8

Molecular Phylogenetic and Evolutionary Studies of Parasitic Plants

Daniel L. Nickrent, R. Joel Duff, Alison E. Colwell, Andrea D. Wolfe, Nelson D. Young, Kim E. Steiner, and Claude W. dePamphilis

The parasitic nutritional mode is a frequently evolved adaptation in animals (Price, 1980), as well as in flowering plants (Kuijt, 1969). Heterotrophic angiosperms can be classified as either *mycotrophs* or as *haustorial parasites*. The former derive nutrients via a symbiotic relationship with mycorrhizal fungi. Haustorial parasites, in contrast, directly penetrate host tissues via a modified root called a haustorium and thereby obtain water and nutrients. Although such categories are often a matter of semantics, we use the term *parasite* in a strict sense to refer to haustorial parasites. Angiosperm parasites are restricted to the dicot subclasses Magnoliidae, Rosidae, and Asteridae; have evolved approximately 11 times; and represent approximately 22 families, 265 genera, and 4,000 species, that is, about 1% of all angiosperms (Fig. 8.1). Owing to their unique adaptations, parasitic plants have long been the focus of anatomical, morphological, biochemical, systematic, and ecological research (Kuijt, 1969; Press and Graves, 1995). For the vast majority of parasitic plants, negative effects upon the host are difficult to detect, yet others (e.g., *Striga, Orobanche*) are serious weeds of economically important crops (Kuijt, 1969; Musselman, 1980; Eplee, 1981; Stewart and Press, 1990; Press and Graves, 1995).

The degree of nutritional dependence on the host varies among haustorial parasites. Hemiparasites are photosynthetic during at least one phase of their life cycle and derive mainly water and dissolved minerals from their hosts. Obligate hemiparasites require a host plant to complete their life cycles whereas facultative hemiparasites do not. Hemiparasites can be found in Laurales (*Cassytha*), Polygalales (*Krameria*), and all families of Santalales. In Solanales (*Cuscuta*) and Scrophulariales, some species are chlorophyllous hemiparasites whereas other species are achlorophyllous holoparasites. Holoparasites represent the most extreme manifestation of the parasitic mode because they lack

This work was supported by grants from the National Science Foundation (DEB 94-07984 to DLN, DEB 91-20258 to CWD, and BIR 93-03630 to ADW), the Special Research Program of the Office of Research Development Administration, SIUC and the University Research Council of Vanderbilt University. Thanks go to C. Augspurger, W. Barthlott, J. Beaman, D. E. Bran, S. Carlquist, W. Forstreuter, J. Leebens-Mack, A. Markey, C. Marticorena, D. McCauley, S. Medbury, M. Melampy, Willem Meijer, B. Molloy, L. Musselman, R. Narayana, M. Nees, J. Paxton, S. Sargent, B. Swalla, W. Takeuchi, and M. Wetherwax for helpful discussions and/or for contributing plant material. The manuscript was improved by the critical comments of M. Bowe and an anonymous reviewer.

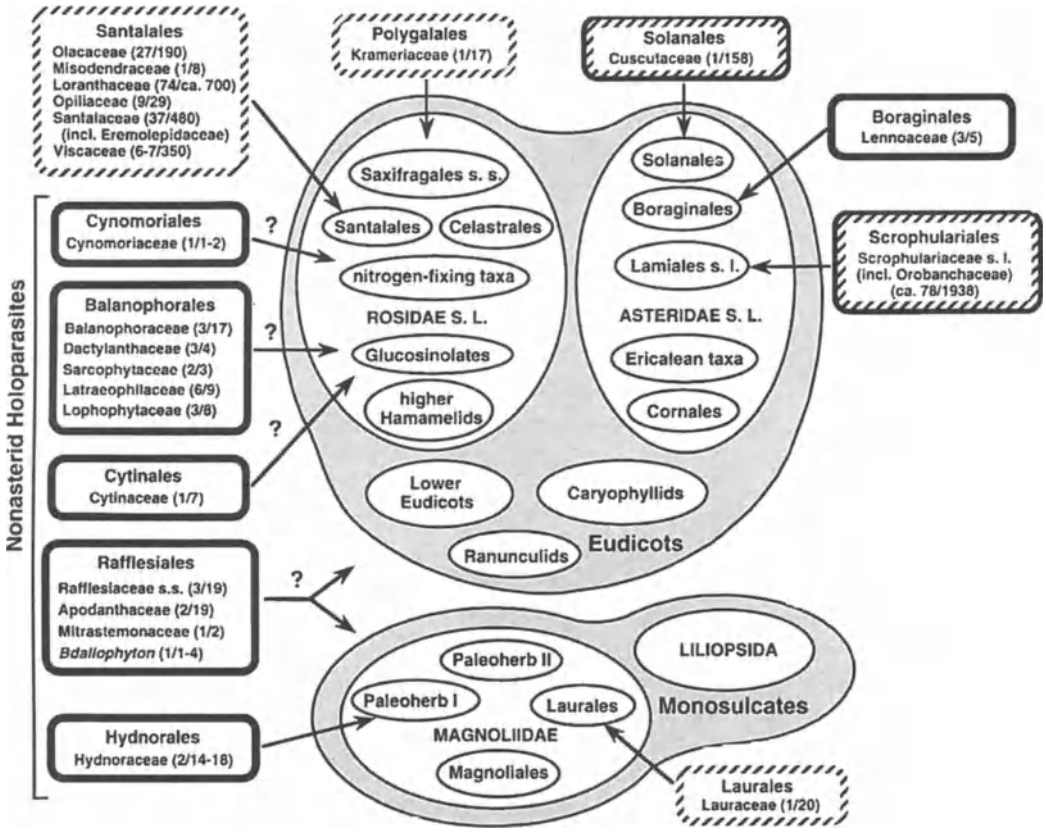

Figure 8.1. The distribution of haustorial parasitism among angiosperms. This generalized diagram incorporates information from global molecular phylogenetic studies using *rbcL* (Chase, Soltis, Olmstead et al. 1993) and nuclear 18S rDNA (Soltis, Soltis, Nickrent et al., 1997). No attempt was made to show all taxa, only to indicate groups that were supported by both studies. Hemiparasitic angiosperms are enclosed within dashed borders and holoparasites by black borders. Both trophic modes occur in Scrophulariaceae s. l. and *Cuscuta*. Arrows that touch a group indicate that strong evidence exists for the placement of that parasitic taxon within the group. Uncertain affinities are indicated by arrows with question marks. The familial classification of the nonasterid holoparasites is modified from Tahktajan (1987); however the placement of these orders is not concordant with his superordinal classification. The number of genera and species is indicated in parentheses following each family name. For Scrophulariaceae s. l., only the parasitic members are tabulated.

photosynthesis and must rely upon the host for both water and inorganic and organic nutrients. Six groups (orders or families—Fig. 8.1) are represented entirely by holoparasites: Balanophorales, Cynomoriaceae, Cytinaceae, Hydnoraceae, Lennoaceae, and Rafflesiales.

The relationships shown in Figure 8.1 are based upon those of Takhtajan (1987) as well as results of recent molecular analyses. For this paper, Santalales are considered, in a strict sense, to include Olacaceae, Misodendraceae, Loranthaceae, Opiliaceae, Santalaceae, and Viscaceae. This composition differs from that of Cronquist (1988) by excluding Balanophoraceae, Medusandraceae, and Dipentodontaceae. Morphological, cytological, and molecular evidence all point toward the separation of Cynomoriaceae from Balanophoraceae and of Cytinaceae from Rafflesiaceae (Takhtajan, 1987; Nickrent and Duff, 1996; Pazy et al., 1996). Traditional classifications have often allied these holoparasites with Santalales; however, considerable variation can be seen in alternate classifications and such an affinity is not apparent from molecular investigations (see below). Given this, the term *nonasterid holoparasites* will be used to distin-

guish these plants from holoparasites in Asteridae such as those found in Scrophulariales, Boraginales, and Solanales. The nonasterid holoparasites are Balanophorales, Cynomoriales, Cytinales, Hydnorales, and Rafflesiales (Fig. 8.1). Results of molecular phylogenetic studies indicate that the nonasterid holoparasites are not closely related to each other (Nickrent and Duff, 1996), hence the term is applied for reference purposes only.

Additional parasitic plants can be found in Scrophulariaceae, a large family (over 250 genera) that is broadly and at present inexactly defined within Scrophulariales/Lamiales (Burtt, 1965; Cronquist, 1981; Thorne, 1992; Olmstead and Reeves, 1995). Although most species are completely autotrophic, the members of two tribes (Buchnereae, Pediculareae; [Pennell, 1935]) display a wide range of parasitic modes from fully photosynthetic, facultative hemiparasites, to nonphotosynthetic holoparasites. Orobanchaceae are a group of nonphotosynthetic holoparasites closely related to the holoparasitic Scrophulariaceae. A continuum of morphological and physiological traits unites the two, as do several "transitional genera" (*Harveya, Hyobanche,* and *Lathraea*) that have been classified alternatively in one family or the other (Boeshore, 1920; Kuijt, 1969; Minkin and Eshbaugh, 1989). Orobanchaceae are alternatively included within Scrophulariaceae (Takhtajan, 1987; Thorne, 1992) or recognized, by tradition, at the family level (Cronquist, 1981). Regardless of rank, most workers are in agreement that Orobanchaceae are derived from within the parasitic Scrophulariaceae. Numerous lines of evidence support this conclusion, including the shared presence of several morphological characters (Boeshore, 1920; Weber, 1980), pollen features, and a derived chloroplast DNA restriction site loss present in both groups (C. dePamphilis, unpubl.). For these reasons, this chapter considers Scrophulariaceae s. l. to include Orobanchaceae (Fig. 8.1).

In this chapter we discuss the results of macromolecular studies of the groups shown in Fig. 8.1. Most attention is, by necessity, directed toward four groups for which DNA sequence data are available, Scrophulariaceae s. l., *Cuscuta,* Santalales, and the nonasterid holoparasites. Our goals are to demonstrate the utility of these molecular markers in documenting phylogenetic relationships (at the genus level and above) and to show how parasitic plants represent unique models that can be used to study molecular evolutionary and genetic processes such as the structure, function, and evolution of plant genomes. For example, the continuum of trophic modes in Scrophulariaceae s. l. from nonparasitic to hemiparasitic to holoparasitic makes this group ideal for investigating questions concerning the evolution of parasitism and the molecular changes that accompany adaptation to a heterotrophic existence. Holoparasitic plants that show increased rates of molecular evolution pose particular problems for phylogenetic analysis but at the same time provide intriguing subjects for studying genome reorganizations that accompany the loss of photosynthesis.

Problems with the Classification of Parasitic Plants

The placement of many parasitic plants within the global angiosperm phylogeny is not disputed. For example, despite questions about familial boundaries and interfamilial relationships within Scrophulariales, it is clear that Scrophulariaceae are allied with other sympetalous dicots of Asteridae s. l. Such is not the case, however, for the nonasterid holoparasites whose higher-level classification still remains problematic. Two processes that occur during the evolution of advanced parasitism are reduction (and/or extreme modification) of morphological features and convergence. The first may involve loss of leaves, chlorophyll, perianth parts, or even ovular integuments. Loss of features confounds phylogenetic analysis unless clear transformation series are apparent. Convergences in parasitic plants are rampant because similar features have evolved in unrelated groups. For example, the squamate habit (i.e., with scale leaves) has evolved independently in aerial parasites of Viscaceae, Santalaceae, and Misodendraceae. One of the most striking examples of convergence involves *Cuscuta* (Asteridae) and *Cassytha* (Magnoliidae), two twining, yellow to orange parasites that are frequently confused

based on a superficial examination of morphology. With reference to the internal haustorial tissues (endophytes) of Rafflesiaceae and Viscaceae, Cronquist (1981, p. 698) states "... parallelism in a number of features is in itself some indication of relationship, to be considered along with other evidence." Because parallelism involves similarities due to analogy *and* homology, and given that the common ancestry of the two families has not been established, this situation should best be described as convergence. A more serious problem is that this statement implies that all characters, whether they evolved via parallelism or convergence, should receive equal consideration when constructing a phylogeny.

