MOLECULAR STUDIES OF PARASITIC PLANTS USING RIBOSOMAL RNA

DANIEL L. NICKRENT and R. JOEL DUFF. Department of Plant Biology, Southern Illinois University, Carbondale, Illinois, USA 62901-6509

CITATION:

Nickrent, D. L. and R. J. Duff. 1996. Molecular studies of parasitic plants using ribosomal RNA. Pp. 28-52. In: M. T. Moreno, J. I. Cubero, D. Berner, D. Joel, L. J. Musselman, C. Parker (eds.), Advances in Parasitic Plant Research, Junta de Andalucia, Dirección General de Investigación Agraria, Cordoba, Spain.

ABSTRACT

Molecular evolutionary and phylogenetic studies of parasitic flowering plants have been advanced through the use of nuclear and plastid-encoded ribosomal RNA genes. Phylogenetic analysis of all families of Santalales using nuclear 18S rDNA sequences supports the basal position of Olacaceae, the monophyly of Opiliaceae, the distinctiveness of Loranthaceae and Viscaceae, and the sister relationship between Santalaceae (including Eremolepidaceae) and Viscaceae. These data do not support a close relationship between Santalales and holoparasites in Balanophoraceae, Hydnoraceae, and Rafflesiaceae. High nucleotide substitution rates in these plants hinder their placement in the global angiosperm phylogeny. Some support is obtained for the placement of Hydnoraceae with the paleoherbs (Magnoliidae). Analysis of 18S rDNA sequences for ten genera of Scrophulariaceae and Orobanchaceae indicate the latter family is paraphyletic. The holoparasites *Orobanche, Harveya, Hyobanche* and *Lathraea* appear on a clade with substitution rates higher than related hemiparasites. Plastid-encoded 16S rDNA was PCR amplified and sequenced from representatives of Balanophoraceae, Hydnoraceae, and Rafflesiaceae, thereby providing preliminary evidence for the existence of a plastid genome. The 16S rRNA sequence of *Cytinus ruber* (Rafflesiaceae or Cytinaceae) has mutations at 8.6% of the 1497 sites, yet structural integrity (and likely functionality) is maintained. Formal relative rate tests document that *Cytinus* is more divergent than any of the previously sequenced plant 16S rRNAs. These holoparasitic plants should be viewed as valuable model organisms that can increase our understanding of the mode and tempo of evolutionary change at the molecular level.

Additional key words: Nuclear rDNA, Phylogeny, Evolution

INTRODUCTION

Within angiosperms, haustorial parasitism has evolved independently at least seven times and more likely at least nine times. Parasitic families are restricted to three of the six dicot subclasses of Cronquist (1988): Magnoliidae, Rosidae and Asteridae (Fig. 1). Strangely, true parasitism has apparently never evolved in monocots (Kuijt, 1969). All parasitic angiosperms develop haustorial attachments to their hosts, yet they vary widely in their nutritional dependence. Chlorophyllous hemiparasites can be found in Laurales (*Cassytha*), Polygalales (*Krameria*), and all families of Santalales. The evolution and biochemistry of two orders, Solanales (*Cuscuta*) and Scrophulariales, are of interest because some species are photosynthetic hemiparasites whereas other species have lost photosynthesis and become holoparasitic. Holoparasites represent the most extreme manifestation of the parasitic mode since they lack photosynthesis and must rely upon the host for water, and inorganic and organic nutrients. Four families (sensu Cronquist, 1988) are represented entirely by holoparasites: Lennoaceae, Balanophoraceae, Hydnoraceae, and Rafflesiaceae.

This paper will focus on the use of ribosomal RNA genes in studies of parasitic flowering plants, specifically molecular phylogenetic and molecular evolutionary analyses of several of the groups shown in Fig. 1. The three distinct genomes of plants (nuclear, plastid and mitochondrial) each contain a complement of the ribosomal RNA genes (rDNA). This paper will deal specifically with projects involving plastid 16S and nuclear 18S rDNA. The goal is not only to demonstrate the utility of these molecular markers in documenting phylogenetic history, but also to show how parasitic plants represent unique models that can be used to study molecular evolutionary and genetic processes.

Nuclear rRNA in Parasitic Plants

In eukaryotes, nuclear rDNA loci occur at nucleolar organizing regions (NORs) on one or several chromosomes. The characteristics and evolution of nuclear rDNA has been extensively discussed (Appels and Honeycutt, 1986; Jorgensen and Cluster, 1988; Hamby and Zimmer, 1992) and only the following brief review will be given. Ribosomal cistrons are repeated hundreds to thousands of times in tandem arrays within the genome. The genes that code for 18S, 5.8S and 26S rRNA occur within a cistron, each separated by spacers (Fig. 2A). The intergenic spacer (IGS) occurs between repeat types of each cistron and is composed of a nontranscribed spacer (NTS) bounded by external transcribed spacers (ETS) on each end. The internal transcribed spacer (ITS) exists in two

pieces, ITS-1 between the 18S and 5.8S rDNA and ITS-2 between the 5.8S and 26S rDNA. After transcription and processing, the 18S rRNA resides in the small ribosomal subunit and the 5.8S and 26S rRNAs reside in the large ribosomal subunit of the mature ribosome. These rRNA molecules fold into higher-order structures that are directly involved in the translation process.

Sequence similarity between the individual cistrons within a single organism is generally very high, possibly due to unequal crossing over during meiosis, gene conversion, slippage, transposition, and RNA mediated changes (Arnheim, 1983; Dover, 1982; 1987). This homogenizing process that occurs in multigene families has been called concerted evolution (Brown *et al.*, 1972; Zimmer *et al.*, 1980; Sanderson and Doyle, 1993) and ribosomal loci represent an extreme case. This characteristic has also made these genes useful for examining deep phylogenetic divergences (Cedergren *et al.*, 1988; Sogin *et al.*, 1986). The ribosomal cistron is not invariant over its entire length, but is a mosaic of slowly and rapidly evolving regions (Fig. 2A). Variable and conserved domains exist on both 18S and 26S rDNA. The actual utility of 18S rDNA in molecular phylogenetic studies of angiosperms has only recently been realized (Nickrent and Soltis, 1995; Soltis *et al.*, submitted). Spacer regions, such as ITS, have recently attracted much interest since their level of variability is high enough to allow phylogenetic analysis within and among genera (Baldwin, 1995; Nickrent *et al.*, 1994).

Molecular Phylogenetic Studies of Santalales

One of the earliest studies to use 18S rRNA sequences in a phylogenetic analysis of angiosperms examined parasitic Santalales (Nickrent and Franchina, 1990). Using direct rRNA sequencing with reverse transcriptase, sequences from representatives of 10 families were analyzed. Although only three parasitic families were examined, the analyses supported the monophyly of Santalales, the sister group relationship between Viscaceae and Santalaceae and the basal position of Olacaceae within the order. This study also showed that 18S rRNA contained sufficient variation to conduct phylogenetic analyses in angiosperms. Since this initial work, 18S sequences of parasites in Santalales have increased rapidly, mainly owing to the advent of the polymerase chain reaction (PCR). Small-subunit rDNA is readily amplified from genomic DNA samples and can then be purified and directly sequenced. Sampling within Santalales has also steadily improved such that currently genomic DNA exists for representatives of all families of the order and, in some cases, all genera within a family.

The following section will discuss the results of preliminary phylogenetic analyses of Santalales using complete 18S rDNA sequences. The order is here defined as the following families: Olacaceae, Loranthaceae, Misodendraceae, Opiliaceae, Santalaceae and Viscaceae. The holoparasite families Balanophoraceae, Rafflesiaceae and Hydnoraceae are often classified in or near Santalales, however, since these relationships are not clear, these families will be treated separately (see below). The methods used to extract DNA, PCR amplify and sequence 18S rDNA, conduct multiple sequence alignments, and generate minimum-length Fitch parsimony trees (via PAUP - Swofford, 1993) are discussed in Nickrent and Soltis (1995). Figure 3 shows the results of parsimony analysis of 18S rDNA sequences from 62 members of Santalales and several outgroup taxa. Specific results are discussed below.

Olacaceae. The family has traditionally been considered the most primitive of the order based upon the presence of two ovular integuments in some taxa and the presence of both autotrophic and hemiparasitic members. Kuijt (1968, 1969) considered the Olacaceae the "plexus" from which all other families in the order were derived. The combination of primitive and specialized features and the very high number of monotypic genera prompted Sleumer (1984) to suggest the family differentiated early during the Cretaceous, prior to the separation of the continents. The Olacaceae is certainly the most problematic one in the order, for indeed extreme variability can be seen in morphological features such as habit (nonparasitic and parasitic), flower sexual condition (unisexual, perfect), petal fusion (distinct, connate), ovary position (hypogynous, epigynous, perigynous), ovular integuments (0, 1, 2), embryo sac type (monosporic, bisporic), and even cotyledon number (2, 3, 4 and 8). Such variation has prompted the erection (and subsequent subsumption) of numerous splinter families such as Aptandraceae, Cathedraceae, Erythropalaceae, Heisteriaceae, Octoknemaceae, Schoepfiaceae, Scorodocarpaceae, Strombosiaceae, and Tetrastylidiaceae. Sleumer (1935 and amended in 1984) divided the family into three subfamilies: Anacolosoideae (formerly Dysolacoideae), Olacoideae, and Schoepfioideae.

At present only seven of the ca. 28 genera in the family are represented in Fig. 3. All genera except *Schoepfia* are basal in the order, which is concordant with previous taxonomic systems. The genera do not form a monophyletic clade, hence the family is paraphyletic. *Schoepfia screberi* Gmelin forms a clade with *Misodendron brachystachyum*

DC and this clade appears basal to the Old World Loranthaceae. This same relationship is supported by analyses of the chloroplast gene *rbc*L (Nickrent, 1996) that included five genera of Olacaceae. *Schoepfia* is distinct from other olacaceous genera in possessing aliform-confluent parenchyma (Sleumer, 1984) and by its ratio of tracheid and vessel features (Reed, 1955). Reed (1955) also noted similarities in the pollen of *Schoepfia* with Santalaceae (a more advanced family of the order). Taken as a whole, it is worth considering possible phylogenetic affinities between *Schoepfia* and the root parasitic Loranthaceae such as *Atkinsonia* and *Nuytsia*. Additional sampling within Olacaceae is required to further address the possible polyphyly of the family. Assistance by colleagues is requested to acquire genera from tropical South America, Africa, and Indomalaya.

Misodendraceae. *Misodendron* consists of ca. 10 mistletoe species parasitic on *Nothofagus* of southern South America. The genus is unusual in having wind-dispersed fruits (via enlarged, plumed staminodes). Relatively little published information on phylogenetic affinities of this monogeneric family exists, although Kuijt (1968, 1969) derives the family from Olacaceae. Results of analyses of 18S rDNA (Fig. 3) and *rbc*L sequences (Nickrent and Soltis, 1995; Nickrent, 1996) show that *Misodendron* clusters with *Schoepfia* (Olacaceae), although connected to it by a long- branch. Given this molecular and biogeographic information, *Misodendron* may represent a relictual taxon that diverged early from loranthaceous or olacaceous stock present on the Gondwanan landmass.

Loranthaceae. Complete 18S rDNA sequences currently exist for 23 of the ca. 75 genera in this family. Results of phylogenetic analyses (Fig. 3) show that the family (with the inclusion of *Schoepfia* and *Misodendron*) is monophyletic An apparent feature of the tree is that significant genetic differentiation has occurred between clades composed of New World mistletoes (e.g. *Gaiadendron*, *Ligaria*, *Psittacanthus*, etc.) and Old World mistletoes (*Loranthus*, *Tapinanthus*, *Amyema*, etc.). The two mistletoes endemic to New Zealand (*Tupeia* and *Alepis*) show affinity with *Lysiana*, an Australian genus. *Gaiadendron* is classified with *Atkinsonia* in tribe Elytrantheae (Danser, 1933) and shares with it (and *Nuytsia*) N=12, the base chromosome number for the family. These three genera are considered the most primitive in the family (Barlow, 1983). The 18S rDNA data do not strongly support a basal position in the family for *Gaiadendron* (see, however, *rbc*L analyses in Nickrent, 1996). Although we have DNA for the root parasite *Nuytsia*, we have yet to generate any sequences. Tissue for another root-parasitic genus, *Atkinsonia*, has yet to be obtained.

