sequences for higher level phylogenetic analysis: 14 promising candidates. Syst. Biol. 41: 483–490.

GADEK, P. A. & C. J. QUIN. 1993. An analysis of relationships within the Cupressaceae sensu stricto based on *rbc*L sequences. Ann. Missouri Bot. Gard. 80: 581–586.

GAUT, B. S., S. V. MUSE, W. D. CLARK & M. T. CLEGG. 1992. Relative rates of nucleotide substitution at the *rbc*L locus of monocotyledonous plants. J. Molec. Evol. 35: 292–303.

GUTELL, R. R. 1993. Collection of small subunit (16S- and 16S-like) ribosomal RNA structures. Nucl. Acids Res. 21: 3051–3054.

———, B. WEISER, C. R. WOESE & H. F. NOLLER. 1985. Comparative anatomy of 16-S-like ribosomal RNA. Prog. Nucl. Acid Res. & Molec. Biol. 32: 155–216.

HAMBY, K. R. & E. A. ZIMMER. 1988. Ribosomal RNA sequences for inferring phylogeny within the grass family (Poaceae). Pl. Syst. Evol. 160: 29–37.

——— & ———. 1992. Ribosomal RNA as a phylogenetic tool in plant systematics. Pp. 50–91 *in* P. S. Soltis, D. E. Soltis & J. J. Doyle (editors), Molecular Systematics of Plants. Chapman and Hall, New York.

HENDRIKS, L., R. DE BAERE, Y. VAN DE PEER, J. NEEFS, A. GORIS & R. DE WACHTER. 1991. The evolutionary position of the rhodophyte *Porphyra umbilicalis* and the basidiomycete *Leucosporidium scottii* among other eukaryotes as deduced from complete sequences of small ribosomal subunit RNA. J. Molec. Evol. 32: 167–177.

HENDY, M. D. & D. PENNY. 1989. A framework for the quantitative study of evolutionary trees. Syst. Zool. 38: 297–309.

HILLIS, D. M. & M. T. DIXON. 1991. Ribosomal DNA: Molecular evolution and phylogenetic inference. Quart. Rev. Biol. 66: 411–453.

HOOT, S. B., A. CULHAM & P. CRANE. 1995. The utility of *atp*B gene sequences in resolving phylogenetic relationships: Comparison with *rbc*L and 18S ribosomal DNA sequences in the Lardizabalaceae. Ann. Missouri Bot. Gard. 82: 194–208.

HORI, H. & S. OSAWA. 1987. Origin and evolution of organisms as deduced from 5S ribosomal RNA sequences. Molec. Biol. Evol. 4: 455–472.

———, B.-L. LIM & S. OSAWA. 1985. Evolution of green plants as deduced from 5S rRNA sequences. Proc. Natl. Acad. U.S.A. 82: 820–823.

HUSS, V. A. R. & M. L. SOGIN. 1990. Phylogenetic position of some *Chlorella* species within the Chlorococcales based upon complete small-subunit ribosomal RNA sequences. J. Molec. Evol. 31: 432–442.

JOHNSON, L. A. & D. E. SOLTIS. 1995. Phylogenetic inference in Saxifragaceae sensu stricto and *Gilia* (Polemoniaceae) using *mat*K sequences. Ann. Missouri Bot. Gard. 82: 149–175.

JORGANSEN, R. A. & P. D. CLUSTER. 1988. Modes and tempos in the evolution of nuclear ribosomal DNA: New characters for evolutionary studies and new markers for genetic and population studies. Ann. Missouri Bot. Gard. 75: 1238–1247.

KANTZ, T. S., E. C. THERIOT, E. A. ZIMMER & R. L. CHAPMAN. 1990. The Pleurastrophyceae and Micromonadophyceae: A cladistic analysis of nuclear rRNA sequence data. J. Phycol. 26: 711–721.

KIM, K.-J., R. K. JANSEN, R. S. WALLACE, H. J. MICHAELS & J. D. PALMER. 1992. Phylogenetic implications of *rbc*L sequence variation in Asteraceae. Ann. Missouri Bot. Gard. 79: 428–445.

KISS, T., A. SZKUKALEK & F. SOLYMOSY. 1989. Nucleotide sequence of a 17S rRNA gene from tomato. Nucl. Acids Res. 17: 2127.

KRON, K. A. & M. W. CHASE. 1993. Systematics of the Ericaceae, Empetraceae, Epacridaceae and related taxa based upon *rbc*L sequence data. Ann. Missouri Bot. Gard. 80: 735–741.