The classifications of parasitic plants are often plagued by past misconceptions or overemphasis on a few conspicuous characters. The parasitic habit itself, known to have had multiple origins, was used by Cronquist (1981) to link Balanophoraceae, Hydnoraceae, and Rafflesiaceae to Santalales. It is not our contention that molecular data are more immune to the effects of convergence, parallelism, and reversal (i.e., homoplasy) than morphological characters, only that homologous DNA sequences provide additional, independent genetic characters that can be rigorously analyzed to test phylogenetic hypotheses. Furthermore, molecular data can be used to explore the dynamics of molecular evolution in parasitic plants by examining the genetic structure and biochemical processes underlying evolutionary change.

THE PLASTOME OF PARASITIC PLANTS

The chloroplast genomes (cpDNA) of photosynthetic plants are circular molecules ranging in size from about 120 to 217 kilobase pairs (kb) (Downie and Palmer, 1992). These genomes typically have a large (about 25 kb) inverted repeat that contains rRNA-encoding and other genes, and separates the large and small single-copy regions of the molecule. A typical chloroplast genome, such as the completely sequenced cpDNA of *Nicotiana* (Shimada and Sugiura, 1991), consists of about 112 genes and potentially functional open reading frames (ORFs). Although a substantial number of the genes (29, approximately 25%) encode proteins for photosynthetic carbon fixation and electron transport, about half (60, 54%) encode proteins or RNAs involved in gene expression (transcription, translation, and related processes). An additional 11 genes (10%; the *ndh* genes) encode putative "chlororespiratory" proteins, based on their sequence similarity to mitochondrial *ndh* genes (Shimada and Sugiura, 1991). Ten or more large ORFs (and many smaller ones) also represent potentially functional protein genes, but their functions are largely unknown. Among the many hundreds of species whose cpDNAs have been characterized using restriction endonuclease mapping (Downie and Palmer, 1992), most retain a very similar gene content and gene order. Molecular systematic research on photosynthetic plants has extensively used phylogenetically informative variation in the presence and absence of restriction endonuclease sites, as well as a very large and rapidly growing comparative database of DNA sequences for the plastid genes *rbcL, matK,* and *ndhF,* and smaller databases for other plastid genes (see Chapter 1). Rare structural mutations such as inversions and insertions or deletions of genes or introns provide additional characters of special phylogenetic significance (reviewed by Downie and Palmer, 1992; see also Chapter 1).

Because many of the parasitic plants discussed herein are nonphotosynthetic, the more general term *plastid* will be used instead of *chloroplast* and ptDNA (plastid DNA or plastome) will be used instead of cpDNA (chloroplast DNA). Given that a large fraction of the plastome is devoted to photosynthetic function and to the expression of photosynthetic and other genes, parasitic plants provide a unique opportunity to determine the extent to which the evolutionary conservation of ptDNA is related to photosynthetic ability. In this regard, parasitic plants may be considered to be natural genetic mutants for the dissection of ptDNA function and evolutionary processes (dePamphilis, 1995, and references therein). At the same time, development of a phylogenetic framework to interpret significant variation among parasitic and nonparasitic lineages is complicated by the absence of many plastid genes from at least some parasitic lineages and by the acceleration of evolu-

tionary rate for many other plastid and nuclear genes (see Rate Variation among Genomes and Lineages).

Epifagus

Epifagus virginiana (beechdrops, Scrophulariaceae s. l.) is a holoparasite native to eastern North America whose sole host is *Fagus grandifolia* (American beech). Although entirely lacking chlorophyll and photosynthetic ability, *Epifagus* retains plastids (Walsh et al., 1980) and ptDNA (dePamphilis and Palmer, 1989). Detailed mapping (dePamphilis and Palmer, 1990) and complete sequencing of the *Epifagus* plastome (Wolfe et al., 1992c) revealed a greatly reduced genome of only 70,028 bp, containing just 42 intact genes. Surprisingly, although the *Epifagus* ptDNA had sustained a large number of deletions (and some small insertions as well) relative to another asterid, *Nicotiana,* the two genomes are almost entirely colinear (dePamphilis and Palmer, 1990). Complete sequencing revealed only one small inversion in *trnL* of the small single-copy region of *Epifagus* relative to *Nicotiana.* Furthermore, the parasite retains a nearly full-sized inverted repeat of 22,735 bp that separates the greatly reduced large and small single-copy regions of the genome.

Comparison of the ptDNA gene content of *Epifagus* with that of *Nicotiana* revealed a highly selective pattern of gene deletion in the holoparasitic plant (Table 8.1). All plastid-encoded photosynthetic genes have been deleted

Table 8.1. Plastid genes in *Epifagus* (42 total[a]) compared to *Nicotiana* (112 total).

	Gene Present in *Epifagus*	Deleted or Pseudogene (ψ) in *Epifagus*
Photosynthesis		
Photosystem I		*psaA, B, C, I, J*
Photosystem II		ψ*psbA,* ψ*B, C, D, E, F,* *psbH, I, J, K, L, M*
Cytochrome b/f		*petA, B, D, G*
ATP synthase		ψ*atpA,* ψ*B, E, F, H, I*
Calvin cycle		ψ*rbcL*
Chlororespiration		*ndhA,* ψ*B, C, D, E, F, G,* *ndhH, I, J, K*
Gene Expression		
rRNA	16S, 23S, 4.5S, 5S	
Ribosomal protein	*rps2, 3, 4, 7, 8, 11,* *rps14, 18, 19,* *rpl2, 16, 20, 33, 36*	*rps15, 16,* ψ*14,* *rpl22,* ψ*23, 32*
Transfer RNA	$trnD_{GUC}$, E_{UUC}, F_{GAA}, $trnH_{GUG}$, I_{CAU}, L_{CAA}, $trnL_{UAG}$, M_{CAU}, N_{GUU}, $trnP_{UGG}$, Q_{UUG}, R_{ACG}, $trnS_{GCU}$, S_{UGA}, W_{CCA}, $trnY_{GUA}$, $trnfM_{CAU}$	ψ$trnA_{UGC}$, ψC_{GCA}, G_{GCC}, $trnG_{UCC}$, ψI_{GAU}, K_{UUU}, $trnL_{UAA}$, ψR_{UCU}, ψS_{GGA}, $trnT_{GGU}$, T_{UGU}, V_{GAC}, $trnV_{UAC}$
RNA polymerase		ψ*rpoA, B, C1, C2*
Maturase	*matK*	
Initiation factor	*infA*	
Other protein genes	*clpP, accD,* ORF1738[b], ORF2216[b]	ORF29, 31, 34, 62, 168, ORF184, 229, 313

[a]Data based on gene mapping (dePamphilis and Palmer, 1990) and complete sequence (Wolfe et al., 1992c). All genes listed are present in *Nicotiana* chloroplast DNA except *infA,* which is a pseudogene in that plant (Wolfe et al., 1992c).

[b]ORF1738 and ORF2216 in *Epifagus* are homologs of *Nicotiana* ORF1901 and ORF2280, respectively (Wolfe et al., 1992c).

from *Epifagus* or remain as small fragments or pseudogenes. Similarly, the 11 *ndh* genes have been lost. A majority of genes of unknown function (ORFs 168, 184, 229, and 313, plus a host of smaller ones) has been completely lost from *Epifagus.* In contrast, 38 (or over 90%) of the retained genes represent components of the plastid machinery for gene expression. All of the four ribosomal RNA genes are retained, as are intact copies of 13 of 19 plastid-encoded ribosomal protein genes and 17 of 30 plastid tRNA genes. The implications of these results are that ptDNA can evolve very rapidly under certain predictable conditions, and that DNA deletions, both large and small, are a dominant mode of structural evolution in that molecule. This is a striking example of the role natural selection plays in shaping genome structure and suggests that rapid structural alteration is possible when photosynthetic constraints are altered.

These results also have important implications for our understanding of the function of ptDNA. RNA polymerase subunit genes (*rpoA, B, C1,* and *C2*), which encode the central enzymes of transcription, are present in tobacco but absent in *Epifagus.* If plastid-encoded RNA polymerase is solely responsible for ptDNA transcription, we would expect the beechdrops ptDNA to be unexpressed. To the contrary, despite the loss of all RNA polymerase genes, the plastome of *Epifagus* is clearly transcribed and is probably also translated (Wolfe et al., 1992c). The evidence comes from several sources, including direct observation of transcripts of plastid rRNA genes (dePamphilis and Palmer, 1990), ribosomal protein genes (Ems et al., 1995), and additional plastid ORFs (Ems et al., 1995). *Epifagus* plastid RNAs are subject to appropriate intron splicing and even experience RNA editing, as observed in other plastid RNAs (Ems et al., 1995). Furthermore, the highly specific pattern of gene retention and bias of synonymous over nonsynonymous base substitutions, particularly in large genes subject to significant levels of sequence divergence (Wolfe et al., 1992b, 1992c; dePamphilis et al., 1997) is clearly indicative of genes that have evolved under functional constraint.

That gene expression in plastids is linked to nuclear-encoded genes is well documented (Mayfield et al., 1995). Expression of plastid genes in the absence of plastid-encoded RNA polymerase would imply that another RNA polymerase, presumably of nuclear origin, is involved in the transcription of beechdrops ptDNA (Morden et al., 1991; Ems et al., 1995). It is interesting to note also that virtually all of the proposed eubacterial-type -35 and -10 promoter regions have either diverged or been deleted in *Epifagus* (Wolfe et al., 1992a, 1992c), consistent with the possibility that these genes now use different promoters, possibly ones capable of interacting with an RNA polymerase of nonplastid origin. Similarly, the loss of 13 of the 30 normally plastid-encoded tRNA genes would suggest that plastid translation, to the extent that it occurs, must also require the import of some tRNAs from outside the plastid (Wolfe et al., 1992c). Finally, the retention of just four protein-coding genes not involved in gene expression (*clpP, accD,* ORF1738, and ORF2216) in the plastome of *Epifagus* suggests that one or more of these genes encodes a protein required for a nonbioenergetic (i.e., not involved in photosynthetic or chlororespiratory) function in the *Epifagus* plastid and may be the raison d'être for the retention of a functional ptDNA (Wolfe et al., 1992c; dePamphilis, 1995).

A related study investigated the phylogenetic distribution of one of the lost tRNA genes, tRNA-cys, in 11 parasitic and nonparasitic plants related to *Epifagus* (Taylor et al., 1991). The polymerase chain reaction was used to amplify fragments that would contain the tRNA-cys locus from each plant. The fragments were then cloned and sequenced to reveal that all of the photosynthetic plants had a normal tRNA-cys gene, whereas all of the nonphotosynthetic plants had lost the gene. This result suggests that the tRNA-cys gene was lost at about the same time as photosynthesis was lost and that the two events may be causally related.

Conopholis

Concurrent with the early work on *Epifagus,* the plastome of another orobanchaceous holoparasite, *Conopholis americana* (squawroot), was being examined at the molecular level. Using heterologous probes, Wimpee et al. (1991) doc-

umented the modification or absence of many photosynthetic genes. Later, the rDNA operon was cloned and sequenced, thereby demonstrating that the 16S/23S spacer lacked tRNA genes and contained substantial deletions (Wimpee et al., 1992a). Experiments using the polymerase chain reaction and hybridizations failed to detect the lost tRNA genes elsewhere in the squawroot genome (Wimpee et al., 1992b). Colwell (1994) conducted restriction site mapping of the plastome of squawroot, documenting its size as 43 kb, making it the smallest ptDNA molecule then observed in plants. This genome is only slightly smaller than that of *Epifagus* if the length of one inverted repeat is removed from the latter genus (70 kb − 22.7 kb = 47.3 kb).