By examining branch lengths, it is apparent that fewer mutations exist among loranthaceous than viscaceous genera. The molecular data strongly support the concept of Barlow (1983) that these two mistletoe families are distinct. Furthermore, more mutations occur among Old World than New World genera of Loranthaceae. A number of relationships support current concepts, such as associations between *Amyema* and *Diplatia*, *Oryctanthus* and *Dendropemon*, etc. Although 18S rDNA does provide some indications of intergeneric affinities, continued sequencing using a faster rate molecule is required to fully resolve relationships in this family. Sequences for *rbc*L for three genera (*Gaiadendron, Moquinella* and *Tupeia*) indicate that this molecule will also provide insufficient phylogenetic signal to address relationships among all genera. Our lab is currently exploring the utility of two chloroplast genes (*ndh*F and *mat*K) as well as 26S rDNA which contains more rapidly evolving domains (expansion segments).

Opiliaceae. Four of the possible nine genera of this family have been sequenced for 18S rDNA: *Agonandra*, *Cansjera*, *Champereia*, and *Opilia*. Results of this analysis (Fig. 3) represents the family as a monophyletic clade between the Santalaceae and Loranthaceae. Opiliaceae is one of the most coherent families within the order as reflected by the overall similarity in its wood structure (Reed, 1955). Similarities between Opiliaceae and Santalaceae can be seen in floral morphology (Fagerlind, 1948), embryology (Johri and Bhatnagar, 1960), and haustorial anatomy (Kubat, 1987). Although Hiepko (1979, 1982) published a taxonomic revision of the Old World members of the family, no intergeneric phylogenetic inferences were made, therefore the classification of Sleumer (1935) must be used. Therein, *Agonandra* (with *Gjellerupia*) was placed in its own tribe, Agondandreae. The remaining seven genera were classified within tribe Opilieae. rDNA analyses place *Opilia*, not *Agonandra*, at a basal position in the family, however, this is supported by only a few steps. In contrast, *rbc*L analyses (Nickrent, 1996) place *Agonandra* at the base of the clade in support of the Sleumer (1935) classification.

Santalaceae and Eremolepidaceae. Pilger (1935) divided the family into three tribes, the Anthoboleae (2 genera - *Anthobolus* and *Exocarpos*), Osyrideae (= Santaleae - 22 genera), and Thesieae (6 genera). Little subsequent taxonomic work has been conducted on higher-level relationships in the family. This analysis utilized sequences from 11 of the 30 genera of Santalaceae. The family does not form a monophyletic group but a grade that eventually culminates in Viscaceae (Fig. 3). Sequences for only two genera (*Osyris* and

Santalum) were used in the *rbc*L analysis by Nickrent (1996), however, they formed a monophyletic group (with Eremolepidaceae - see below) sister to Viscaceae. Several relationships are worth noting. *Buckleya* and *Pyrularia*, two relictual genera that have representatives in eastern North America and eastern China, form a clade at the base of the family. The morphologically similar north temperate genera *Geocaulon* and *Comandra*, form a clade with *Thesium*. An unresolved polytomy that includes *Osyris*, *Nestronia* and *Santalum* in one clade and *Dufrenoya* and *Dendrotrophe* in another is also present.

Three santalalean genera, *Antidaphne, Eubrachion* and *Lepidoceras*, were placed in the family Eremolepidaceae by Kuijt (1988) and allied with primitive Loranthaceae. The results of this 18S rDNA sequence analysis places two representatives (*Antidaphne* and *Eubrachion* - DNA has yet to be obtained from *Lepidoceras)* within the Santalaceae (Fig. 3). *Eubrachion* forms a clade with *Exocarpos,* the one representative of tribe Anthoboleae, and *Antidaphne* occupies a position intermediate between Santalaceae and Viscaceae. Results of analyses of *rbc*L and 18S rDNA sequences (Nickrent and Soltis, 1995; Nickrent, 1996) place both *Antidaphne* and *Eubrachion* on a clade with Santalaceae. These results support the statements made by Wiens and Barlow (1971) that Eremolepidaceae is not closely related to Viscaceae and that the three genera are not related to each other based upon karyological and morphological evidence. Embryological features for these two families, summarized by Bhandari and Vohra (1983), support an association between Eremolepidaceae and Santalaceae. Taken together, these results are intriguing since, prior to the present molecular study, the only aerially parasitic Santalaceae were Old World genera such as *Phacellaria, Dendromyza,* and *Cladomyza*. The status of Eremolepidaceae ultimately depends upon the placement of the third genus, *Lepidoceras*, which has resided with viscaceous mistletoes, the Loranthaceae (Kuijt 1968) and later Eremolepidaceae (Kuijt 1988).

Viscaceae. Analyses of nuclear 18S rDNA sequences of the Viscaceae confirms its monophyly, derivation from Santalaceae, and advanced position in the order (Fig. 3). Contrary to Bandahri and Vohra (1983), the family is distinct from Loranthaceae and, relative to it, exhibits increased evolutionary rates. This can be seen by examining branch lengths on Fig. 3, especially for advanced hemiparasites such as *Arceuthobium*. Sequences from both 18S rDNA and *rbc*L have been obtained from several representatives from each of the seven genera in the family. The three species of *Arceuthobium* (Old and New World species) form a clade that is sister to two Australian species of *Notothixos*. *Ginalloa*

arnottiana Korth. (from Borneo) clusters with two species of *Korthalsella* (from Hawaii and New Zealand). The next clade is composed of the New World genera *Dendrophthora* and *Phoradendron*. The last clade is composed of two species of *Viscum*. The order of branching for the above major clades is not resolved following bootstrap analysis, hence only the relationships between *Phoradendron* and *Dendrophthora* and *Korthalsella* and *Ginalloa* are strongly supported.

These results differ in several ways from the relationships proposed by Wiens and Barlow (1971) who used mainly chromosomal and biogeographical information. They proposed two major lines, one composed of *Viscum*, *Notothixos* and *Ginalloa*, the second composed of *Phoradendron*, *Dendrophthora*, *Korthalsella* and *Arceuthobium*. Given their similar distributions and shared karyotype of 12 or 13 large chromosomes, it is logical to derive a relationship between *Ginalloa* and *Notothixos,* as was done by Wiens and Barlow (1971). This relationship is not supported by either rDNA (Fig. 3) or *rbc*L (Nickrent, 1995; Nickrent and Soltis, 1995; Nickrent, 1996) sequence analysis which both link *Korthalsella* with *Ginalloa*. The array of chromosome numbers seen in *Viscum* (N=10, 11, 12, and 13) suggests that extensive genetic variance exists for this feature. The basal position of this genus following bootstrap analysis (results not shown), although not strongly supported, suggests radiation of the major viscaceous lines from a *Viscum*-like ancestor. Further studies of the molecular evolution in this family using 26S rDNA or other molecules is required to elucidate relationships in this family.

Molecular Phylogenetic Studies of Rafflesiales and Balanophoraceae

Over the past century, traditional means of classifying Balanophoraceae, Hydnoraceae, and Rafflesiaceae have met with difficulty owing to the extreme reduction and/or modification of morphological structures that have accompanied the evolution of these lineages. Moreover, the features that remain are often enigmatic, i.e. they are so unusual (derived) as to confound assessment of homology via character state transformation series. The disagreement regarding classification of these families can be seen by comparing three current systems for angiosperms. Cronquist (1981, 1988) placed Balanophoraceae in Santalales and classified Hydnoraceae and Rafflesiaceae together in the sister order Rafflesiales. The system of Takhtajan (1980, 1987) was similar in that Hydnoraceae and Rafflesiaceae were placed in Rafflesiales, however, this order was derived from within subclass Magnoliidae, not Rosidae. Takhtajan (1980) stated that Balanophoraceae (in Balanophorales) were "probably near to and derived from Santalales, but the affinity is not

fully clear." The system of Thorne (1992) differed little from the above by classifying Balanophorales near Santalales and Rafflesiales in superorder Rafflesianae.

Resolution of phylogenetic relationships in angiosperms has become increasingly refined using molecular markers such as *rbc*L and 18S rDNA (Chase *et al.*, 1993; Nickrent and Soltis, 1995). For this reason, the analysis of appropriately selected molecular markers holds great promise in placing the holoparasites within the global phylogeny of all angiosperms. Unfortunately, the chloroplast genome of these plants is likely highly modified, hence genes such as *rbc*L are not available for analysis (see below). As already demonstrated for Santalales, 18S rDNA sequences are appropriate molecular markers for addressing phylogenetic relationships. Complete 18S rDNA sequences have been obtained for representatives of Balanophoraceae, Hydnoraceae, and Rafflesiaceae. Analyses of sequences from these three holoparasite families showed increased nucleotide substitution rates as was first observed in 18S rDNA of *Arceuthobium*. Using relative rate tests, it was found that the holoparasites were, on average, 3.5 times faster than nonparasitic and hemiparasitic plants (Nickrent and Starr, 1994). Elevated substitution rates in genes normally considered extremely conservative, such as nuclear ribosomal genes, complicates phylogenetic analysis. For example, when these divergent sequences are included in an analysis with nonparasites, the resulting topologies frequently show aberrations such as "long-branch attractions" (Felsenstein, 1978) that artifactually links clearly unrelated taxa with faster rates. When fewer than 20 nonparasitic "outgroup" taxa are used, the parasites with long-branches migrate to the base of the angiosperm clade, intermediate between the angiosperms and the outgroup Gnetales. This same result is seen using parsimony, neighbor-joining and maximum likelihood methods. Further discussion of rates tests will be given in the section "Molecular Evolutionary Studies - Relative Rate Tests."

Recent sequencing efforts have greatly expanded the number of available 18S rDNA sequences in plants (Nickrent and Soltis, 1995; Soltis *et al.*, submitted.). At present, over 400 complete 18S sequences exist for angiosperms, i. e. sufficient density to avoid erroneous topologies stemming from insufficient taxon sampling (Chase *et al.*, 1993). This large dataset has allowed a re-examination of the placement of the above three holoparasite families. The results presented below should be considered extremely preliminary since no statistical tests (e.g. bootstrap analyses) have yet been conducted.

18S rDNA sequences from 223 angiosperms representing all major lineages were obtained (Soltis *et al.*, submitted) and aligned. The following holoparasites were then added: *Balanophora fungosa* J. R. & G. Forst., *Corynaea crassa* Hook. f.*, Cynomorium coccineum* Mich.*, Helosis cayennensis* (Schwartz) Spreng.*, Ombrophytum subterraneum* (Asplund) B. Hansen, *Rhopalocnemis phalloides* Jungh., and *Scybalium jaimaicense* (Sw.) Schott. & Endl. (Balanophoraceae), *Cytinus ruber* Fritsch, *Rafflesia keithii* Meijer, and *Rhizanthes zippelii* (Blume) Spach (Rafflesiaceae s. lat.), and *Hydnora africana* Thunb. and *Prosopanche americana* (R. Br.) Baillon (Hydnoraceae) for a total of 245 sequences. Given the size of this dataset, only heuristic search strategies could be used. Since branch swapping did not go to completion, the tree discussed below is not optimal (minimum length).