LANE, D. J., B. PACE, G. J. OLSEN, D. A. STAHL, M. L. SOGIN & N. R. PACE. 1985. Rapid determination of 16S ribosomal RNA sequences for phylogenetic analyses. Proc. Natl. Acad. U.S.A. 82: 6955–6959.

LES, D. H., D. K. GARVIN & C. F. WIMPEE. 1991. Molecular evolutionary history of ancient aquatic angiosperms. Proc. Natl. Acad. U.S.A. 88: 10119–10123.

MADDISON, D. R. 1991. Discovery and importance of multiple islands of most-parsimonious trees. Syst. Zool. 40: 315–328.

MADDISON, W. P. & D. R. MADDISON. 1992. MacClade, Analysis of Phylogeny and Character Evolution, version 3.01. Sinauer Associates, Sunderland, Massachusetts.

MARTIN, P. G. & J. M. DOWD. 1991. A comparison of 18S ribosomal RNA and rubisco large subunit sequences for studying angiosperm phylogeny. J. Molec. Evol. 33: 274–282.

——— & ———. 1993. Partial sequences of ribosomal RNA of Papilionaceae and related families. Phytochemistry 33: 361–363.

MESSING, J., J. CARLSON, G. HAGEN, I. RUBENSTEIN & A. OLESON. 1984. Cloning and sequencing of the ribosomal RNA genes in maize: the 17S region. DNA 3: 31–40.

MICHAELS, H. J., K. M. SCOTT, R. G. OLMSTEAD, T. SZARO, R. K. JANSEN & J. D. PALMER. 1993. Interfamilial relationships of the Asteraceae: Insights from *rbc*L sequence variation. Ann. Missouri Bot. Gard. 80: 742–751.

MISHLER, B. D., K. BREMER, C. J. HUMPHRIES & S. P. CHURCHILL. 1988. The use of nucleic acid sequence data in phylogenetic reconstruction. Taxon 37: 391–395.

Miyamoto, M. M. & J. Cracraft. 1991. Phylogenetic inference, DNA sequence analysis, and the future of molecular systematics. Pp. 3–17 *in* M. M. Miyamoto & J. Cracraft (editors), Phylogenetic Analysis of DNA Sequences. Oxford Univ. Press, New York.

Morgan, D. R. & D. E. Soltis. 1993. Phylogenetic relationships among members of Saxifragaceae sensu lato based on *rbc*L sequence data. Ann. Missouri Bot. Gard. 80: 631–660.

———, ———, & K. R. Robertson. 1994. Systematic and evolutionary implications of *rbc*L sequence variation in Rosaceae. Amer. J. Bot. 81: 890–903.

Mullis, K. B. & F. A. Faloona. 1987. Specific synthesis of DNA *in vitro* via a polymerase-catalyzed chain reaction. Meth. Enzymol. 155: 335–350.

Nairn, C. J. & R. J. Ferl. 1988. The complete nucleotide sequence of the small-subunit ribosomal RNA coding region for the cycad *Zamia pumila*: Phylogenetic implications. J. Molec. Evol. 27: 133–141.

Neefs, J.-M., Y. Van de Peer, P. De Rijk, S. Chapelle & R. De Wachter. 1993. Compilation of small subunit RNA structures. Nucl. Acids Res. 21: 3025–3049.

Nickrent, D. L. 1994. From field to film: Rapid sequencing methods for field collected plant species. BioTechniques 16: 470–474.

——— & A. Colwell. 1994. Elevated nucleotide substitution rates in nuclear 18S rDNA from heterotrophic angiosperms. Amer. J. Bot. 81: 176. [Abstract.]

——— & C. R. Franchina. 1990. Phylogenetic relationships of the Santalales and relatives. J. Molec. Evol. 31: 294–301.

——— & M. L. Sargent. 1991. An overview of the secondary structure of the V4 region of eukaryotic small-subunit ribosomal RNA. Nucl. Acids Res. 19: 227–235.

——— & E. M. Starr. 1994. High rates of nucleotide substitution in nuclear small-subunit (18S) rDNA from holoparasitic flowering plants. J. Molec. Evol. 39: 62–70.

Nishizawa, Y. & A. Hirai. 1987. Nucleotide sequence and expression of the gene for the large subunit of rice ribulose 1,5-bisphosphate carboxylase. Jap. J. Genet. 62: 389–395.

Olmstead, R. G. & P. A. Reeves. 1995. Evidence for the polyphyly of the Scrophulariaceae based on chloroplast *rbc*L and *ndh*F sequences. Ann. Missouri Bot. Gard. 82: 176–193.