Cuscuta

Other molecular genetic investigations have focused on *Cuscuta,* a genus sometimes placed in its own family (Cuscutaceae) but widely recognized to share a common ancestor with nonparasitic Convolvulaceae. *Cuscuta,* like Scrophulariaceae, includes hemiparasitic species (e.g., *C. reflexa*) as well as holoparasitic species (e.g., *C. europaea*) that lack thylakoids, chlorophyll, RUBISCO (ribulose bisphosphate carboxylase/oxygenase) and light-dependent CO_2 fixation, but (strangely) retain *rbcL* (Machado and Zetsche, 1990). Although the plastome of *Cuscuta* is yet to be sequenced completely, significant progress is being made. A 6-kb portion of ptDNA from *C. reflexa* that includes *petG, trnV, trnM, atpE, atpB,* and *rbcL* was sequenced (Haberhausen et al., 1992; Haberhausen and Zetsche, 1992). Further work on this species (Bömmer et al., 1993) resulted in sequences for 16S rDNA, *psbA, trnH,* ORF740, ORF77, *trnL,* and ORF55. Later, a 9-kb portion of ptDNA from this same species that included a large portion of inverted repeat A spanning a segment from *trnA* to *trnH* was cloned and sequenced (Haberhausen and Zetsche, 1994) Although some short sequences were identical to those of *Nicotiana* (e.g., *trnI*), many deletions were observed. For example, *rpl2* and *rpl23* were both deleted, and ORF2280 was reduced to only 740 bp. These results show that, like *Epifagus, C. reflexa* has experienced major deletions in the plastome. The complete loss of ribosomal protein genes such as *rpl2* invites questions about how the translational apparatus of the plastid functions and its relationship to the nuclear and mitochondrial genomes. The plastid genome of the holoparasite *C. europaea* has undergone additional deletions in comparison to *C. reflexa,* such as the loss of the cis-spliced intron of *rps12* (Freyer et al., 1995). Despite this alteration, correctly processed transcripts and ribosomal proteins are produced, thus providing evidence that the translational apparatus of the plastid may remain functional following significant structural alterations.

Do the Nonasterid Holoparasites Retain a Plastome?

As described in the foregoing discussion, the plastomes of *Epifagus, Conopholis,* and *Cuscuta* have experienced extensive losses of photosynthetic genes, yet all have retained intact and functional ribosomal cistrons. It was therefore reasoned that if any portion of the plastome remained in nonasterid holoparasites, it would likely be the ribosomal cistron. The first evidence that rDNA was present in these holoparasites was obtained when plastid 16S rDNA from *Cytinus* (Cytinaceae) was amplified and sequenced (Nickrent et al., 1995; Nickrent and Duff, 1996). In pairwise comparisons to other angiosperms, this 16S sequence contained approximately three times more base substitutions, yet this rRNA retained all major secondary structural features (because most changes were compensatory), thus suggesting functionality. Since these initial studies, 16S rDNA sequences have been obtained from all nonasterid holoparasite lineages (Nickrent et al., 1997a). These represent the most diverse 16S rRNA sequences documented among angiosperms and contain several structural features unknown among land plants.

Southern blots using both homologous and heterologous probes developed for five plastid genes (16S rDNA and four ribosomal protein genes) resulted in positive hybridizations to digested total genomic DNA from *Cytinus* (Nickrent et al., 1997b). These preliminary data suggest the *Cytinus* genome is approximately 20 kb

in size, thus the smallest yet documented for angiosperms. Hybridizations using more derived holoparasites such as *Rafflesia* and *Balanophora* were negative, thus suggesting their plastome is either absent or even more extensively modified than that of *Cytinus*. Recently, sequences of PCR-amplified products derived from the plastid 16S–23S spacer have been obtained from *Cytinus* and *Cynomorium* (D. Nickrent and R. Duff, unpubl.), further suggesting the retention of a plastome in these plants.

MOLECULAR PHYLOGENETIC STUDIES OF SCROPHULARIACEAE S. L.

Systematic and Phylogenetic Problems

Until recently, most opinions on the circumscription of Scrophulariaceae could be traced to concepts from the nineteenth century, such as the recognition of three subfamilies (Pseudosolaneae, Antirrhinoideae, and Rhinanthoideae) and a distinct Orobanchaceae. Rhinanthoideae were traditionally considered to include the parasitic tribes Buchnereae and Euphrasieae, as well as the nonparasitic tribes Veronicae and Digitaleae. Anatomical evidence was used to refute a presumed relationship between Rhinanthoideae and the slightly zygomorphic Pseudosolaneae and Solanaceae (Thieret, 1967). Other family-level issues concern the presumed "edges" of the family, including whether to recognize the Selagineae and/or the Globulariaceae as formal taxa within Scrophulariaceae. A general trend has been the recognition of various small tribes (including Angeloneae, Russelieae, Collinsieae, Melospermeae, and many others) and especially the splitting of three large and heterogeneous tribes, the Digitaleae (Hong, 1984), Gratioleae, and Cheloneae (Thieret, 1967) into smaller, presumably monophyletic tribes. Attempts to analyze cladistically morphological variation across the entire family have met with little success, leading An-Ming (1990) to speculate about the possible paraphyly, or even polyphyly, of the family as presently defined.

Although a close relationship between Orobanchaceae and parasitic Scrophulariaceae seems assured, genera such as *Harveya, Hyobanche,* and *Lathraea,* have been alternatively placed in either family (Kuijt, 1969). Does the evolutionary series nonparasite·hemiparasite·holoparasite explain the pathway to all holoparasites? Are the transitional genera and Orobanchaceae monophyletic or are they independent lineages that have converged on a nonphotosynthetic lifestyle? If the latter is true, how many times has photosynthesis been lost? The availability of numerous extant species displaying all modes of parasitism makes these questions tractable via systematic analyses.

The molecular phylogeny of parasitic Scrophulariaceae has been examined using sequences from three plastid genes: *rps2, matK,* and *rbcL.* The first two are relatively rapidly evolving plastid genes that encode nonphotosynthetic proteins (see Chapter 1) and that we have found are present in all parasitic Scrophulariaceae regardless of photosynthetic ability. Although the photosynthetic gene *rbcL* is absent from some holoparasites (see Phylogeny of Scrophulariaceae s. l. using *rbcL*), it is present in most taxa. Phylogenetic analyses have revealed much about the molecular evolution of this key plastid gene in relation to photosynthesis.

Phylogeny of Scrophulariaceae s. l. Using rps2

Phylogenetic analysis of plastid ribosomal protein gene *rps2* was conducted for 55 species of Scrophulariaceae and related families (Fig. 8.2), including 32 species of parasites (dePamphilis et al., 1997). Nonparasites were selected to reflect the patterns shown following analysis of a larger data set (58 genera of nonparasitic Scrophulariaceae and relevant outgroup taxa; C. dePamphilis unpubl.). Parsimony analysis of *rps2* identifies a single clade (bootstrap value of 53%) containing a broad sampling of parasitic Scrophulariaceae s. l. With the exception of *Lathraea,* the "transitional" genera *Harveya* and *Hyobanche* are present on the *rps2* consensus tree as a clade (88% bootstrap support) of hemiparasitic Scrophulariaceae; without these genera, the remaining orobanchaceous taxa are weakly supported as monophyletic. These data indicate that hemiparasites gave rise to holoparasites a minimum of five separate times (dePamphilis et al.,

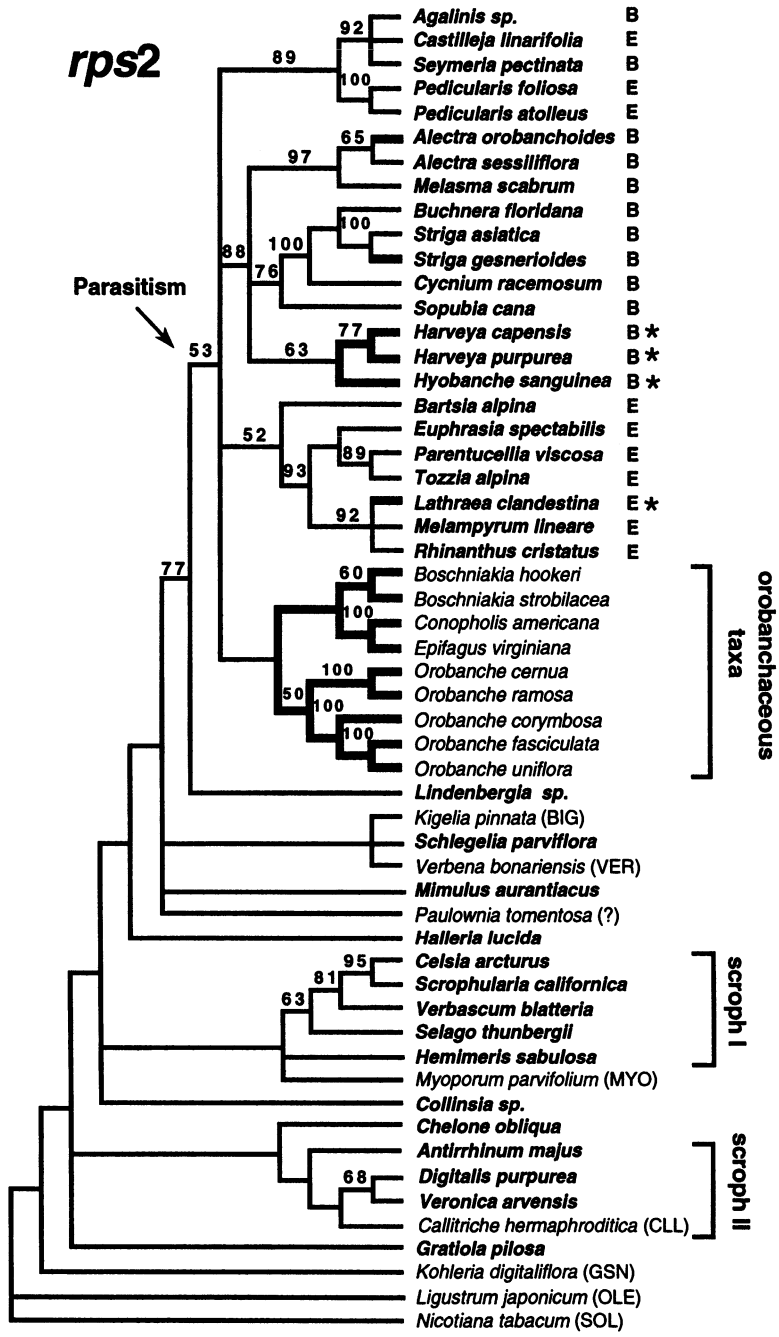

Figure 8.2. Phylogenetic reconstruction of Scrophulariaceae and outgroups based on maximum parsimony analysis of plastid *rps2* gene sequences. The tree was constructed using 174 phylogenetically informative base substitutions. Strict consensus of 78 trees at 660 steps (Consistency Index [CI] without autapomorphies = 0.492; Retention Index [RI] = 0.657). Bootstrap values shown above branches; nodes without values were supported in less than 50% of the replications. The arrow indicates a parasite clade including genera traditionally assigned to Orobanchaceae. Nonphotosynthetic holoparasites are indicated by thick lines. Taxa in bold font are traditionally placed in Scrophulariaceae. The traditional parasitic tribes Buchnereae and Euphrasieae are indicated by B and E, respectively. Taxa followed by an asterisk (*) are "transitional" genera classified alternatively in Scrophulariaceae or Orobanchaceae. Scroph I and II refers to clades identified in Olmstead and Reeves (1995). Abbreviations for outgroups and families in Scrophulariales: BIG = Bignoniaceae; CLL = Callitrichaceae, GSN = Gesneriaceae; MYO = Myoporaceae; OLE = Oleaceae; SOL = Solanaceae; VRB = Verbenaceae. The "?" for *Paulownia* indicates its higher-level classification is uncertain.

1997). Seven of the nine genera classified as members of Euphrasieae occur in a single clade; the remaining two genera (*Castilleja* and *Pedicularis*) occur in a clade with two genera of Buchnereae (*Agalinis* and *Seymeria*).

In contrast to the apparent monophyly of the parasitic taxa, the nonparasitic Scrophulariaceae are only weakly resolved with *rps2* and appear as a complex paraphyletic assemblage with representatives of other families in Scrophulariales such as Bignoniaceae, Verbenaceae, and Myoporaceae. Low bootstrap values suggest that the power of this short gene to resolve phylogenetic relationships in the nonparasitic lineages is low. Yet, two of the nonparasitic lineages recovered correspond to the scroph I and scroph II clades identified by Olmstead and Reeves (1995) in a phylogenetic analysis using *rbcL* and *ndhF* sequences. Other genera not sampled by Olmstead and Reeves (*Gratiola, Collinsia, Halleria, Mimulus,* and *Lindenbergia* plus the entire parasitic lineage) are not components of these clades. These results suggest that at least three and possibly additional major lineages may represent what has now clearly been identified as a polyphyletic Scrophulariaceae (Olmstead and Reeves, 1995). Neither *Veronica* nor *Digitalis,* genera usually included within the rhinanthoid subfamily, are close relatives of the parasitic plants. In contrast, *Lindenbergia,* a nonparasitic genus originally classified in the Gratioleae (Wettstein, 1897), occurs as sister to all other parasitic Scrophulariaceae.