The tree resulting from this analysis retained the major topological features found in the more extensive analysis reported in Soltis *et al.* (submitted). For example, the clade composed of Santalales (strict sense) was allied with Polygalaceae and Fabaceae within a large group composed of Rosid and Dilleniid taxa. The placement of Santalales as sister to Polygalales is of interest since parasitism has also evolved in one family of the latter (Krameriaceae). None of the holoparasite families (Balanophoraceae, Hydnoraceae, and Rafflesiaceae) were associated with the santalalean clade. The two genera of Hydnoraceae, *Prosopanche* and *Hydnora*, were placed as sister taxa in a clade composed of several families referred to as paleoherbs (sensu Donoghue and Doyle, 1989), i.e. Aristolochiaceae, Chloranthaceae, Lactoridaceae, Piperaceae, and Saruraceae. Evidence from several molecular and morphological datasets supports the basal position of the paleoherbs within the angiosperms. It is of interest that the relationship between Hydnoraceae and Aristolochiaceae is fully concordant with the concepts proposed by Solms-Laubach (1894) and Harms (1935). More recently, Cocucci (1983) used embryological and floral characters to derive Aristolochiaceae, Rafflesiaceae, and Hydnoraceae from annonaceous stock. The molecular analyses reported here provide additional evidence linking Hydnoraceae to magnoliid ancestors.

When analyzed together, Rafflesiaceae and Balanophoraceae exhibit artifactual long-branch attractions. The clade composed of *Rafflesia* and *Rhizanthes* is intercalated in the middle of Balanophoraceae, an unlikely relationship. In this analysis, this composite clade was placed with taxa allied with Saxifragales, an equally unlikely relationship. The 18S rDNA analysis does not support a relationship between *Cytinus* and *Rafflesia/Rhizanthes*. These taxa

were separated at the tribal level by Harms (1935) and, on the basis of pollen morphology, *Cytinus* has been segregated to Cytinaceae by Takhtajan *et al.* (1985). The classification of *Cytinus* in its own family is concordant with results from these molecular analyses that place it in Rosidae. These results indicate that Rafflesiaceae *sensu latu* is a paraphyletic group composed of at least two and possibly three distinct families. Preliminary molecular analyses of 18S rDNA sequences indicates that the small-flowered Rafflesiaceae (*Apodanthes, Bdallophyton, Mitrastemma*, and *Pilostyles* - *Berlinianche* yet to be sampled) are not closely related to the large-flowered members (*Rafflesia* and *Rhizanthes* - *Sapria* yet to be sampled). Since 18S rDNA analyses strongly suggest placement of Hydnoraceae in Magnoliidae, and since all the major angiosperm classifications have placed Rafflesiaceae and Hydnoraceae together in Rafflesiales, it is tempting to follow tradition and place Rafflesiaceae in Magnoliidae, despite lack of direct support from molecular data. The position of *Cytinus* in Rosidae counteracts this inclination since this family may provide a link to the more derived members of the family. All systematic discussion must be considered speculative at this point since the extreme rate heterogeneity displayed by these plants is clearly affecting their position relative to slower-rate plants.

Despite long-branch attractions, relationships among the seven genera of Balanophoraceae can be addressed. When Balanophoraceae are analyzed without Rafflesiaceae, using the gnetophytes *Ephedra* and *Gnetum* as outgroups, the phylogram shown in Fig. 4 results. One clade is composed of the New World genera *Corynaea*, *Helosis*, *Ombrophytum* and *Scybalium* and the other clade of the Old World genera *Balanophora, Rhopalocnemis*, and *Langsdorffia*. *Langsdorffia* occurs basal to the Old World branch, a position concordant with biogeographic evidence since it is the only genus in the family present in both the Old and New World (Central and South America and Madagascar). These relationships are partly in accord with the classification of Harms (1935) who proposed six subfamilies. Contrary to that classification, where the Indonesian *Rhopalocnemis* was placed with the New World genera in Helosidoideae, rDNA sequence data show it is very closely related to *Balanophora*. The most distantly related component of this clade is *Cynomorium*. Given the amount and types of mutations present in 18S rDNA from this genus and other Balanophoraceae indicate it is distinct. For this reason, segregation at the family level as proposed by Tahktajan (1987) and Thorne (1992) is favored.

Molecular Phylogenetic Studies of Scrophulariales

Within Scrophulariales, two families are traditionally defined that contain parasitic members: Scrophulariaceae and Orobanchaceae. These families contain greater than 45 genera and 1650 parasitic species and also exhibit the entire range of trophic conditions found in parasitic plants (hemiparasites to holoparasites). The traditional concept maintaining these two families as separate is not being followed in recent morphologically-based (e.g. Takhtajan, 1987; Minkin and Eshbaugh, 1989; Thorne, 1992; Judd *et al.*, 1994) and molecularly-based (e.g. Olmstead and Reeves, 1995; dePamphilis, 1995) classifications. The study by Olmstead and Reeves (1995) utilized sequences from the chloroplast genes *rbc*L and *ndh*F (separately and in combination); however, they did not include any members of the hemiparasitic rhinanthoid Scrophulariaceae or Orobanchaceae. Wolfe and dePamphilis (1995) have generated over 35 new *rbc*L sequences for these families and have discovered that intact *rbc*L genes are present in some nonphotosynthetic species. This contrasts with the condition in *Epifagus* (Orobanchaceae) where *rbc*L is present as a pseudogene (434 bp vs. 1452 bp in length). They also observed that pseudogene formation can occur independently within a genus (*Orobanche*).

A second chloroplast gene, *rps*2 (ribosomal small-subunit 2) has been sequenced and analyzed by dePamphilis (1995). Sampling in this study included parasitic and nonparasitic Scrophulariaceae, Orobanchaceae and several additional taxa from Scrophulariales. Results from this analysis showed that: 1) Orobanchaceae is derived from within Scrophulariaceae making the latter family paraphyletic, 2) taxa traditionally placed in other families are related to Scrophulariaceae (e.g. Bignoniaceae, Verbenaceae), and 3) parasitism arose just once within this lineage whereas loss of photosynthesis has occurred on many independent occasions. The *rps*2 tree was not fully resolved, however, some relationships received strong support such as a clade containing *Epifagus, Conopholis* and *Boschniakia*, a second with *Harveya* and *Hyobanche*, a third with two species of *Orobanche*, and a fourth with *Pedicularis* and *Castilleja*. *Lathraea* is the last member to emerge from this large polytomy.

Colwell (1994) conducted phylogenetic analyses of hemiparasitic Scrophulariaceae and holoparasitic Orobanchaceae using nuclear 18S rDNA sequences. Although this molecule usually provides insufficient characters for analysis at the rank of family and below, increased rates of nucleotide substitution in Orobanchaceae provided enough variation to address relationships at this level. Colwell (1994) showed results of analyses of nine ingroup (parasitic) plants, however, subsequent sequencing (Colwell and Nickrent, unpublished)

has increased this number to 18. Phylogenetic analysis was conducted using 18 ingroup 18S rDNA sequences (Scrophulariaceae and Orobanchaceae) and three outgroup sequences (*Glycine, Lycopersicon* and *Ipomoea*). Two equally parsimonious minimum-length trees were obtained and the consensus is shown in Fig. 5. A clade containing *Linaria* and *Chamophila* is sister to the remaining Scrophulariaceae plus Orobanchaceae suggesting (as with *rps2*, above) that parasitism arose just once. The topology of this tree is fully concordant with that obtained with *rps2* and in fact provides further resolution of relationships. The hemiparasites *Pedicularis, Orthocarpus* and *Castilleja* form a clade separate from the holoparasites. Three genera have been placed in either Scrophulariaceae or Orobanchaceae by different taxonomists: *Harveya, Hyobanche* and *Lathraea*. This 18S rDNA analysis shows these three genera to be components of a clade containing three species of *Orobanche*. Although taxon sampling is still incomplete, it appears that members of this clade are undergoing accelerated rates of nucleotide substitution compared with their relatives. The 18S tree also indicates that the genus *Orobanche* is monophyletic, yet it is not closely related to *Epifagus, Conopholis* and *Boschniakia* thus making Orobanchaceae paraphyletic. This study provides evidence that 18S rDNA can be used to address phylogenetic questions in Orobanchaceae. Of the 15 "nontransitional" genera classified in this family, only four have been sampled, hence the inclusion of more Old World representatives is needed to fully address the remaining questions.

General Features of Parasite Plastid Genomes

Among the three distinct genomes in plant cells, the circular plastid genome is certainly the best characterized in terms of structure and function (Sugiura, 1992). Green plants have from one to several hundred chloroplasts per cell and each chloroplast may contain from 7 to greater than 200 plastid genomes (Maguire *et al.*, 1995). The complete chloroplast DNA molecule has been sequenced for *Marchantia* (Ohyama *et al.*, 1986), *Pinus* (Wakasugi *et al.*, 1994), *Oryza* (Hiratsuka *et al.*, 1989), *Nicotiana* (Shinozaki *et al.*, 1986) and *Epifagus* (dePamphilis and Palmer, 1990). The cpDNAs of these plants vary widely in size: 119 kb in pine, 121 kb in liverwort, 134 kb in rice, and 156 kb in tobacco. The genome of *Epifagus virginiana* (L.) Bart. (beechdrops) is only 71 kb owing to extensive losses of photosynthetic genes that have accompanied its loss of photosynthesis (dePamphilis and Palmer, 1990). *Epifagus* has a large single copy region (LSC) of 18 kb (vs. 87 kb in tobacco), a small single copy region (SSC) of 3.6 kb (vs. 18.5 kb) that contains only two genes (Wolfe, *et al.*, 1992), but strangely the inverted repeat regions are

present and full sized (25 kb). Of the 42 genes remaining in *Epifagus*, 38 are involved in protein synthesis (e.g. rRNA, tRNA, ribosomal protein genes), yet expression relies upon import of nuclear-encoded tRNAs and RNA polymerase (Morden *et al.*, 1991). The selective maintenance of these specific genes (versus loss of nearly all photosynthetic genes) provided inferential evidence of their functionality, but direct evidence was obtained by Ems *et al.* (1995) using Northern blot analyses. *Conopholis americana* (L.) Wallr. (squawroot), another holoparasite in Orobanchaceae, has also been the subject of molecular studies. Using heterologous probes, Wimpee *et al.* (1991) documented the modification or absence of many photosynthetic genes. Colwell (1994) conducted restriction site mapping of the plastid genome of squawroot documenting its size as 43 kb, i.e. the smallest ptDNA molecule yet observed in plants. Much of this reduction is due to the loss of one copy of the inverted repeat.

In addition to parasitic Scrophulariaceae and Orobanchaceae, molecular genetic investigations have also been focused upon *Cuscuta*. This genus, sometimes placed in its own family (Cuscutaceae), is widely recognized to share a common ancestor with nonparasitic Convolvulaceae. As shown in Fig. 1, *Cuscuta* (like Scrophulariaceae), includes hemiparasitic species (e.g. *C. reflexa* Roxb.) as well as holoparasitic species (e.g. *C. europaea* L.) that lack thylakoids, chlorophyll, Rubisco and light-dependent CO_2 fixation but (strangely) retain *rbc*L (Machado and Zetsche, 1990). Although the plastid genome of *Cuscuta* is yet to be fully sequenced, significant progress is being made. Bömmer *et al.* (1993). cloned and sequenced a 9 kb portion of ptDNA from *C. reflexa* that includes 16S rDNA, *psb*A, *trn*H, ORF 740, ORF 77, *trn*L, and ORF 55. Later, Haberhausen and Zetsche (1994) cloned and sequenced a 9 kb portion of ptDNA from this same species that included a large portion of inverted repeat A spanning a segment from *trn*A to *trn*H. Although some sequences were identical to *Nicotiana* (e.g. *trn*I), many deletions were observed. For example, *rpl*2 and *rpl*23 were both deleted and ORF2280 was reduced to only 740 bp. These results show that, like *Epifagus*, *Cuscuta* has experienced major deletions in the plastid genome. The complete loss of ribosomal protein genes such as *rpl*2 invites questions about how such the translational apparatus of the plastid functions and its relationship to the other two subcellular genomes.