———, B. Bremer, K. M. Scott & J. D. Palmer. 1993. A parsimony analysis of the Asteridae sensu lato based on *rbc*L sequences. Ann. Missouri Bot. Gard. 80: 700–722.

———, H. J. Michaels, K. M. Scott & J. D. Palmer. 1992. Monophyly of the Asteridae and identification of their major lineages inferred from DNA sequences of *rbc*L. Ann. Missouri Bot. Gard. 79: 249–265.

Olsen, G. J. 1987. Earliest phylogenetic branchings: Comparing rRNA-based evolutionary trees inferred with various techniques. Cold Spring Harbor Symp. Quant. Biol. 52: 825–837.

Palmer, J. D. 1985. Evolution of chloroplast and mitochondrial DNA in plants and algae. Pp. 131–240 *in* R. J. MacIntyre (editor), Monographs in Evolutionary Biology: Molecular Evolutionary Genetics. Plenum Press, New York.

———, R. K. Jansen, H. J. Michaels, M. W. Chase & J. M. Manhart. 1988. Analysis of chloroplast DNA variation. Ann. Missouri Bot. Gard. 75: 1180–1206.

Price, R. A. & J. D. Palmer. 1993. Phylogenetic relationships of the Geraniaceae and Geraniales from *rbc*L sequence comparisons. Ann. Missouri Bot. Gard. 80: 661–671.

Qiu, Y.-L., M. W. Chase, D. H. Les & C. R. Parks. 1993. Molecular phylogenetics of the Magnoliidae: Cladistic analyses of nucleotide sequences of the plastid gene *rbc*L. Ann. Missouri Bot. Gard. 80: 587–606.

Rathgeber, J. & I. Capesius. 1990. Nucleotide sequence of the intergenic spacer and the 18S ribosomal RNA gene from mustard (*Sinapis alba*). Nucl. Acids Res. 18: 1288.

Rieseberg, L. & D. E. Soltis. 1991. Phylogenetic consequences of cytoplasmic gene flow in plants. Evol. Trends Pl. 5: 65–84.

Ritland, K. & M. T. Clegg. 1987. Evolutionary analyses of plant DNA sequences. Amer. Naturalist 130: S74–S100.

Rodman, J., R. A. Price, K. Karol, E. Conti, K. J. Sytsma & J. D. Palmer. 1993. Nucleotide sequences of the *rbc*L gene indicate monophyly of mustard oil plants. Ann. Missouri Bot. Gard. 80: 686–699.

Sanderson, M. J. & J. J. Doyle. 1993. Reconstruction of organismal gene phylogenies from data on multigene families: Concerted evolution, homoplasy, and confidence. Syst. Biol. 4: 4–17.

Sanger, F., S. Nicklen & A. R. Coulson. 1977. DNA sequencing with chain terminating inhibitors. Proc. Natl. Acad. U.S.A. 74: 5463–5467.

Savard, L. & M. Lalonde. 1991. Sequence of 18S rDNA of actinorhizal *Alnus glutinosa* (Betulaceae). Pl. Molec. Biol. 16: 725–728.

Schlegel, M., H. J. Elwood & M. L. Sogin. 1991. Molecular evolution in hypotrichous ciliates: Sequence of the small subunit ribosomal RNA genes from *Onychodromus quadricornutus* and *Oxytricha granulifera* (Oxytrichidae, Hypotrichida, Ciliophora). J. Molec. Evol. 32: 64–69.

Senecoff, J. & R. B. Meagher. 1992. *In vivo* analysis of plant RNA structure: Soybean 18S ribosomal and ribulose-1,5-bisphosphate carboxylase small subunit RNAs. Pl. Molec. Biol. 18: 219–234.

Simovic, N., D. Wolyn & G. Jelenkovic. 1992. Sequence analysis of 18S ribosomal RNA gene in *Fragaria* × *ananassa* Dutch. cultivated octoploid strawberry. Pl. Molec. Biol. 18: 1217–1220.

Smith, A. B. 1989. RNA sequence data in phylogenetic reconstruction: Testing the limits of its resolution. Cladistics 5: 321–344.

Smith, S. 1992. Genetics Data Environment, Version 2.0. Harvard Genome Lab, Cambridge, Massachusetts.

Sogin, M. L., H. J. Elwood & J. H. Gunderson. 1986. Evolutionary diversity of eukaryotic small-subunit rRNA genes. Proc. Natl. Acad. U.S.A. 83: 1383–1387.