Phylogeny of Scrophulariaceae s. l. Using rbcL

Studies employing *rbcL* sequencing for phylogenetic reconstruction in Scrophulariaceae have been, until recently, restricted to nonparasitic plants (Olmstead and Reeves, 1995; Wolfe and dePamphilis, 1995, 1997; dePamphilis et al., 1997). Recent studies by Wolfe and dePamphilis (1995, 1997) have revealed accelerated synonymous and nonsynonymous substitution rates for *rbcL* in several lineages of derived members of Scrophulariaceae. In comparison to nonparasitic plants of Scrophulariales, the evolutionary rate of *rbcL* in those parasites that retain this gene is faster, thus enabling more resolved phylogenetic reconstructions. The *rbcL* locus from 30 additional taxa representing 20 genera of Scrophulariaceae has been sequenced (Wolfe and dePamphilis 1995, 1997). Phylogenetic reconstruction based on the new sequences and those from Olmstead and Reeves (1995) reveals that Scrophulariaceae, as traditionally circumscribed, are not monophyletic (Fig. 8.3). Although Olmstead and Reeves (1995), arrived at the same conclusion based on a combined analysis of *rbcL* and *ndhF* sequences, specific relationships between major clades differ between the two studies. The "scroph I" and "scroph II" clades were distinct in Olmstead and Reeves (1995) but in the present analysis they occur in one clade with representatives of other Scrophulariales.

One of the major results of the phylogenetic analysis based on both *rps2* (Fig. 8.2) and *rbcL* (Fig. 8.3) sequences is that there is a single origin of the parasitic habit in Scrophulariaceae. Taxa traditionally placed in Orobanchaceae are intercalated with hemi- and holoparasitic genera of the Scrophulariaceae, hence Orobanchaceae (as traditionally defined) are paraphyletic. The genus *Orobanche* is also not monophyletic according to the *rbcL* strict consensus tree (Fig. 8.3). The *rps2* results suggested at least five independent losses of photosynthesis. A strict interpretation of the *rbcL* tree indicates seven losses of photosynthesis, assuming that it is unlikely that the nonphotosynthetic ancestors of *Orobanche, Boschniakia,* and others gave rise to photosynthetic hemiparasites. Some of these relationships, such as lack of monophyly for *Orobanche* or a distant relationship between *Harveya* and *Hyobanche,* are not in accord with *rps2* or 18S rDNA data (see next section); hence further examination of the phylogenetic utility of *rbcL* in these plants is in order.

In those holoparasites where *rbcL* no longer produces a functional product, it is considered a pseudogene. In most cases of pseudogene formation, a functional gene is retained in addition to a degenerated duplicated copy (e.g., β-globin genes [Li, 1983]); however, in Scrophulariaceae no duplication has occurred, and only the primary *rbcL* pseudogene remains. These data indicate that primary pseudogene formation occurred independently three times (Fig. 8.3). The most surprising discovery from this *rbcL* phylogenetic analysis is

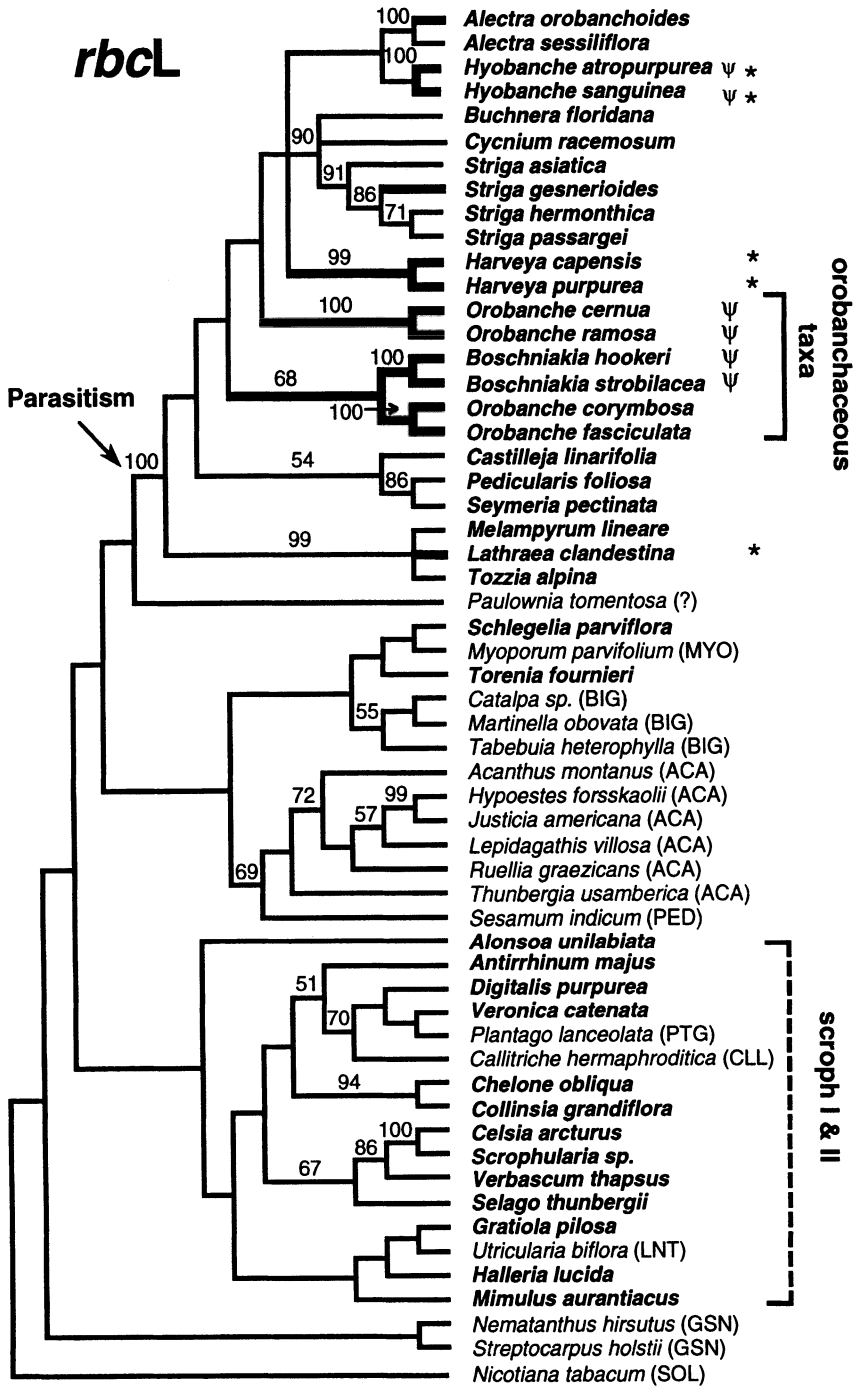

Figure 8.3. Phylogenetic reconstruction of Scrophulariaceae based on maximum parsimony analysis of *rbcL* nucleotide sequences. Taxa with primary pseudogenes are indicated as ψ (see Fig. 8.6 and text). Strict consensus of 12 most-parsimonious trees with 1,482 steps (CI = 0.550, RI = 0.521). The arrow indicates a parasite clade including genera traditionally assigned to Orobanchaceae. Nonphotosynthetic holoparasites are indicated by thick lines. Taxa in bold font are traditionally placed in Scrophulariaceae. The clade identified as scroph I and II refers to taxa described in Olmstead and Reeves (1995), but here also includes representatives of other families. Taxa followed by an asterisk (*) are "transitional" genera classified alternatively in Scrophulariaceae or Orobanchaceae. Abbreviations for families in Scrophulariales in addition to those in Fig. 8.2: ACA = Acanthaceae, GLB = Globulariaceae, LNT= Lentibulariaceae, PED = Pedaliaceae, PTG = Plantaginaceae.

that many holoparasitic Scrophulariaceae (e.g., *Alectra, Harveya, Lathraea, Orobanche,* and *Striga*) retain an open reading frame for the gene (Fig. 8.3). The independence of structural mutations among and/or within genera for the pseudogene sequences (see Structural Analysis of RUBISCO) suggests that disruption of the gene occurred independently following the loss of photosynthesis in *Boschniakia, Conopholis, Hyobanche, Epifagus,* and *Orobanche* (dePamphilis and Palmer, 1990; Wolfe et al., 1992c; Colwell, 1994; Wolfe and dePamphilis, 1997). Combined, these data suggest that a *rbcL* ORF may have been retained in these holoparasitic lineages after the loss of photosynthesis and that perhaps the gene was functional at some point following the adaptation to heterotrophy.

Phylogeny of Scrophulariaceae s. l. Using Nuclear 18S rDNA

A general review of the organization of nuclear 18S rDNA and its use in phylogenetic studies of plants can be found in Chapter 1. Phylogenetic relationships among hemi- and holoparasitic Scrophulariaceae using 18S rDNA sequences were determined by Colwell (1994). That study examined nine parasitic genera; subsequent sequencing has increased to 18 the number of 18S rDNA sequences of Scrophulariaceae s. l. now available for analysis (Nickrent and Duff, 1996). Using *Glycine, Lycopersicon,* and *Ipomoea* as outgroups, a clade is retrieved (Fig. 8.4) containing the nonparasitic *Linaria* and *Chionophila;* this clade is sister to the remaining Scrophulariaceae s. l. These results based on 18S rDNA sequences indicate, as did *rps2* and *rbcL,* that parasitism arose just once in this group. Although sampling was limited compared to analyses of the chloroplast genes, a number of relationships are concordant, such as clades containing: (1) *Pedicularis, Orthocarpus,* and *Castilleja;* (2) *Harveya* and *Hyobanche;* and (3) *Epifagus* and *Conopholis.* Unlike the results from the chloroplast genes, 18S rDNA sequence analysis resulted in a strong association (91% bootstrap support) between the "transitional" genus *Lathraea* and an *Orobanche* clade that is composed of only North American taxa. The rDNA tree also indicates (as does *rps2*) that this clade is not closely related to the North American genera *Epifagus, Conopholis,* and *Boschniakia,* making Orobanchaceae (in the tra-

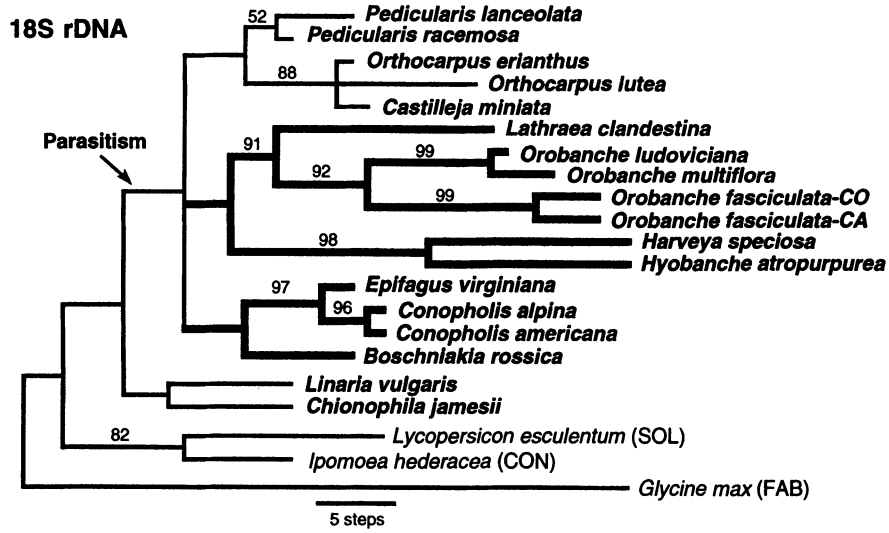

Figure 8.4. Parsimony analysis of nuclear 18S rDNA sequences from Scrophulariaceae s. l. Strict consensus phylogram of four trees at 246 steps (CI without autapomorphies = 0.553; RI = 0.643. Bootstrap values (from 100 replications) shown above branches; nodes without values were supported in less than 50% of the replications. The arrow indicates a parasite clade including genera traditionally assigned to Orobanchaceae. Nonphotosynthetic holoparasites are indicated by thick lines. *Linaria* and *Chionophila* are nonparasitic Scrophulariaceae. *Lycopersicon* (Solanaceae), *Ipomoea* (Convolvulaceae) and *Glycine* (Fabaceae) were included as outgroups.

ditional sense) paraphyletic. Although nuclear 18S rDNA sequences usually provide insufficient data at the rank of family and below, increased rates of nucleotide substitution, especially in orobanchaceous and transitional genera, have allowed greater resolution of some relationships. From the same taxa sampled for *rps2* and *rbcL*, 18S rDNA sequences should be obtained, thereby allowing direct comparisons of resulting trees and analyses of combined data sets.