Plastid rRNA in Parasitic Plants

In plants, the ribosomal cistrons are present on both inverted repeats (when present) and are composed of genes in the following order: 16S, *trn*I, *trn*A, 23S, 4.5S, and 5S (Fig. 2B).

This arrangement is extremely conserved, yet some variation can be seen in Orobanchaceae. *Conopholis* retains an intact 16S and 23S rDNA, however, the intervening *trn*I is absent and *trn*A has become a pseudogene (Wimpee *et al.*, 1992). The flanking 4.5S and 5S rDNAs are intact, but the spacers between them are only 70% of the typical length. *Epifagus* also retains intact 16S and 23S rDNA, but both *trn*I and *trn*A have become pseudogenes while conserving the length of the entire cistron (Wolfe *et al.*, 1992).

Prior to 1995, there were no published data available regarding the presence of a plastid genome (or genes) in Balanophoraceae, Hydnoraceae or Rafflesiaceae. Despite experiencing extensive losses in photosynthetic and other genes, both *Epifagus* and *Conopholis* have intact and functional ribosomal cistrons. Given this, it was reasoned that if any genes were present in these holoparasites, very likely they would be rDNA. The goal of these initial studies was to use PCR to amplify 16S rDNA from each of these lineages.

A multiple sequence alignment was conducted on 16S rDNA using cyanobacterial outgroups as well as published plant sequences. Conserved regions were identified and oligonucleotide primers developed for PCR and sequencing. PCR was conducted using the 8 forward [GGA GAG TTC GAT CCT GGC TCA G] and the 1461 reverse [GGT GAT CCA GCC GCA CCT TCC AG] primers (numbers based upon position on tobacco). To date, 16S rDNA has been amplified from genomic DNA samples of representatives of Balanophoraceae, Hydnoraceae and Rafflesiaceae (Nickrent *et al.*, 1995). We have also amplified 23S rDNA and the spacer region between the 16S and 23S rDNA in several taxa indicating that a large portion (if not a full) ribosomal cistron is present. Although compelling, these data do not yet prove the existence of a plastid genome since the possibility of migration to nucleus has not been excluded.

A secondary structure of the 16S rRNA of *Cytinus ruber* (Rafflesiaceae or Cytinaceae) is presented in Fig. 6 based upon the model of Gutell *et al.* (1985) for *Zea*. This model corresponds well to previously proposed ones for cyanobacteria and green plant plastid rRNAs as evidenced by conservation of all major structural features. Mutations relative to the 16S rRNA of *Nicotiana* are indicated by highlighted bases. Overall, 130 single base change mutations (8.6%) can be seen on this molecule, i.e. greater than twice as many as would be seen in a comparison of *Nicotiana* to the more distantly related liverwort *Marchantia*. Similar (or higher) mutational levels were observed for the other holoparasite 16S sequences. Given the large number of changes in the *Cytinus* sequence, it is

possible that this sequence is that of a pseudogene. To investigate the effect of each type of mutation on the maintenance of helices, mutations were identified and categorized. Fully compensated changes occur when each paired base in a helix changes but canonical or noncanonical base pairing is retained. Fully compensated changes can be found in 12 pairs representing 24 mutational events. Compensated changes, when one member of a base pair in a helix changes, yet canonical or noncanonical pairing is maintained (e.g. G–C → G • U), was observed at 37 sites. Changes in unpaired regions (loops and bulges) were the largest class of mutation, constituting 60 changes. Finally, eight changes were found that disrupted a base pair present in the *Nicotiana* 16S rRNA secondary structure. This exercise shows that, despite frequent mutation, the majority do not result in obvious disruptions of the secondary structure. The conclusion is that the 16S rRNA in *Cytinus* is very likely functional, although Northern blot and/or *in vitro* studies are required for additional confirmation.

Plastid 16S rDNA sequences have been used infrequently for inferring phylogenetic relationships in plants, owing mainly to the paucity of sufficient numbers of sequences and the overall slow evolutionary rate of this gene. The relationships among seven dicots and two monocots using 16S rRNA sequences was shown in a phylogram reported by Wolfe *et al.* (1992). That study documented the overall conservative nature of this molecule and the higher nucleotide substitution rates in *Conopholis* and *Epifagus* relative to other dicots.

The following sequences (with Genbank accession numbers) were used in this phylogenetic analysis: *Anacystis nidulans* [X03538], *Anabaena* sp. [X59559], *Microcystis aeruginosa* [U03402], *Marchantia polymorpha* [X04465], *Pinus thunbergii* [D17510], *Oryza sativa* [X15901], *Zea mays* [X86563], *Pisum sativum* [X51598], *Glycine max* [X06428], *Daucus carota* [X73670], *Cuscuta reflexa* [X72584], *Nicotiana tabacum* [Z00044], *Epifagus virginiana* [X62099], and *Conopholis americana* [X58864]. The sequence of *Brassica* and *Vicia* were obtained from Gao *et al.* (1989, 1990, respectively). The 16S sequence of *Cytinus ruber* is available via Genbank number U47845. The above 16S rDNA sequences were aligned and subjected to cladistic analysis using PAUP (Swofford 1993). Alignment was generally unambigous and secondary structural features were used to guide alignment in variable regions.

The branch and bound algorithm found a single most parsimonious minimum length tree of 808 steps. To test the statistical stability of the clades, bootstrap analysis (100 replications using a heuristic search strategy) was conducted and the resulting tree is shown in Fig. 7.

The clade containing the two cyanobacteria (*Anacystis* and *Microcystis*) is separated from the green plant (chloroplast) clade by 104 substitutions. As expected, the angiosperms form a monophyletic clade, separate from the liverwort and pine, with an 83% bootstrap value. Within angiosperms, a clade containing Asteridae (plus *Brassica*) is recovered in 79% of the bootstrap trees. The bootstrap value for the legume clade is 70% whereas the monocots (plus *Cytinus*) is 77%. These two clades, however, are recovered in less than 50% of the bootstrap trees and therefore form a polytomy with the Asteridae clade (above). An unusual feature of the tree is the position of *Cytinus* with the monocots - clearly a case of long-branch attraction. The branch leading to *Cytinus* has 102 substitutions, i.e. as great as the branch separating the chloroplast clade from the cyanobacteria.

Mapped onto the phylogram shown in Fig. 7 are instances where the plastid inverted repeat has been lost. This has occurred independently in several lineages, i.e. *Conopholis*, two legumes (*Pisum* and *Vicia* but not *Glycine*) and very likely *Cytinus*. It is of interest to note that in the legumes, 47 and 34 mutations have occurred since the divergence of *Vicia* and *Pisum* (respectively) from their common ancestor with *Glycine*. These numbers are similar to values obtained when deriving *Epifagus* and *Conopholis* from their common ancestor. Strangely, only *Conopholis* has lost the inverted repeat, not *Epifagus*. It is possible that the presence of the inverted repeat imparts a stabilizing effect on the dynamics of plastid evolution and that increased rates are seen when this effect is removed. The presence of similar numbers of substitutions for *Conopholis* and *Epifagus* may indicate that insufficient time has elapsed for the appearance of major differences.

Molecular Evolutionary Studies - Relative Rate Tests

The molecular phylogenetic approach to reconstructing evolutionary relationships is closely tied to the concept of a molecular clock (Zuckerkandl and Pauling, 1965). The strict molecular clock infers that for organisms with equal generation times, their respective lineages will have equal rates of nucleotide substitution. For plants, a strict molecular clock cannot be applied universally to all clades and to all gene loci. The generation–time hypothesis is the most frequently cited cause for rate invariance (Bosquet *et al.*, 1992); however, this explanation is often difficult to apply to plants and may be only partially responsible for the phenomenon (Nickrent and Starr, 1994).

The most dramatic cases of rate invariance in plants have been detected following relative rate tests using nuclear 18S rDNA sequences from parasitic plants. Five parasitic flowering plants showed unusually high substitution rates, yet these lineages do not have fast

generation times (Nickrent and Starr 1994). It was hypothesized that the underlying cause of such divergent rDNA sequences was relaxation of selectional constraints on rRNA structure and function, possibly as a result of small effective population size and molecular drive. These processes, however, are occurring in other eukaryotes that do not show increased substitution rates, hence they do not provide a universal explanation. Accelerated substitution rates are not restricted to a single parasitic plant clade (cf. Viscaceae, Balanophoraceae, Rafflesiaceae) nor are they found universally within a single clade (cf. Viscaceae with Loranthaceae). It appears that rate acceleration of nuclear rDNA occurs only in lineages that exhibit a reduction in photosynthesis and an advanced state of nutritional dependence upon the host. It is not presently understood why these changes in life history are correlated with changes at nuclear rRNA loci. Studies are currently underway (Colwell and Nickrent, in prep.) comparing evolutionary rates of 18S rDNA in all the major parasite and mycotroph lineages.

The 16S rDNA phylogenetic tree shown in Fig. 7 indicates (by means of branch lengths) the highly diverged sequence of *Cytinus*. The sequences used in this analysis were subjected to formal relative rate tests (Wu and Li, 1985) as described in Nickrent and Starr (1994). The purpose of the test is to determine whether the rates of substitution are the same or different between two lineages - here taxon 1 and 2 (see Fig. 8). A hypothetical common ancestor of taxa 1 and 2 is designated X. The actual number of substitutions (transitions and transversions) on the branch from X to 1 (K_1) cannot be determined directly but can be calculated assuming that the time since divergence from X to taxon 1 and 2 is the same. The null hypothesis is that the number of substitutions from taxon 1 to taxon 3 (K_{13}) is equal to that from taxon 2 to 3 (K_{23}), or $K_{13} - K_{23} = 0$. A significant deviation of this value from 0 indicates that the two lineages are evolving at different rates. Generally, when $K_{13} - K_{23}$ / standard error > 3, the probability of rejecting the null hypothesis is less than 1%. Taxon 3 (the reference) must be an unambiguous outgroup to taxa 1 and 2 for this test to be meaningful. For this reason, *Pinus* was chosen as the reference since the topology shown in Fig. 7 does not place *Cytinus* with the dicots.

The results of the relative rate tests shown in Fig. 8 clearly indicate the gross rate asymmetry found in the branch leading to *Cytinus*. In the study by Wolfe *et al.* (1992), most substitutions were seen along the *Epifagus* and *Conopholis* branch, hence the authors concluded that the rate of substitution was higher in these holoparasites. The formal

relative rate tests reported here do not show these two plants to exhibit any greater rate asymmetry than in comparisons of other plants (such as *Vicia*).

CONCLUSIONS

This paper has focused upon the utility of nuclear and plastid ribosomal RNA in molecular phylogenetic and molecular evolutionary studies of parasitic angiosperms. The use of nuclear 18S rDNA in examining the phylogeny of the component families and genera of Santalales has been demonstrated. This molecular phylogeny clearly indicates the progressive evolution from root parasitic (and nonparasitic) Olacaceae to the most advanced clade, the aerially parasitic Viscaceae. Mistletoes are represented by at least four, independently evolved groups (Viscaceae, Loranthaceae, Misodendraceae, and Santalaceae/Eremolepidaceae). Data from nuclear-encoded rDNA and plastid-encoded *rbc*L indicate that *Antidaphne* and *Eubrachion* (Eremolepidaceae) are best classified within Santalaceae. The difficulty in placing the holoparasite families Balanophoraceae, Hydnoraceae, and Rafflesiaceae within the overall classification of angiosperms can be attributed to their extremely derived vegetative and floral features combined with very high rates of molecular evolution. Preliminary evidence from 18S rDNA analyses place Hydnoraceae with the paleoherbs (Magnoliidae).