Soltis, D. E., P. S. Soltis, M. T. Clegg & M. L. Edgerton. 1990. *rbc*L sequence divergence and phylogenetic relationships in Saxifragaceae sensu lato. Proc. Natl. Acad. U.S.A. 87: 4640–4644.

Specht, T., J. Wolters & V. A. Erdmann. 1991.

Compilation of 5S rRNA and 4.8S rRNA gene sequences. Nucl. Acids Res. 19: 2189–2191.

SWANN, E. C. & J. W. TAYLOR. 1993. Higher taxa of Basidiomycetes: An 18S rRNA gene perspective. Mycologia 85: 923–936.

SWOFFORD, D. L. 1993. PAUP—Phylogenetic Analysis Using Parsimony, version 3.1. Center for Biodiversity, Illinois Natural History Survey, Champaign, Illinois.

——— & G. J. OLSEN. 1990. Phylogeny reconstruction. Pp. 411–501 *in* D. M. Hillis & C. Moritz (editors), Molecular Systematics. Sinauer Associates, Sunderland, Massachusetts.

TAKAIWA, F., K. OONO & M. SUGIURA. 1984. The complete nucleotide sequence of a rice 17S rRNA gene. Nucl. Acids Res. 12: 5441–5448.

THOMPSON, W. F. & R. B. FLAVELL. 1988. DNase I sensitivity of ribosomal RNA genes in chromatin and nucleolar dominance in wheat. J. Molec. Biol. 204: 535–548.

TROITSKY, A. V., Y. F. MELEKHOVETS, G. M. RAKHIMOVA & V. K. BOBROVA. 1991. Angiosperm origin and early stages of seed plant evolution deduced from rRNA sequence comparison. J. Molec. Evol. 32: 253–261.

UNFRIED, I., U. STOCKER & P. GRUENDLER. 1989. Nucleotide sequence of the 18S rRNA gene from *Arabidopsis thaliana* Col0. Nucl. Acids Res. 17: 7513.

VAN DE PEER, Y., J.-M. NEEFS, P. DE RIJK & R. DE WACHTER. 1993. Reconstructing evolution from eukaryotic small-ribosomal-subunit RNA sequences: Calibration of the molecular clock. J. Molec. Evol. 37: 221–232.

WADA, H. & N. SATOH. 1994. Phylogenetic relationships among extant classes of echinoderms, as inferred from sequences of 18S rDNA, coincide with relationships deduced from the fossil record. J. Molec. Evol. 38: 41–49.

WHEELER, W. C. & R. L. HONEYCUTT. 1988. Paired sequence difference in ribosomal RNAs: Evolutionary and phylogenetic implication. Molec. Biol. Evol. 5: 90–96.

WILSON, M. A., B. GAUT & M. T. CLEGG. 1990. Chloroplast DNA evolves slowly in the palm family. Molec. Biol. Evol. 7: 303–314.

WOESE, C. R. 1987. Bacterial evolution. Microbiol. Rev. 51: 221–271.

WOLTERS, J. & V. A. ERDMANN. 1986. Cladistic analysis of 5S rRNA and 16S rRNA secondary and primary structure—The evolution of eukaryotes and their relation to Archaebacteria. J. Molec. Evol. 24: 152–166.

XIANG, Q.-Y., D. E. SOLTIS, D. R. MORGAN & P. S. SOLTIS. 1993. Phylogenetic relationships of *Cornus* L. sensu lato and putative relatives inferred from *rbc*L sequence data. Ann. Missouri Bot. Gard. 80: 723–734.

ZIMMER, E. A., R. K. HAMBY, M. L. ARNOLD, D. A. LEBLANC & E. C. THERIOT. 1989. Ribosomal RNA phylogenies and flowering plant evolution. Pp. 205–214 *in* B. Fernholm, K. Bremer & J. Jörnvall (editors), The Hierarchy of Life. Elsevier Science, Amsterdam.

———, S. L. MARTIN, S. M. BEVERLEY, Y. W. KAN & A. C. WILSON. 1980. Rapid duplication and loss of genes coding for the a chains of hemoglobin. Proc. Natl. Acad. U.S.A. 77: 2158–2162.

ZURAWSKI, G., P. R. WHITFELD & W. BOTTOMLEY. 1986. Sequence of the gene for the large subunit of ribulose 1,5-bisphosphate carboxylase from pea chloroplasts. Nucl. Acids Res. 14: 3975.

———, B. PERROT, W. BOTTOMLEY & P. R. WHITFELD. 1981. The structure of the gene for the large subunit of ribulose 1,5-bisphosphate carboxylase from spinach chloroplast DNA. Nucl. Acids Res. 14: 3251–3270.