MOLECULAR PHYLOGENETIC STUDIES OF SANTALALES AND NONASTERID HOLOPARASITES

Systematic and Phylogenetic Problems

Traditional classifications place Santalales in Rosidae, often near Celastrales (Cronquist, 1981; Takhtajan, 1987). In the system of Cronquist (1981), Santalales comprise ten families (including Balanophoraceae) and this order is allied with Rafflesiales (Rafflesiaceae and Hydnoraceae). The phylogenetic system proposed by Takhtajan (1987) is similar, except for the placement of the nonasterid holoparasites (Fig. 8.1). Takhtajan (1987) places Rafflesiaceae, Hydnoraceae, and Balanophoraceae (minus Cynomoriaceae) in subclass Magnoliidae, not Rosidae. For Santalales, concepts of interfamilial relationships date to the nineteenth century. A noncladistic phylogeny of the order was proposed by Fagerlind (1948); however, this system was overly influenced by the presence of *Allium* type embryo sacs and reductions seen in the gynoecium, thereby resulting in unlikely relationships such as the derivation of Balanophoraceae from Viscaceae. As with others before and after him, Fagerlind's system failed to deal with convergent features in unrelated groups.

At least three systems for interfamilial relationships in Santalales have been proposed that differ mainly in their circumscription of Eremolepidaceae and Santalaceae. The first places Santalaceae as sister to Viscaceae and Eremolepidaceae as sister to Loranthaceae (Kuijt, 1968, 1969). The second also places Santalaceae as sister to Viscaceae, but does not consider Eremolepidaceae as distinct from the former family (Wiens and Barlow, 1971). The third hypothesis (Bhandari and Vohra, 1983) allies Loranthaceae with Viscaceae (the latter including Eremolepidaceae). All three of these systems include Opiliaceae as a part of Olacaceae; the latter family being the least derived in the order because it contains both parasitic and nonparasitic members. These differing hypotheses have provided the impetus for molecular phylogenetic investigations (Nickrent and Franchina, 1990).

Phylogeny of Santalales Using Nuclear 18S rDNA and rbcL

At present, 62 nuclear 18S rDNA (Nickrent and Duff, 1996) and 37 *rbcL* sequences exist for representatives of six families of Santalales (Fig. 8.1). For Viscaceae and Misodendraceae, all genera have been sampled. Phylogenetic analyses for the same suite of 37 taxa using 18S rDNA and *rbcL* sequence data, separately and in combination, allow the utilities of these molecules to be compared directly (Table 8.2). The

Table 8.2. Comparison of the results of phylogenetic analyses of Santalales using 18S rDNA, *rbcL*, and 18S rDNA/*rbcL* combined.[a]

Tree Characteristic	18S rDNA	*rbcL*	18S + *rbcL*
Length of shortest tree (steps)	1047	1134	2212
Number of shortest trees	15	8	11
Number of characters	1741	1410	3151
Number of variable characters	433	468	901
Number of potentially informative characters	226	275	501
Consistency Index/C.I. minus uninformative sites	0.5/0.39	0.54/0.44	0.51/0.40
Homoplasy Index/H.I. minus uninformative sites	0.5/0.61	0.46/0.56	0.48/0.60
Retention Index	0.57	0.61	0.58
Rescaled Consistency Index	0.29	0.33	0.30

[a]Sequences for the same suite of 38 taxa were used in a heuristic search using PAUP.

number of variable characters and the number of phylogenetically informative characters for the two molecules are similar, as are the consistency indices. The analysis of the 37-taxon 18S rDNA data set yields trees generally concordant with those obtained following analysis of the 62-taxon matrix. Trees summarizing family-level relationships derived from the 18S rDNA and *rbcL* analyses are generally quite similar and differ mainly in the placement of Opiliaceae (Fig. 8.5). Given the overall concordance between the topologies, the data sets were combined and analyzed together. The topology of the combined-analysis phylogram is congruent with that obtained from the separate 18S rDNA data set, including the position of Opiliaceae; hence it will be used in the following discussion of relationships.

In agreement with traditional phylogenetic systems, Olacaceae represent the least derived family in the order. Despite limited generic sampling, neither 18S nor *rbcL* phylograms indicate that the family is monophyletic (Fig. 8.5). In the combined 18S/*rbcL* analysis, *Schoepfia* is sister to *Misodendrum,* and this clade is sister to Loranthaceae. Anatomical evidence supporting the distinctiveness of *Schoepfia* from other Olacaceae include its possession of aliform-confluent parenchyma (Sleumer, 1984) and tracheid/vessel features (Reed, 1955). The latter study also noted similarities in the pollen of *Schoepfia* with the more derived family Santalaceae. Taken as a whole, it is worth considering possible phylogenetic relationships between *Schoepfia* and genera of Loranthaceae that are root parasites. Biogeographic information and the long branch connecting *Schoepfia* with *Misodendrum* suggest the latter genus may represent a relictual taxon that evolved on the Gondwanan landmass along with the aerial parasites of Loranthaceae. The three genera of Loranthaceae used for the combined analysis form a clade (100% bootstrap) that is sister to the *Misodendrum/Schoepfia* clade. Analysis of a larger 18S data set (23 of the 74 genera, data not shown) resolves two clades, one composed of New World and the other of Old World mistletoes (Nickrent and Duff, 1996). A more rapidly evolving gene is required to provide sufficient phylogenetic signal to address relationships among all genera of this family. The combined analysis indicates that Opiliaceae form a well-supported (100% bootstrap) clade that is sister to the remaining members of the order. In contrast to traditional classifications, sequence data indicate that Opiliaceae are distinct from and evolutionarily more derived than Olacaceae. Using *rbcL* data alone, Opiliaceae are intercalated between a group of New World and Old World Santalaceae (the latter including Eremolepidaceae).

Analyses of 18S rDNA and *rbcL* sequences separately and in combination show that Santalaceae are not monophyletic but a grade that culminates in Viscaceae (Fig. 8.5). Two groups of Santalaceae are resolved representing mainly New World genera (*Acanthosyris, Buckley, Jodina,* and *Pyrularia*) and Old World genera (*Exocarpos, Osyris,* and *Santalum*) plus Eremolepidaceae (*Antidaphne, Eubrachion,* and *Lepidoceras*). These latter three genera are not monophyletic using either molecular data set, thus supporting the concept of Wiens and Barlow (1971) who used karyological and morphological evidence to suggest that Eremolepidaceae are aerial parasites of Santalaceae. Viscaceae are the most derived family of the order, and bootstrap support for the clade is high using either 18S or *rbcL* sequences. Molecular analyses indicate that Viscaceae is distinct from Loranthaceae, as was previously shown by data derived from biogeography, floral morphology, and cytology (Barlow, 1964; Barlow and Wiens, 1971; Wiens and Barlow, 1971; Barlow, 1983). Strongly supported intergeneric relationships are obtained only for *Phoradendron/Dendrophthora* and *Korthalsella/Ginalloa.* Despite complete generic sampling in the family, 18S and *rbcL* sequence data, and accelerated substitution rates, relationships among all genera are not fully resolved, suggesting that the evolution of the viscaceous genera may represent a "hard polytomy" (Maddison, 1989), that is, a true rapid radiation.

The foregoing discussion demonstrates that analyses of a nuclear and plastid gene are concordant in resolving relationships within Santalales. In addition, hypotheses proposed by the three traditional classification systems can now be evaluated in light of new evidence. For example, it is apparent that aerial parasitism has evolved in Santalales in at least four separate

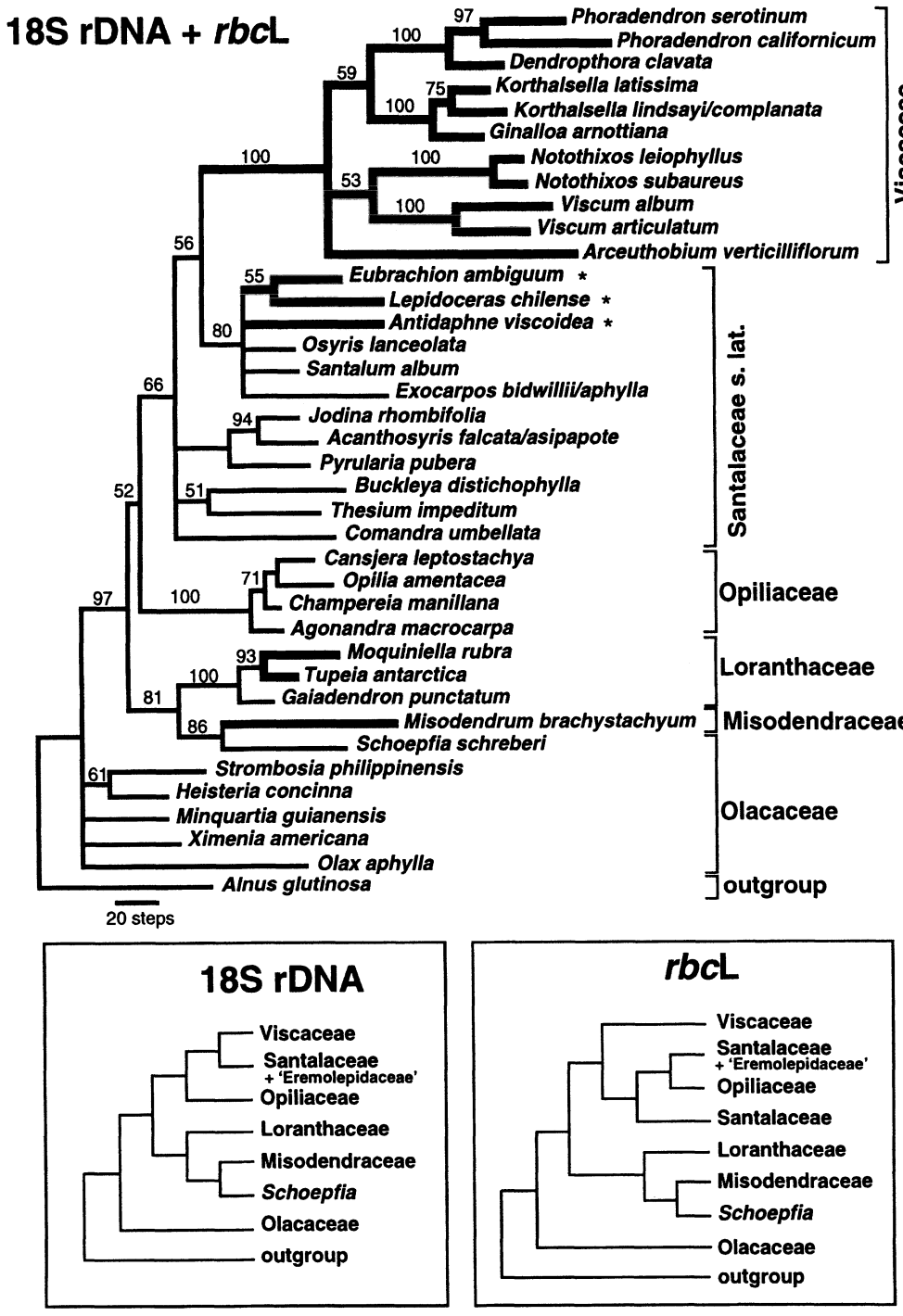

Figure 8.5. Strict consensus phylogram of 11 trees of length 2,212 derived from a parsimony analysis of nuclear 18S rDNA and *rbcL* sequences combined from 37 taxa of Santalales and *Alnus* (outgroup). For taxa with two specific epithets, 18S rDNA was sequenced from the first species and *rbcL* from the second. The phylograms enclosed as insets indicate the topologies of trees derived from analyses of the 18S rDNA and *rbcL* sequences separately. Numbers above the nodes indicate bootstrap values (from 100 replications). Nodes without values were supported in less than 50% of the replications. See Table 8.2 for additional details pertaining to this analysis. The genera indicated by an asterisk have traditionally been classified in Eremolepidaceae but are here considered part of Santalaceae s. l. The multiple origins of aerial parasitism are indicated by taxa with thick lines.