Nuclear 18S rDNA sequence analysis is in agreement with recent classifications that recognize Orobanchaceae as a component of Scrophulariaceae. 18S rDNA data are concordant with analyses of *rps*2 in that both recover a monophyletic group composed of rhinanthoid Scrophulariaceae and Orobanchaceae, thus indicating the unique origin of parasitism in the family. Genera traditionally considered transitional between Scrophulariaceae and Orobanchaceae (*Lathraea, Harveya* and *Hyobanche*) emerged as components of the *Orobanche* clade. The utility of 18S rDNA sequences in addressing subfamilial phylogenetic relationships has been demonstrated but not fully realized owing to incomplete sampling.

Holoparasitic plants (such as *Epifagus* and *Conopholis*) have undergone extreme reorganization of the plastid genome, mainly manifested by the loss of photosynthetic genes. Sequences from plastid rDNA from representatives of Balanophoraceae, Hydnoraceae, and Rafflesiaceae have been obtained thereby suggesting that these plants have also retained a plastid genome. These plants have the most divergent 16S rRNAs ever documented, yet structural studies provide evidence that these molecules are still

functional. Extreme rate heterogeneity, as compared with other vascular plants, was demonstrated for 16S rDNA in *Cytinus ruber* using relative rate tests. This amount of change at highly conserved ribosomal loci provides unprecedented opportunities to study the molecular evolution of the plastid and nuclear genomes. For example, the study of these "fast rate" rRNAs can provide general insight into the structure and function of all rRNA molecules by providing compensatory mutations in regions previously thought to be invariant. Unlike *Orobanche* or *Striga*, which are the topic of many papers at this symposium, many of the plants discussed here are rare and endangered. For this reason, they are worthy of conservation efforts and should be viewed as valuable model organisms that can increase our understanding of the mode and tempo of evolutionary change at the molecular level.

ACKNOWLEDGMENTS

The laboratory assistance of Joseph Agan, Jamie Beam, Mark Berning, Brent Beilschmidt, Erica Grimm, Jennifer Patrick, and Ellen Starr is gratefully acknowledged. Special thanks go to Alison Colwell who provided several unpublished 18S sequences of Scrophulariaceae. Sincere thanks to the colleagues who collected samples of the myriad parasite species reported herein. We could not have achieved a worldwide perspective on the evolution of these plants without the assistance of the following people: Carol Augspurger, Wilhelm Barthlott, Donaldo E. Bran, Sherwin Carlquist, Claude dePamphilis, James Fetzner, Wolfgang Forstreuter, Jorge Gomez, Gregori Hambali, Danny Joel, Linda Mabasa, James Mauseth, Scott Medbury, Mike Melampy, Willem Meijer, Brian Molloy, Lytton Musselman, R. Narayana, Jack Paxton, Roger Polhill, Sarah Sargent, Robert Scharpf, Kim Steiner, S. A. Robertson, Wayne Takeuchi, Steven Teo, and James Trice. Support for this project was provided by grants from the National Science Foundation (BSR-89-18385 and DEB 94-07984) and the Special Research Program of the Office of Research Development and Administration, SIUC.

REFERENCES

Appels, R. and Honeycutt, R.L. (1986) rDNA: Evolution over a billion years. In: S.K. Dutta (ed.) DNA systematics, Vol. II, Plants. CRC Press, Boca Raton, Florida. pp. 81-135.

Arnheim, N. (1983) Concerted evolution of multigene families. In: M. Nei and R. K. Koehn (eds.), Evolution of Genes and Gene Proteins. Sinauer Associates. Sunderland, MA. pp. 38-61.

Baldwin, B.G., Sanderson, M.J., Porter, J.M., Wojciechowski, M.F., Campbell, C.S. and Donoghue, M.J. (1995) The ITS region of nuclear ribosomal DNA: A valuable source of evidence on angiosperm phylogeny. Annals of the Missouri Botanical Garden 82:247-277.

Barlow, B.A. (1983) Biogeography of Loranthaceae and Viscaceae. In: D.M. Calder and P. Bernhardt (eds.) The Biology of Mistletoes. Academic Press, New York. pp. 19-45.

Bhandari, N.N. and Vohra, S.C.A. (1983) Embryology and Affinities of Viscaceae. In: M. Calder and P. Bernhardt (eds.) The Biology of Mistletoes. Academic Press, New York. pp. 69-86.

Bosquet, J., Strauss, S.H., Doerksen, A.H. and Price, R.A. (1992) Extensive variation in evolutionary rate of *rbc*L gene sequences among seed plants. Proceedings of the National Academy of Sciences U.S.A. 89:7844-7848.

Brown, D.D., Wensink, P.C. and Jordan, E. (1972) A comparison of the ribosomal DNA's of *Xenopus laevis* and *Xenopus mulleri*: The evolution of tandem genes. Journal of Molecular Biology 63:57-73.

Bömmer, D., Haberhausen, G. and Zetsche, K. (1993) A large deletion in the plastid DNA of the holoparasitic flowering plant *Cuscuta reflexa* concerning two ribosomal proteins (*rpl*2, *rpl*23), one transfer RNA (*trn*I) and an ORF 2280 homologue. Current Genetics 24:171-176.

Cedergren, R., Gray, M.W., Abel, Y. and Sankoff, D. (1988) The evolutionary relationships among known life forms. Journal of Molecular Evolution 28:98-112.

Chase, M.W., Soltis, D.E., Olmstead, R.G., Morgan, D., Les, D.H., Mishler, B.D., Duvall, M.R., Price, R.A., Hills, H.G., Qiu, Y.-L., Kron, K.A., Rettig, J.H., Conti, E., Palmer, J.D., Manhart, J.R., Sytsma, K.J., Michaels, H.J., Kress, W.J., Karol, K.G., Clark, W.D., Hedrén, M., Gaut, B.S., Jansen, R.K., Kim, K.-J., Wimpee, C.F., Smith, J.F., Furnier, G.R., Strauss, S.H. , Xiang, Q.-Y., Plunkett, G.M., Soltis, P.S., Swensen, S.M., Williams, S.E., Gadek, P.A., Quinn, C.J., Eguiarte, L.E., Golenberg, E., Learn,

G.H., Jr., Graham, S.W., Barrett, S.C.H., Dayanandan, S. and Albert, V.A. (1993) Phylogenetics of seed plants: An analysis of nucleotide sequences from the plastid gene *rbc*L. Annals of the Missouri Botanical Garden 80:528-580.

Cocucci, A.E. (1983) New evidence from embryology in angiosperm classification. Nordic Journal of Botany 3:67-73.

Colwell, A.E. (1994) Genome evolution in a non-photosynthetic plant, *Conopholis americana*. Ph.D. Dissertation, Washington University, St. Louis, Missouri, U.S.A.. 222 pp.

Cronquist, A. (1981) An integrated system of classification of flowering plants. Columbia Univ. Press, N.Y.

Cronquist, A. (1988) The evolution and classification of flowering plants. New York Botanical Gardens, N.Y.

Danser, B.H. (1933) A new system for the genera of Loranthaceae-Loranthoideae, with a nomenclator for the Old World species of this subfamily. Verhandelingen der Koninklijke Akademie van Wetenschappen te Amsterdam Afdeeling Natuurkunde Sect. 2, 29:1-128.

dePamphilis, C.W. and Palmer, J.D. (1990) Loss of photosynthetic and chlororespiratory genes from the plastid genome of a parasitic flowering plant. Nature (London) 348:337-339.

dePamphilis, C.W. (1995) Genes and genomes. In: M.C. Press and J.D. Graves (eds.) Parasitic Plants. Chapman and Hall, London. pp. 176-205.

Donoghue, M.J. and Doyle, J.A. (1989) Phylogenetic analysis of angiosperms and the relationships of Hamamelidae. In: P.R. Crane and S. Blackmore (eds.) Evolution, Systematics, and Fossil History of the Hamamelidae. Clarendon Press, Oxford. pp. 17-45.

Dover, G.A. (1982) Molecular drive: a cohesive mode of species evolution. Nature 299:111-117.

Dover, G.A. (1987) DNA turnover and the molecular clock. Journal of Molecular Evolution 26:47-58.

Ems, S.C., Morden, C.W., Dixon, C.K., Wolfe, K.H., dePamphilis, C.W. and Palmer, J.D. (1995) Transcription, splicing and editing of plastid RNAs in the nonphotosynthetic plant *Epifagus virginiana*. Plant Molecular Biology 29:721-733.

Fagerlind, F. (1948) Beitrage zur Kenntnis der Gynaceummorphologie und Phylogenie der Santalales-Familien. Svensk Botanisk Tidskrift. 42:195-229.

Felsenstein, J. (1978) Cases in which parsimony or compatibility will be positively misleading. Systematic Zoology 27:401-410.

Gao, J., Wang, X. and Wang, Q. (1989) The complete nucleotide sequence of *Brassica napus* chloroplast 16S rRNA gene. Acta Genetica Sinica 16:263-268.

Gao, J., Wang, X., Wang, Q., and Tan, J. (1990) Analysis of the primary structure of the leader sequence of the 16S rRNA gene from *Vicia faba*. Science in China, Series B. 33:592-598.

Gutell R.R. (1993) Collection of small subunit (16S- and 16S-like) ribosomal RNA structures. Nucleic Acids Research 21:3051-3054.

Haberhausen, G. and Zetsche, K. (1994) Functional loss of all ndh genes in an otherwise relatively unaltered plastid genome of the holoparasitic flowering plant *Cuscuta reflexa*. Plant Molecular Biology 24:217-222.

Hamby, K.R. and Zimmer, E.A. (1992) Ribosomal RNA as a phylogenetic tool in plant systematics. In: P. S. Soltis, D. E. Soltis, and J. J. Doyle (eds.) Molecular Systematics of Plants. Chapman and Hall, New York. pp. 50-91.

Harms, H. (1935) Rafflesiaceae, Hydnoraceae, and Balanophoraceae. In: A. Engler and H. Harms (eds.), Die Natürlichen Planzenfamilien, Part 16b. Leipzig. pp. 243-281; pp. 282-295; pp. 296-339.

Hiepko, P. (1979) A revision of Opiliaceae I. Genera of the eastern Old World, excluding *Opilia*. Willdenowia 9:13-56.

Hiepko, P. (1982) A revision of Opiliaceae II. *Opilia* Roxb. Willdenowia 12:161-182.

Hiratsuka, J., Shimada, H., Whittier, R., Ishibashi, T., Sakamoto, M., Mori, M., Kondo, C., Honji, Y., Sun, C. R., Meng, B. Y., Li, Y. Q., Kanno, A., Nishizawa, Y., Hirai, A., Shinozaki, K. and Sugiura, M. (1989) The complete sequence of the rice (*Oryza sativa*) chloroplast genome: intermolecular recombination between distinct tRNA genes accounts for a major plastid DNA inversion during the evolution of cereals. Molecular and General Genetics 217:185-194.

Johri, B.M. and Bhatnagar, S.P. (1960) Embryology and taxonomy of the Santalales I. Proceedings of the National Institute of Science, India, Part B., Biological Sciences 26:199-220.

Jorgansen, R.A. and Cluster, P.D. (1988) Modes and tempos in the evolution of nuclear ribosomal DNA: new characters for evolutionary studies and new markers for genetic and population studies. Annals of the Missouri Botanical Garden 75:1238-1247.

Judd, W.S., Sanders, R.W. and Donoghue, M.J. (1994) Angiosperm family pairs: preliminary phylogenetic analyses. Harvard Papers in Botany 5:1-51.

Kubat, R. (1987) Report of the first investigations of parasitism in Opiliaceae (Santalales). In: H. C. Weber and W. Forstreuter (eds.) Proceedings of the 4th. International Symposium on Parasitic Flowering Plants, Marburg, Germany. pp. 489-492.

Kuijt, J. (1968) Mutual affinities of Santalalean families. Brittonia 20: 136-147.

Kuijt, J. (1969) The biology of parasitic flowering plants. University of California Press, Berkeley.

Kuijt, J. (1988) Monograph of Eremolepidaceae. Systematic Botany Monographs 18:1-60.