lineages (Misodendraceae, Loranthaceae, "Eremolepidaceae," and Viscaceae; thick lines on Fig. 8.5). Aside from such phylogenetic implications, what molecular evolutionary insights might be gained by comparing Santalales to other parasitic plants? In contrast to Scrophulariaceae, the holoparasitic trophic mode has apparently never evolved among the diverse lineages of root and stem parasitic Santalales. Although Olacaceae have been proposed as the evolutionary origination point for Balanophoraceae (Cronquist, 1981), such an association is not supported by nuclear 18S rDNA sequence analyses (see next section). If such a relationship is real, then no intermediate (less derived) taxa exist that might provide a link between the two families. It is noteworthy that, like Balanophoraceae, base substitution rates in 18S rDNA are increased among Viscaceae in general and *Arceuthobium* in particular. For the latter genus, base substitution rates are also increased for *rbcL*. Significantly, only about 30% of the carbon needed for growth is fixed in shoot tissue of *Arceuthobium* (Hull and Leonard, 1964). Because these parasites are obtaining a large portion of their photosynthate from the host, can we expect the eventual evolution of holoparasitism? It appears that such a trophic mode is not compatible with the aerially parasitic habit. Seedlings of *Arceuthobium* (Tocher et al., 1984) and those of other mistletoes are actively photosynthetic and must exist autotrophically prior to attachment to the host branch. Such is not the case for root holoparasites that germinate in response to chemical stimulants (carried in a moist rhizosphere) and attach to closely proximal host roots (Press and Graves, 1995). Mistletoe seeds do not undergo dormancy, are not adapted to long-term storage, and germinate in the absence of chemical cues; therefore, attachment to the host must be rapid. These features impose critical evolutionary constraints upon mistletoes that do not exist for root holoparasites.

Phylogenetic Studies of the Nonasterid Holoparasites Using Nuclear 18S rDNA Sequences

As discussed previously, traditional means of classifying the nonasterid holoparasites have met with difficulty owing to the extreme reduction and/or modification of morphological structures that have accompanied the evolution of these lineages. Although sequences of *rbcL* have proved extremely valuable in elucidating relationships throughout the angiosperms (see Chapter 17), this gene does not amplify for the nonasterid holoparasites using standard PCR procedures (D. Nickrent, unpubl.). Because nuclear 18S rDNA sequences have proved useful for addressing phylogenetic relationships in other parasitic plants, data from this molecular marker were obtained for representatives of all nonasterid holoparasites (Fig. 8.1). Relative rate tests showed that 18S sequences in these holoparasites were evolving, on average, 3.5 times faster than nonparasitic and most hemiparasitic plants (Nickrent and Starr, 1994), thus complicating their use in phylogenetic analyses. When these divergent sequences were included in analyses with nonparasites, the resulting topologies frequently showed aberrant relationships attributed to "long-branch attractions" (Felsenstein, 1978). Similar results were seen using parsimony, neighbor-joining, and maximum likelihood methods.

Long-branch artifacts are sometimes the result of incomplete sampling where intermediate lineages that could serve to "break up" a long branch have been omitted (Chase, Soltis, Olmstead et al., 1993). Recently, increased taxon density for nuclear 18S rDNA has been achieved such that, at present, more than 400 complete nuclear 18S sequences exist for angiosperms (Nickrent and Soltis, 1995; Soltis, Soltis, Nickrent et al., 1997). A 223-taxon data set (that included a greater sampling of monosulcate taxa than previous studies) allowed a reexamination of the placement of the nonasterid holoparasites. Sequences from the following holoparasites were added to the alignment: Balanophorales (*Balanophora fungosa, Corynaea crassa, Helosis cayennensis, Ombrophytum subterraneum, Rhopalocnemis phalloides,* and *Scybalium jaimaicense*), Cynomoriales (*Cynomorium coccineum*), Cytinales (*Cytninus ruber*), Rafflesiales (*Rafflesia keithii* and *Rhizanthes zippelii*), and Hydnorales (*Hydnora africana* and *Prosopanche americana*). Given the size of this data set, only heuristic search strategies could be used and branch swapping did not go to completion.

The tree resulting from this analysis (not shown) retained the major topological features found in the more extensive analysis reported in Soltis, Soltis, Nickrent et al. (1997). None of the nonasterid holoparasites were associated with Santalales. As seen in previous analyses, Rafflesiales and Balanophorales exhibited long-branch attractions, for example *Rafflesia* and *Rhizanthes* were intercalated in the middle of Balanophorales—an unlikely relationship. This composite clade was placed with taxa allied with Saxifragales, an equally unlikely relationship. This analysis showed that *Cytinus* and *Rafflesia* were only distantly related, thus supporting the segregation of Cytinaceae as a distinct family of Rosidae s. l. (Takhtajan et al., 1985; Takhtajan, 1987). The disposition of the remaining families of Rafflesiales has been hampered by the lack of nuclear 18S rDNA sequences from Apodanthaceae, Mitrastemonaceae, and *Bdallophyton.* 18S rDNA sequences of representatives of Balanophorales show that Cynomoriaceae differ markedly (more than 100 steps) from other members of the order, thus supporting the segregation of this family as proposed by Tahktajan (1987) and Thorne (1992).

In large- and small-scale analyses, Hydnoraceae are consistently placed as sister to the "paleoherb I" clade (sensu Donoghue and Doyle, 1989), that is, Aristolochiaceae, Chloranthaceae, Lactoridaceae, Piperaceae, all members of Magnoliidae. In contrast to Balanophorales and Rafflesiales, 18S rDNA rate increases for Hydnoraceae are less than half that seen in other holoparasites such as *Rafflesia* and *Balanophora;* hence, long-branch attraction appears to be less problematic. Furthermore, when taxa such as *Rafflesia* and *Balanophora* are included in the same analysis as Hydnoraceae, the association of the latter family with the paleoherbs is retained. The relationship between Hydnoraceae and Aristolochiaceae is concordant with the concepts proposed by Solms-Laubach (1894), Harms (1935), and more recently Cocucci (1983). The results with Hydnoraceae demonstrate the utility of molecular phylogenetic methods in providing independent evidence useful in resolving long-standing problems in the higher-level classification of these unusual holoparasites.

MOLECULAR EVOLUTIONARY STUDIES OF PARASITIC PLANTS

Molecular evolutionary processes occurring in parasitic plants represent some of the most extreme and dynamic known among all angiosperms. As evidenced by advances made from studying *Epifagus, Conopholis,* and *Cuscuta,* these plants can serve as important genetic models for facilitating the characterization of genome evolution. It is therefore worthwhile to compare and contrast genetic processes found in evolutionarily distinct groups in an attempt to uncover common themes, mechanisms, and evolutionary trends. As shown below for RUBISCO, fine-scale examination of the types of mutations that have occurred at the molecular level help illuminate the progression of change among diverse members of Scrophulariaceae.

Structural Analysis of RUBISCO Large Subunit Protein in Scrophulariaceae

The structure and function of the RUBISCO large subunit has been explored in great depth (Roy and Nierzwicki-Bauer, 1991; van der Vies et al., 1992; Schreuder et al., 1993; Adam, 1995; Kellogg and Juliano, 1997 and references therein). A comparative analysis of the amino acid substitutions of 499 flowering plants (Kellogg and Juliano, 1997) reveals a conservative mode of evolution of the protein in that there is a high degree of toggling among biochemically similar amino acids along the length of the polypeptide. The important structural motifs are particularly conserved in most flowering plants. To examine the effect of relaxed functional constraints on the evolution of the RUBISCO large subunit in parasitic Scrophulariaceae, several approaches were used, including: (1) an analysis of the synonymous and nonsynonymous substitution rates (Table 8.3); (2) mapping amino acid substitutions onto the strict consensus tree using MacClade 3.0 (Maddison and Maddison, 1992); (3) a PAM250 analysis (Dayhoff et al., 1978); and (4) assessment of substitutions in important structural motifs (Table 8.4); Wolfe and dePamphilis, 1997).

Average synonymous (K_s) and nonsynonymous (K_n) substitution rates and K_s/K_n for *rbcL*

Table 8.3. Average substitution rates of *rbcL* sequences for parasitic Scrophulariaceae and average PAM250 scores for RUBISCO large subunit polypeptide.

Parasitic Plant Category	No. Species	K_S[a]	K_N[a]	K_S/K_N	PAM250[b]
Nonparasitic	16	0.2022	0.0281	7.19	1113
Photosynthetic hemiparasitic	7	0.2227	0.0223	9.98	1124
Nonphotosynthetic[c] hemiparasitic	11	0.2463	0.0270	9.12	1112
Holoparasitic[d]	6	0.2482	0.0450	5.52	902

[a]Synonymous (K_S) and nonsynonymous (K_N) substitution rates based on Jukes-Cantor corrections determined using the computer program MEGA (Kumar et al., 1993).

[b]*Nicotiana tabacum* as a reference protein.

[c] Hemiparasitic lineages with holoparasitic members and holoparasitic *rbcL* ORFs (see Fig. 8.3).

[d] *rbcL* pseudogene sequences only.

from different categories of parasitic Scrophulariaceae are listed in Table 8.3. As heterotrophy increases in Scrophulariaceae, K_s also increases. In contrast, K_n does not appear to be affected greatly until the selective constraint is removed and pseudogenes are formed. There is a bias toward nonsynonymous substitutions in parasitic plants with *rbcL* open reading frames, as evidenced by the higher K_s/K_n ratios compared to nonparasitic Scrophulariaceae. However, the average K_s/K_n for *rbcL* pseudogene sequences is much lower, suggesting a relatively higher rate of nonsynonymous substitutions after selection is no longer a factor.

The *rbcL* nucleotide sequences were translated to inferred protein sequences. Deletions were maintained in the pseudogene sequences, but insertions were removed to retain the accuracy of the amino acid alignment. The step changes along the length of the polypeptide were plotted against the strict consensus tree based on *rbcL* sequences using MacClade. The primary finding of this analysis is the apparent loss of conserved sequences along the RUBISCO large subunit polypeptide with increasing heterotrophy. For nonparasitic Schrophulariaceae, there are long sequences in which no amino acid replacements are observed. This

Table 8.4. PAM250 analysis of important *rbcL* structural motifs for representative Scrophulariaceae.[a]

		Structural Motifs							
Taxon	No. of Events	α	β	Intradimer	Between Subunits	Dimer–dimer	Active Site	Other	PAM250 Difference
Holoparasites									
Orobanche ramosa	20	6	1	3	2	2	0	9	103
Boschniakia hookeri	13	4	2	0	3	0	1	4	64
Boschniakia strobilacea	10	2	2	0	2	0	1	3	57
Hyobanche sanguinea	7	3	1	1	1	1	0	3	41
Hyobanche atropurpurea	6	4	1	1	0	0	1	1	25
Other Holoparasites									
Lathraea clandestina	6	1	0	1	0	1	0	4	32
Orobanche corymbosa	2	0	1	0	0	0	0	1	5
Orobanche fasciculata	1	0	0	0	0	0	0	1	2
Hemiparasites									
Alectra sessiliflora	1	0	0	0	0	0	0	1	6
Striga hermonthica	1	0	1	0	0	0	0	0	8
Nonparasites									
Alonsoa unilabiata	1	0	1	0	0	0	0	0	3
Chelone obliqua	1	1	0	0	0	0	0	0	5

[a]See text for further explanation and interpretation.

same pattern was found by Kellogg and Juliano (1997) for 499 *rbcL* sequences from the Chase, Soltis, Olmstead et al. (1993) study. Kellogg and Juliano (1997) noted that important structural domains were conserved across a wide sampling of flowering plants and that there were additional conserved sequences of no known function. Hemi- and holoparasitic plants of Scrophulariaceae have fewer conserved regions across the length of the RUBISCO large subunit polypeptide. However, it is important to note that many biochemically similar amino acid replacements (e.g., replacing leucine with isoleucine) do not affect the structure and/or function of the protein and are, therefore, neutral replacements. Plotting the amino acid replacements against the strict consensus tree gives a first indication of molecular evolution of the protein, but does not elucidate the effect of molecular evolution on the structure and/or function of the protein.