Machado, M.A. and Zetsche, K. (1990) A stuctural, functional and molecular analysis of plastids of the holoparsites *Cuscuta reflexa* and *Cuscuta europaea*. Planta 181:91-96.

Maguire, M.J., Goff, L.J. and Coleman, A.W. (1995) In situ plastid and mitochondrial DNA determination: implication of the observed minimal plastid genome number. American Journal of Botany 82:1496-1506.

Minkin, J.P. and Eshbaugh, W.H. (1989) Pollen morphology of the Orobanchaceae and rhinanthoid Scrophulariaceae. Grana 28:1-18.

Morden, C.W., Wolfe, K.H., dePamphilis, C.W., and Palmer, J.D. (1991) Plastid translation and transcription genes in a nonphotosynthetic plant: intact, missing and pseudo genes. European Molecular Biology Organization Journal 10:3281-3288.

Nickrent, D.L. (1995) Molecular Systematics. In: F.G. Hawksworth and D. Wiens (eds.) Dwarf mistletoes: biology, pathology, and systematics. USDA Forest Service Agricultural Handbook 709, Washington DC. pp. 155-170

Nickrent, D.L. (1996) Malaysia's unique biological diversity: new insights from molecular evolutionary studies of parasitic flowering plants. In: M.A. Salleh and J. Beaman (eds.) Proceedings of the International Symposium & Workshop on Conservation Biology: Molecular, Biotechnological, and Conventional Approaches. University of Malaysia, Sarawak.

Nickrent, D.L. and Franchina, C.R. (1990) Phylogenetic relationships of the Santalales and relatives. Journal of Molecular Evolution 31:294-301.

Nickrent, D.L. and Soltis, D.E. (1995) A comparison of angiosperm phylogenies based upon complete 18S rDNA and *rbc*L sequences. Annals of the Missouri Botanical Garden 82:208-234.

Nickrent, D.L. and Starr, E.M. (1994) High Rates of nucleotide substitution in nuclear small-subunit (18S) rDNA from holoparasitic flowering plants. Journal of Molecular Evolution 39:62-70.

Nickrent, D.L., Ouyang, Y. and dePamphilis, C.W. (1995) Presence of plastid genes in representatives of the holoparasitic families Balanophoraceae, Hydnoraceae, and Rafflesiaceae. American Journal of Botany [Abstract] 82:75.

Nickrent, D.L., Schuette, K.P. and Starr, E.M. (1994) A molecular phylogeny of *Arceuthobium* based upon rDNA internal transcribed spacer sequences. American Journal of Botany 1149-1160.

Ohyama, K., Fukuzawa, H., Kohchi, T., Shirai, H., Sano, T., Sano, S., Umesono, K., Shiki, Y., Takeuchi, M., Chang, Z., Aota, S., Inokuchi, H. and Ozeki, H. (1986) Chloroplast gene organization deduced from complete sequence of liverwort *Marchantia polymorpha* chloroplast DNA. Nature (London) 322:572-574.

Olmstead, R.G. and Reeves, P.A. (1995) Evidence for the polyphyly of the Scrophulariaceae based on chloroplast *rbc*L and *ndh*F sequences. Annals of the Missouri Botanical Garden 82:176-193.

Pilger, R. (1935) Santalaceae. In: A. Engler and H. Harms (eds.) Die Natürlichen Pflanzen Familien, 2nd. Ed, Vol. 16b. Leipzig. pp. 52-91.

Reed, C.F. (1955) The comparative morphology of the Olacaceae, Opiliaceae, and Octoknemaceae. Memorias da Sociedade Broteriana 10: 29-79.

Sanderson, M.J. and Doyle, J.J. (1993) Reconstruction of organismal and gene phylogenies from data on multigene families: concerted evolution, homoplasy, and confidence. Systematic Biology 4:4-17.

Shinozaki, K., Ohme, M., Tanaka, M., Wakasugi, T., Hayashida, N., Matsubayashi, T., Zaita, N., Chunwongse, J., Obokata, J., Yamaguchi-Shinozaki, K., Ohto, C., Torazawa, K., Meng, B. Y., Sugita, M., Deno, H., Kamogashira, T., Yamada, K., Kusuda, J., Takaiwa, F., Kato, A., Tohdoh, N., Shimada, H. and Sugiura, M. (1986) The complete nucleotide sequence of the tobacco chloroplast genome. Plant Molecular Biology Reporter 4:110-147.

Sleumer, H.O. (1935) Olacaceae and Opiliaceae. In: A. Engler and H. Harms (eds.), Die Natürlichen Planzenfamilien, Part 16b. Leipzig. pp. 5-32; pp. 33-41.

Sleumer, H.O. (1984) Olacaceae. Monograph No. 38, Flora Neotropica, New York Botanical Garden, New York.

Sogin, M.L., Elwood, H.J. and Gunderson, J.H. (1986) Evolutionary diversity of eukaryotic small-subunit rRNA genes. Proceedings of the National Academy of Sciences U.S.A. 83:1383-1387.

Solms-Laubach, H. (1894) Hydnoraceae. In: A. Engler and K. Prantl (eds.), Die Natürlichen Planzenfamilien, Part III, Leipzig. pp. 282-285

Sugiura, M. (1992) The chloroplast genome. Plant Molecular Biology 19:149-168.

Swofford, D.L. (1993) PAUP: phylogenetic analysis using parsimony, version 3.1. Smithsonian Institution, Washington, D.C.

Takhtajan, A.L. (1980) Outline of the classification of flowering plants (Magnoliophyta). Botanical Review 46: 225-359.

Takhtajan, A. L. (1987) Sistema magnoliofitov. Nauka, Leningrad [in Russian].

Takhtajan, A.L., Meyer, N.R. and Kosenko, V.N. (1985) Morfologiya pyl'tsy i klassifikatsiya semeystva Rafflesiaceae s. l. Botanicheskii Zhurnal 70: 153-162.

Thorne, R.F. (1992) An updated phylogenetic classification of the flowering plants. Aliso 13:365-389.

Wakasugi, T., Tsudzuki, J., Ito, S., Nakashima, K., Tsudzuki, T. and Sugiura, M. (1994) Loss of all ndh genes as determined by sequencing the entire chloroplast genome of the black pine *Pinus thunbergii*. Proceedings of the National Academy of Sciences U.S.A. 91: 9794-9798.

Wiens, D. and Barlow, B.A. (1971) The cytogeography and relationships of the viscaceous and eremolepidaceous mistletoes. Taxon 20: 313-332.

Wimpee, C.F., Wrobel, R.L. and Garvin, D.K. (1991) A divergent plastid genome in *Conopholis americana*, an achlorophyllous parasitic plant. Plant Molecular Biology 17:161-166.

Wimpee, C.F., Morgan, R. and Wrobel, R.L. 1992. An aberrant plastid ribosomal RNA gene cluster in the root parasite *Conopholis americana*. Plant Molecular Biology 18:275-285.

Wolfe, A.D. and dePamphilis, C.W. (1995) Systematic implications of relaxed functional constraints on the RUBISCO large subunit in parasitic plants of the Scrophulariaceae and Orobanchaceae. American Journal of Botany [Abstract] 82:6.

Wolfe, K.H., Katz-Downie, D.S., Morden, C.W. and Palmer, J.D. (1992) Evolution of the plastid ribosomal RNA operon in a nongreen parasitic plant: accelerated sequence evolution, altered promoter structure, and tRNA pseudogenes. Plant Molecular Biology 18:1037-1048.

Wu, C.-I. and Li, W.-H. (1985) Evidence for higher rates of nucleotide substitution in rodents than in man. Proceedings of the National Academy of Sciences U.S.A. 82:1741-1745.

Zimmer, E.A., Martin, S.L., Beverley, S.M., Kan, Y.W. and Wilson, A.C. (1980) Rapid duplication and loss of genes coding for the a chains of hemoglobin. Proceedings of the National Academy of Sciences U.S.A. 77:2158-2162.

Zuckerkandl, E. and Pauling, L. (1965) Molecules as documents of evolutionary history. Journal of Theoretical Biology 8:357-366.

Figure 1. Occurrence of haustorial parasitism in angiosperms mapped upon the balloon phylogeny of subclasses proposed by Cronquist (1981, 1988). The presence of hemiparasitism is indicated by grey borders surrounding Santalales and Polygalales and holoparasitism by means of black borders. Both trophic conditions occur in Cuscutaceae and Scrophulariales. Concepts derived from results of molecular analyses (this paper) are incorporated, such as affinity of Hydnoraceae with Magnoliidae and inclusion of Eremolepidaceae in Santalaceae.

Figure 1

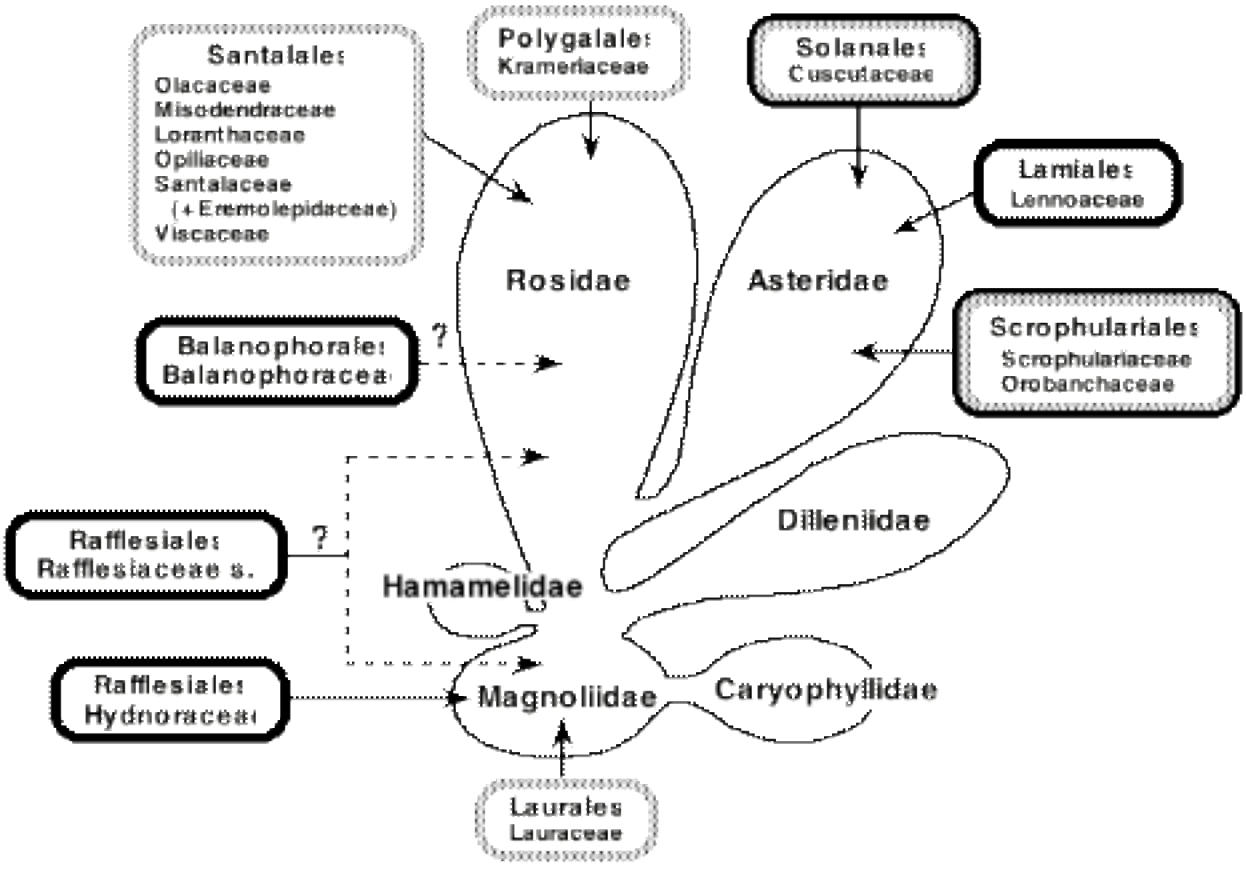


Figure 2. Comparison of plant nuclear and plastid ribosomal DNA cistrons. A. Nuclear rDNA cistron showing small-subunit (18S), large-subunit (26S), and 5.8S rDNAs. Also shown are the intergenic spacer (IGS), external transcribed spacer (ETS), nontranscribed spacer (NTS) and the internal transcribed spacers (ITS-1, -2). Variable domains are shaded (V1-V9 on 18S, D1-D12 on 26S). B. Plastid rDNA cistron as found in tobacco and most other higher plants. The spacer between the small-subunit (16S) and large-subunit (23S) rDNAs contains two tRNA genes (*trn*I_{GAU} and *trn*A_{UGC}), each containing introns.