To assess the effect of amino acid replacements on the structure and function of the RUBISCO large subunit polypeptide, Wolfe and dePamphilis (1995, 1997) conducted a PAM250 analysis (Dayhoff et al., 1978) on the translated protein sequences for selected taxa from all categories of parasitic plants in Schrophulariaceae. Dayhoff et al. (1978), using empirical data for many proteins (hemoglobins, fibrinopeptides, cytochrome c), constructed a table (the PAM250 matrix) of amino acid substitution probabilities; in other words, a table that shows the probabilities of the substitution of one amino acid for another. Amino acids that have similar biochemical properties are more likely to be substituted with one another than with biochemically dissimilar amino acids, and this is reflected in the values given in the PAM250 matrix. These substitution probabilities are measured in terms of PAM (percent accepted mutations) with one unit representing one amino acid substitution per 100 residues.

Using the reference protein from the outgroup, *Nicotiana tabacum,* PAM250 scores were calculated over the entire length of the polypeptide and from the critical structural motifs (Tables 8.3, 8.4) by assigning a PAM250 value for each amino acid of *Nicotiana* and comparing the PAM250 score for each amino acid substitution in the other taxa. The total PAM250 scores for the entire polypeptide and for each critical structural motif were compared between the reference protein and each taxon. Although the amount of amino acid replacement increases with increasing heterotrophy, the overall effect of these replacements on the protein structure as inferred from the PAM250 analysis is minimal. The only PAM250 scores significantly different from the reference protein are pseudogene sequences. The translated protein sequences from the holoparasite *rbcL* ORFs are nearly identical to those of photosynthetic plants in terms of their inferred structure. Hence, despite increased synonymous and nonsynonymous substitutions in nonphotosynthetic hemiparasites, the functionality of RUBISCO does not appear to be negatively impacted, which implies that RUBISCO is functional in parasitic plants even after the loss of photosynthesis (see next section).

Plastome Reduction in Holoparasitic Plants

For those holoparasites whose plastomes have been more fully characterized at the sequence level, the loss of photosynthesis is accompanied by the loss or modification of photosynthetic genes, a process that reduces the overall genome size. Empirical evidence for genome reduction can be seen in Scrophulariaceae s. l., *Cuscuta,* and the nonasterid holoparasites. The genes responsible for plastid gene expression (rDNA, ribosomal proteins) are most often retained, as evidenced by data on more extreme holoparasites such as *Cytinus.* In *Epifagus,* plastid-encoded tRNA genes appear more labile as evidenced by their loss and apparent replacement by nuclear-encoded tRNAs (Wolfe et al., 1992b). Can the reductional trend be taken to the extreme where the plastid and/or its plastome is lost entirely? Based on ultrastructural studies of holoparasite plastids, Dodge and Lawes (1974) concluded that plastids are retained as storage organs or for other functions. Because the plastome of *Epifagus* contains only four genes that are not involved in gene expression (dePamphilis and Palmer, 1990; Wolfe et al., 1992c), any one may serve some indispensable function in the

nonphotosynthetic plastid, thus explaining the retention of the entire plastome.

Studies of the *rbcL* locus in holoparasitic Scrophulariaceae (Fig. 8.6) are especially informative in that they show a range in the degree of structural modifications experienced by different lineages, thereby implying that such changes occur independently and result in different degrees of pseudogene formation. It is of interest that the distribution of *rbcL* pseudogenes on the phylogeny of Scrophulariaceae (Fig. 8.3) is different among Old and New World species of *Orobanche* (the Old World species *Orobanche cernua* and *O. ramosa* have pseudogenes, whereas the New World *O. corymbosa* and *O. fasciculata* do not). It is therefore important to determine not just the presence of *rbcL* genes but also whether they are expressed in holoparasites.

It is clear from recent studies of parasitic Scrophulariaceae and Cuscutaceae (Machado and Zetsche, 1990; Haberhausen et al., 1992; Delavault et al., 1995; Wolfe and dePamphilis, 1997) that *rbcL* ORFs are maintained in many independent lineages of parasitic plants. In two holoparasites, *Lathraea* (Bricaud et al., 1986; Thalouarn et al., 1989; Thalouarn et al., 1991; Delavault et al., 1995, 1996) and *Cuscuta* (Press et al., 1986, 1991; Machado and Zetsche, 1990; Haberhausen and Zetsche, 1992), the *rbcL* locus is minimally expressed in terms of gene transcription and/or protein activity. No *rbcL* expression was detected in several species of *Orobanche* (Thalouarn et al., 1994); however, such studies have not been conducted on species with intact *rbcL* ORFs (e.g., *O. corymbosa* or *O. fasciculata*). An examination of the 5′ and 3′ untranslated regions (UTRs) of the *rbcL* locus in *Orobanche* (Wolfe and dePamphilis, 1997) revealed that the promoter and/or ribosome binding sites are maintained in *O. corymbosa* and *O. fasciculata,* but the 5′ UTRs of the pseudogene sequences have structural mutations or nucleotide substitutions. The 3′ UTRs of all species of *Orobanche* examined to date (Wolfe and dePamphilis, 1997) have intact stem-loop structures but have major deletions upstream of the palindromic sequence. Phylogenetic analyses, gene expression studies, and sequence analyses of *rbcL* all suggest that the retention of *rbcL* ORFs in *O. corymbosa, O. fasciculata,* and in other lineages of holoparasitic Scrophulariaceae can be explained by the recency of their transition to holoparasitism.

Wolfe and dePamphilis (1997) presented three hypotheses to explain the retention of *rbcL*

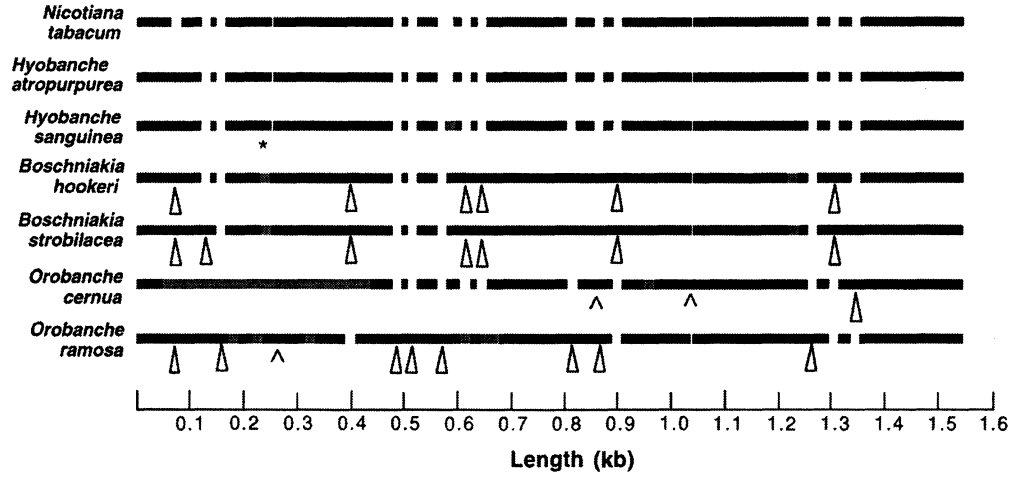

Figure 8.6. Structural maps of *rbcL* pseudogenes from orobanchaceous holoparasites compared with *Nicotiana tabacum.* Solid black lines indicate sequence not affected by structural changes in most other taxa. Blank regions demarcate insertions from other taxa. Open triangles represent multinucleotide insertions, ^ = single nucleotide insertions. Gray lines are multinucleotide deletions, * = single nucleotide deletion. Note that the *rbcL* locus of *Hyobanche atropurpurea* lacks structural mutations, but is a pseudogene, presumably because of a premature stop codon. *Conopholis* and *Epifagus rbcL* loci are not shown (missing in *Conopholis,* unalignable fragment in *Epifagus*).

in holoparasitic plants of Scrophulariaceae and Cuscutaceae: (1) the plant still requires the RUBISCO protein for low levels of autotrophic carbon fixation; (2) the oxygenase activity may function in glycolate metabolism; and (3) its maintenance or loss is a result of stochastic events. For the latter, the presence or absence of the gene in closely related members of a lineage may simply reflect the fact that insufficient time has passed for the accumulation of deleterious mutations. Whereas the carboxylase activity of RUBISCO fixes carbon dioxide into reduced carbon (which is subsequently metabolized by the plant), the oxygenase activity is a competitive reaction involved in glycine and serine biosynthesis (glyoxylate metabolism). Holoparasitic plants receive reduced carbon from their host, eliminating the need for RUBISCO activity (Press et al., 1991); however, some species maintain very low levels of photosynthetic pigment production and photosynthesis, or perform photosynthesis only during a discrete phase of the life cycle such as seedling development or host attachment. This may explain the retention of a minimally expressed *rbcL* in the stem parasite, *Cuscuta reflexa* (Haberhausen et al., 1992). It is unclear how important the glycolate pathway is for amino acid biosynthesis in photosynthetic or parasitic plants. Press et al. (1986) reported a decrease in photorespiration with increasing heterotrophy. Although unlikely, it is possible that the oxygenase activity of RUBISCO is maintained in parasitic plants until the adaptation to heterotrophy is complete, and acquisition of host amino acids reduces the necessity for photorespiration. Glycolate metabolism has also been suggested as a possible cause for the retention of *rbcL* in the plastome of the nonphotosynthetic euglenoid *Astasia longa* which has experienced deletions of nearly all other photosynthetic genes (Gockel et al., 1994).

To determine whether stochastic factors were the sole reason for maintenance of *rbcL* ORFs in *O. corymbosa* and *O. fasciculata,* Wolfe and dePamphilis (1997) developed a probability model using the following assumptions: (1) *O. corymbosa* and *O. fasciculata* had a nonphotosynthetic ancestor, and all *rbcL* sequence divergence occurred without photosynthetic constraint; (2) the probability of random mutations generating a stop codon is 4/63. Under these assumptions, and if all mutations accumulating in the *rbcL* sequences are independent, then, if the sequences are pseudogenes, the probability of retaining an ORF without a stop codon is: $P = (59/63)^n$, where n is the number of nucleotide substitution differences between them. Using this probability model, Wolfe and dePamphilis (1997) calculated a probability of 2.79%; therefore it is unlikely that chance alone explains the maintenance of the *rbcL* ORF in *O. corymbosa* and *O. fasciculata.*

As suggested by the differential occurrence of pseudogenes in different species of *Orobanche,* the short time scale for evolutionary diversification may explain differences in the plastome structure of *Epifagus* and *Conopholis,* shown to be sister taxa by nuclear 18S rDNA, plastid 16S rDNA, and *rps2* data. The 16S/23S rDNA spacer of *Epifagus* contains two nearly full-length tRNA pseudogenes whereas these genes are essentially absent in the spacer of *Conopholis,* thus suggesting a recent loss. Moreover, both 18S rDNA and *rps2* data indicate that *Boschniakia* is sister to the *Epifagus/Conopholis* clade, hence the ancestor of the latter two taxa was also a holoparasite (whose plastome must have contained inverted repeats). Therefore, the loss of one copy of the inverted repeat in *Conopholis* took place after its divergence from *Epifagus,* likely less than 5 myr ago (Muller, 1981). The case of the malarial *Plasmodium* (an apicomplexan) contrasts with such rapid plastome reorganizations. Apicomplexans are thought to have evolved from photosynthetic chromophytes, hence the malarial plastome has been retained for possibly 800 myr or more (Escalante and Ayala, 1995). As with holoparasitic plants, the plastome of *Plasmodium* has a compact organization and retains a specific suite of genes involved in gene expression *and* inverted repeats. These two examples demonstrate the complexity of selectional factors involved in either maintaining or deleting major structural features of the plastome. Although the selective environments of malarial parasites and plants are clearly quite different, both have followed parallel pathways in shaping the overall plastome structure, and both suggest that this genome may contribute essential

metabolic products required by the parasite (Wilson et al., 1996).