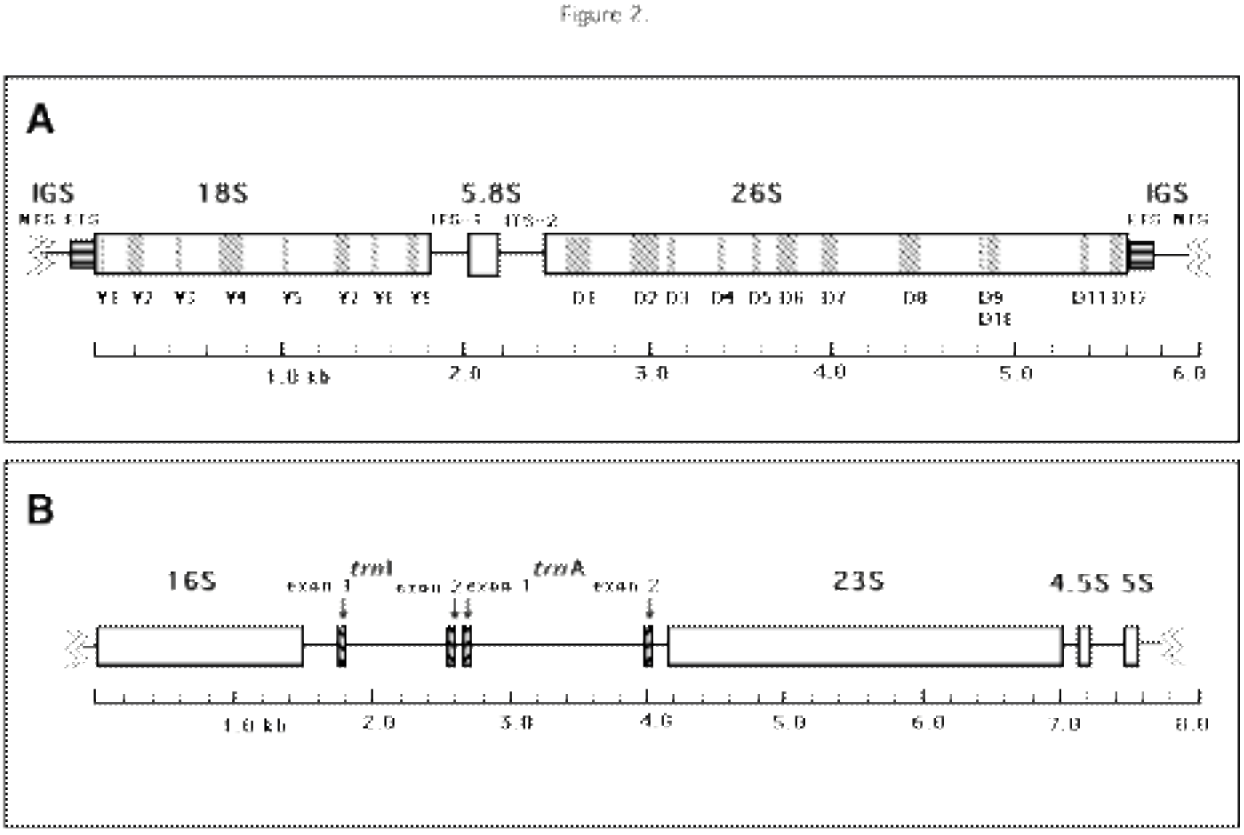


Figure 3. The strict consensus tree derived from a heuristic search with 62 18S rDNA sequences from representatives of Santalales. A total of 144 trees of length 1538 steps were found and are summarized by this consensus phylogram where branch lengths are indicated. The consistency index (minus uniformative sites) is 0.331 and the retention index is 0.610. All Olacaceae are found at the base of the tree with the exception of *Schoepfia* (⊕) which is sister to Misodendraceae within the Loranthaceae. Eremolepidaceae, represented by *Antidaphne* and *Eubrachion* (*) are components of Santalaceae (see text).

Figure 3

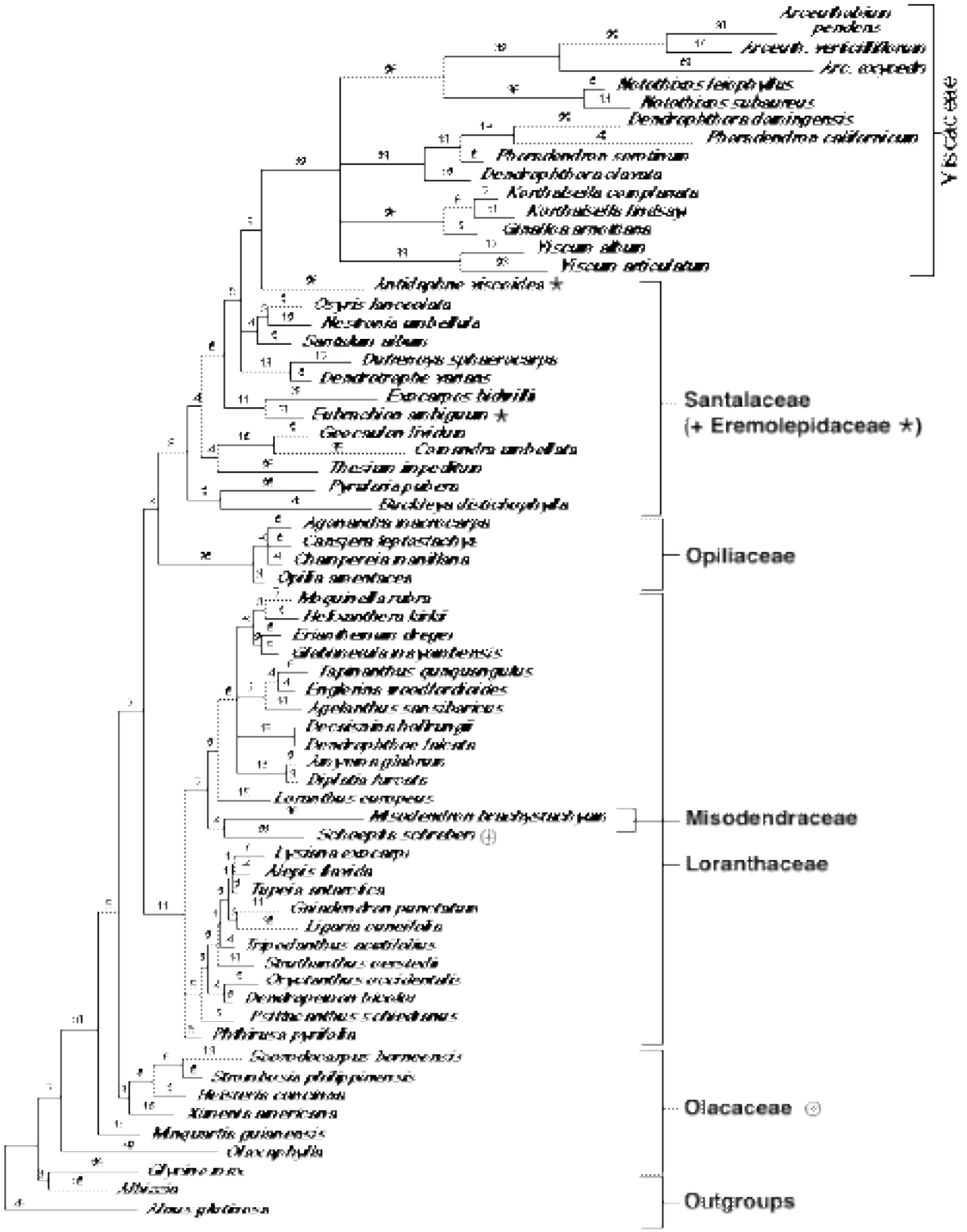

Arceuthobium pendens
Arceuth. verticilliflorum
Arc. oxycedri
Notothixos leiophyllus
Notothixos subaureus
Dendrophthora dominguensis
Phoradendron californicum
Phoradendron serotinum
Dendrophthora clavata
Korthalsella complanata
Korthalsella lindsayi
Ginalloa arnottiana
Viscum album
Viscum articulatum
Viscaceae
Antidaphne viscoidea ★
Osyris lanceolata
Nestronia umbellula
Santalum album
Dufrenoya spheroscarpa
Dendrotrophe varians
Exocarpos bidwillii
Eubrachion ambiguum ★
Geocaulon lividum
Comandra umbellata
Thesium impeditum
Pyrularia pubera
Buckleya distichophylla
Agonandra macrocarpa
Cansjera leptostachya
Champereia manillana
Opilia amentacea
Misodendraceae
Loranthaceae
Olacaceae ⊙
Outgroups
Santalaceae
(+ Eremolepidaceae ★)
Opiliaceae

Figure 4. The single tree of length 784 steps found via a branch and bound search of 18S rDNA sequences of Balanophoraceae (and *Cynomorium*). The consistency index (minus uniformative sites) is 0.685 and the retention index is 0.677. Number of steps are indicated above and bootstrap values (from 100 replications) are indicated below the branches.

Fig. 4

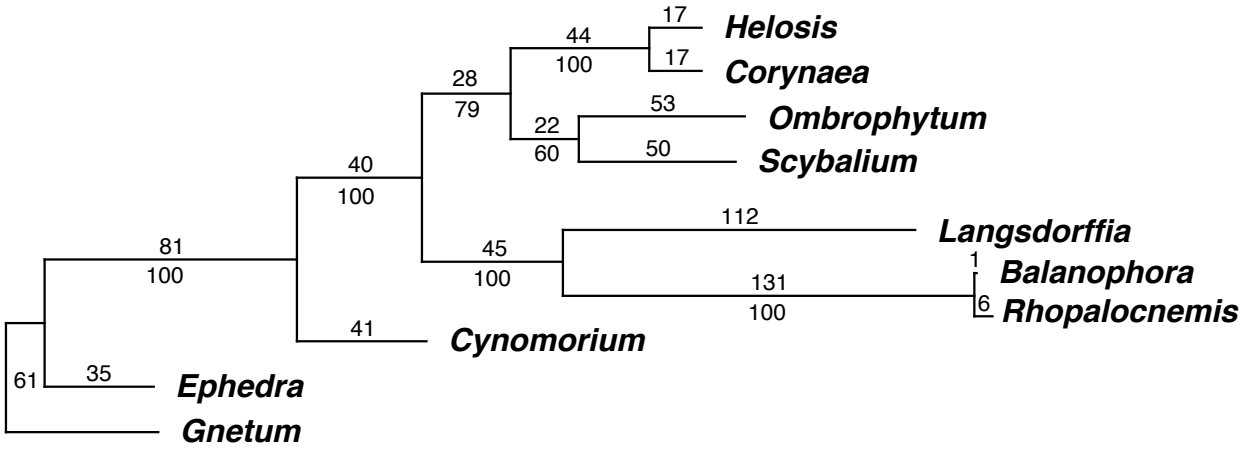


Figure 5. The strict consensus 18S rDNA tree derived from a branch and bound search with 18 sequences from representatives of Orobanchaceae and Scrophulariaceae. Four trees of length 246 steps were found and are summarized by this consensus phylogram. Number of steps are indicated above the branches. The consistency index (minus uniformative sites) is 0.553 and the retention index is 0.643. The evolution of parasitism marks a monophyletic group composed of rhinanthoid Scrophulariaceae and Orobanchaceae.

Figure 5

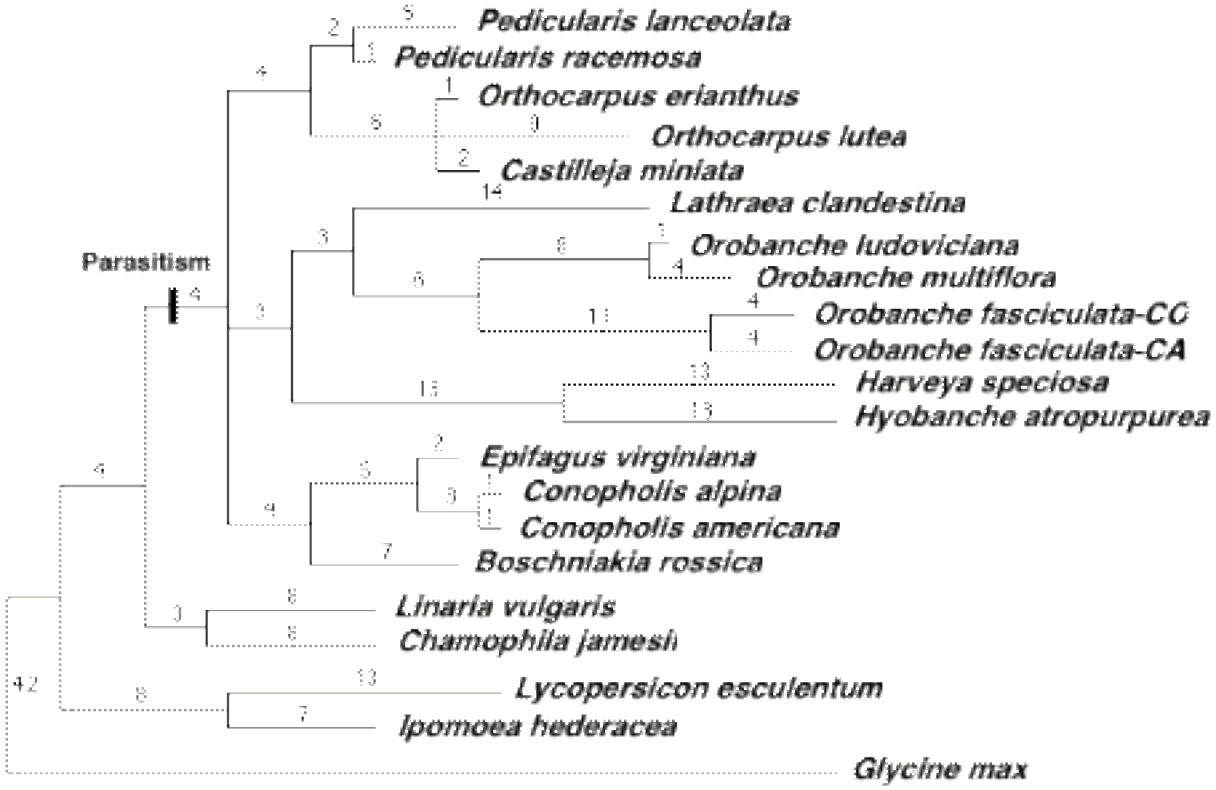
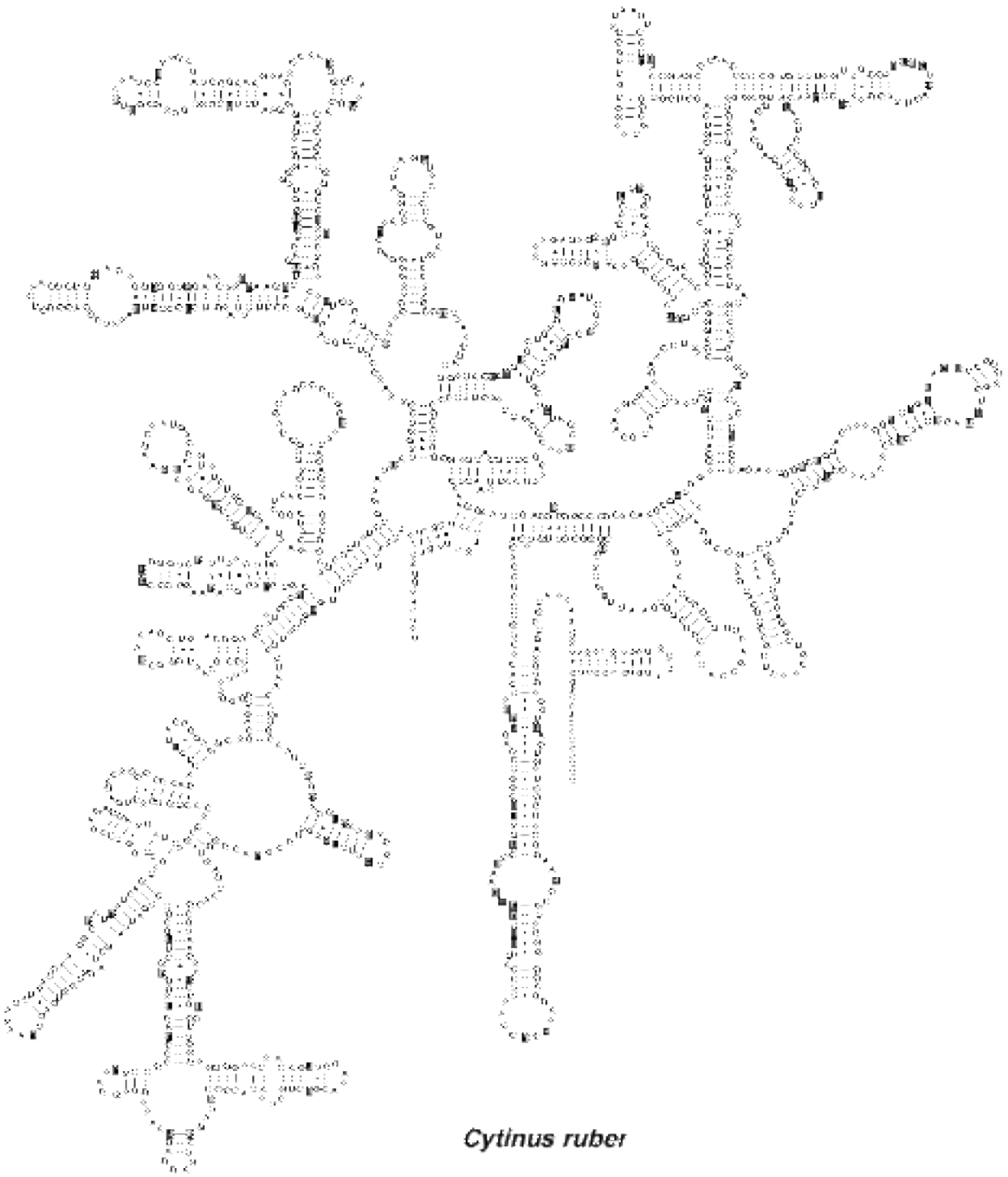


Figure 6. Secondary structural model for plastid-encoded 16S rRNA from *Cytinus ruber* (Cytinaceae or Rafflesiaceae). Bases in reverse font indicate mutations relative to the sequence of *Nicotiana* (totalling 130). Bases in lower case occur at the primer sites, hence were not determined.

Figure 6



Cytinus ruber

Figure 7. The single tree of length 808 steps found via a branch and bound search of a matrix containing16S rDNA sequences from plant plastids and cynobacterial outgroups (*Anacystis*, *Microcystis*, and *Anabaena*). The consistency index (minus uniformative sites) is 0.634 and the retention index is 0.655. Number of steps are indicated above and bootstrap values (from 100 replications) are indicated below the branches. Clades recovered in less than 50% of the trees are indicated by an asterisk (*). The loss of the inverted repeat is indicated by □ IR.

Fig. 7

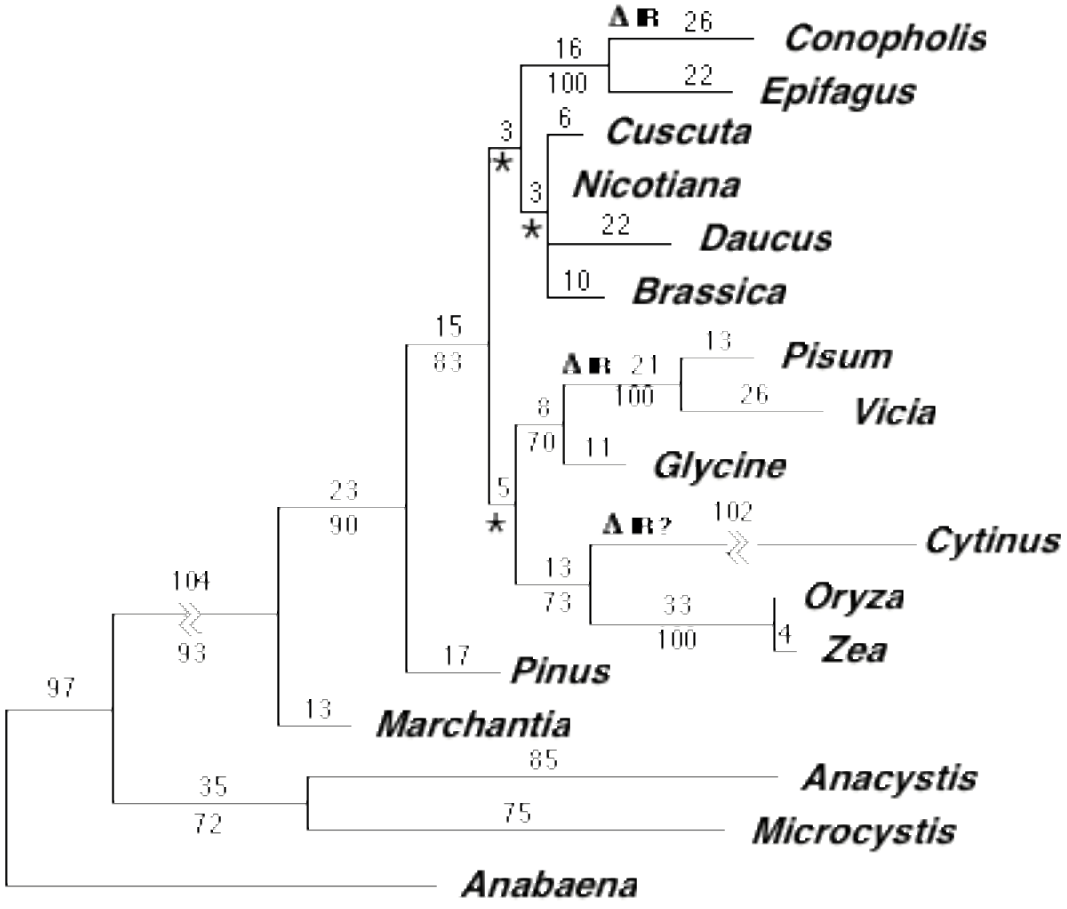


Figure 8. Histogram showing results of relative rate tests using plastid 16S rDNA sequences from 12 angiosperms and *Pinus* (as the reference taxon 3). The three-taxon tree uses K_1, the number of nucleotide substitutions per site for taxon 1, K_2 for taxon 2 and K_3 for taxon 3. The difference in nucleotide substitutions per site ($K_{13} - K_{23}$) is multiplied by 1000 for graphical purposes. Standard error values are included above the bar for each taxon.

Fig. 8