RATE ACCELERATIONS

One of the fundamental questions in evolutionary biology centers around the concept of the molecular clock, specifically the application of a strict molecular clock that infers equal rates of nucleotide substitution for organisms with equal generation times (the generation-time effect). As has been shown for *rbcL* in monocots (Bousquet et al., 1992; Gaut et al., 1992), *rps2* in parasitic Scrophulariaceae (dePamphilis et al., 1997), and 18S rDNA in Santalales (Nickrent and Starr, 1994), a strict molecular clock cannot be universally applied, at least for these loci and taxa. As discussed below in Rate Variation among Genomes and Lineages, rate accelerations occur in genes from all three subcellular genomes in parasitic plants, hence these organisms provide ideal models for studying molecular evolutionary questions such as the effect of rate heterogeneity on phylogeny reconstruction.

Nuclear 18S rRNA Genes

The first documentation of rate acceleration in nuclear 18S rDNA in plants was the study by Nickrent and Starr (1994) wherein relative rates tests were conducted using one hemiparasite (*Arceuthobium*) and four holoparasites (*Balanophora, Prosopanche, Rafflesia,* and *Rhizanthes*). These tests showed that 18S rDNA sequences in these holoparasites were evolving, on average, at least three times faster than nonparasitic and most hemiparasitic plants. Rate heterogeneity of this magnitude has not been detected among over 200 angiosperm sequences that have been examined (D. Nickrent, unpubl.); however, statistically significant rate increases have been measured in some Scrophulariaceae, *Cuscuta, Pholisma* (Lennoaceae), and mycoheterotrophic angiosperms (A. Colwell and D. Nickrent, unpubl.). These four nonasterid holoparasites and the dwarf mistletoe represent four distinct orders, thus accelerated substitution rates have occurred independently in an ancestor of each of these lineages. In Santalales, accelerated rates are not universally seen, as many other nonparasites and hemiparasites show no indication of acceleration. These data suggest that substitution rate acceleration of nuclear rDNA occurs only in lineages where photosynthesis is absent or diminished, thereby requiring 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 rDNA loci, although relaxation of selectional constraints on rRNA structure and function, small effective population size, and molecular drive have been proposed (Nickrent and Starr, 1994).

Plastid Genes: rps2 and rbcL

Phylogenetic comparisons of *Epifagus* to *Nicotiana* and more distant outgroups revealed longer branches leading to *Epifagus* for virtually all of the retained protein-coding genes and plastid-encoded rRNA genes (Morden et al., 1991; Wolfe et al., 1992b, 1992c). Increased numbers of substitutions were observed at both synonymous and nonsynonymous sites between the parasitic and nonparasitic lineages. These studies provided the first indication that the plastome of *Epifagus* may be subject to increased rates of molecular evolution, even at functionally constrained genes. As shown on the *rps2* phylogeny (dePamphilis et al., 1997), the branch lengths for both *Epifagus* and *Conopholis* are considerably longer than those of other Scrophulariaceae, and formal relative rates tests show that both have increased evolutionary rates relative to nonparasitic and most hemiparasitic lineages (dePamphilis et al., 1997). Also evolving at an accelerated rate are the ptDNAs of *Orobanche ramosa* (nonsynonymous only), the *Buchnera-Striga-Cycnium* clade (nonsynonymous sites only), and *Euphrasia* (synonymous sites only). These complex patterns of rate acceleration indicate not only that rate asymmetries in this group are not uniquely tied to the loss of photosynthesis, but that they may have more than one distinct underlying cause, resulting in rate accelerations at synonymous and nonsynonymous sites.

For *rbcL,* synonymous and nonsynonymous substitution rates in nonparasitic and hemiparasitic (photosynthetic) Scrophulariaceae are very similar to those of all other flowering plants examined to date (Wolfe and dePamphilis, 1995,

1997). This similarity is probably the result of the strong functional constraints imposed on the protein to maintain important structural motifs such as the alpha-beta barrels forming the active site, the active site itself, intradimer and dimer–dimer interactions, and intersubunit interactions (Kellogg and Juliano, 1997). Parasitic plants that have lost photosynthesis are no longer under selective constraints to maintain the structure and function of proteins involved in photosynthetic reactions. This has become evident in examination of holoparasitic Scrophulariaceae. For example, the entire *rbcL* locus is missing in *Conopholis* (Colwell, 1994), and only a remnant of the gene remains in *Epifagus* (dePamphilis and Palmer, 1990; Wolfe et al., 1992c).

Plastid Genes: 16S rDNA

Plastid-encoded 16S rDNA sequences have seldom been used in phylogenetic studies of angiosperms given their overall sequence conservation (see Chapter 1). The 16S rDNA sequences from five nonasterid holoparasites were included in a phylogenetic analysis with other land plants, algae, and cyanobacteria to show their affinity to angiosperms and graphically demonstrate rate increases. The topology of the strict consensus phylogram (Fig. 8.7) retains the major features obtained by others using different methods of phylogenetic reconstruction such as distance-based methods and neighbor joining (summarized by Nelissen et al., 1995). Despite their extremely long branches, the holoparasite 16S rDNA sequences are most closely related to other angiosperm plastid 16S rDNAs. This position is significant because previous analyses using nuclear 18S rDNA sequences of extreme holoparasites resulted in the migration of these long-branch taxa to the base of the tree (Nickrent and Starr, 1994). Similar perturbations of the position of divergent taxa on mitochondrial

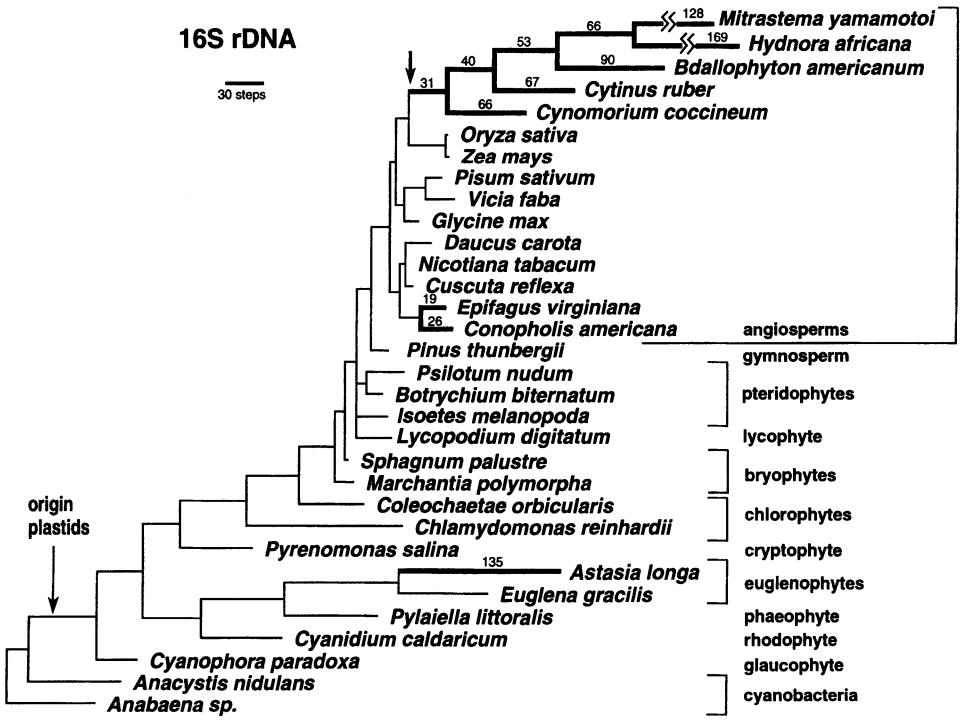

Figure 8.7. Strict consensus phylogram of four minimum-length trees of 2,576 steps derived from a parsimony analysis of 30 plastid-encoded 16S rDNA sequences and two cyanobacterial outgroup taxa (*Anacystis* and *Anabaena*). Numbers above the branches of the nonphotosynthetic taxa indicate number of substitutions (steps) and are shown in bold. The nonasterid holoparasite clade (arrow) is present within the angiosperms; however, intergeneric relationships are influenced by long-branch attraction artifacts.

16S rRNA trees has been observed (Olsen, 1987). For this character-based analysis (parsimony) of the holoparasites, the types of changes, not simply the number of changes, apparently determined the placement of the clade. On the other hand, interrelationships *among* the nonasterid holoparasites are likely artifactual because topological positions simply reflect an increasing number of substitutions.

Most angiosperm 16S rDNA sequences differ by 2–3% when compared with tobacco, whereas the nonasterid holoparasites show an increasingly greater number of mutations: *Cynomorium* (7.3%), *Cytinus* (8.0%), *Bdallophyton* (12.7%), *Mitrastema* (14.9%), *Hydnora* (19.4%), *Pilostyles* (30.4%), and *Corynaea* (35.9%) (Nickrent et al., 1997a). The high sequence variation suggests that these 16S rDNA sequences may possibly be pseudogenes. As with *Cytinus,* however, these rRNAs can be folded into secondary structures and most retain the typical complement of 50 helices (Nickrent et al., 1997a). The single exception is the 16S rRNA of *Pilostyles,* likely the most unusual 16S rRNA structure yet documented; it not only lacks four helices (9, 10, 11, and 37), but also contains large insertions on helices 6, 23-1, 29, and 48. Further molecular and biochemical work is needed to determine whether these genes are expressed and whether the rRNA is functional.

To quantify rate heterogeneity among various 16S rDNA sequences, formal relative rates tests were conducted. The parametric test described by Wu and Li (1985) as implemented for nuclear 18S rDNA sequences in holoparasites by Nickrent and Starr (1994) was used. In the absence of information on actual divergence times, this method employs a variance estimation to determine whether substitution rates (K) between two lineages differ. Because the time of divergence of the holoparasites relative to monocots and dicots is uncertain, *Marchantia* was chosen as the reference. Use of the phylogenetically closer *Pinus* as the reference did not change the results. The magnitude of rate increase among the nonasterid holoparasites matches the overall number of substitutional differences described above with *Cynomorium* at one extreme and *Hydnora* at the other (Fig. 8.8). The highly divergent sequences of *Pilostyles* and *Corynaea* would show K values even greater than *Hydnora* and are therefore not shown. The

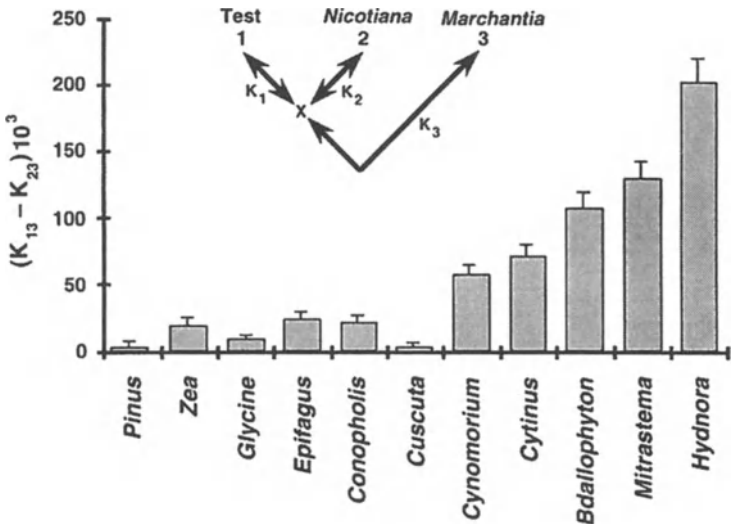

Figure 8.8. Histogram showing results of relative rate tests using plastid 16S rDNA sequences. The 11 land plants indicated on the abscissa are the test organisms (taxon 1), *Nicotiana* was taxon 2, and *Marchantia* was the reference or outgroup (taxon 3). The three-taxon tree uses K_1, the number of 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 1,000 for graphical purposes. Standard error values are included above the bar for each taxon. The fastest rates are observed in the five nonasterid holoparasites.

16S rDNA sequences of *Epifagus* and *Conopholis* exhibited slightly accelerated rates relative to most angiosperms (in agreement with Wolfe et al., 1992a), but not of the magnitude seen in the nonasterid holoparasites.

Mitochondrial rDNA

Increased nucleotide substitution rates among both nuclear 18S rDNA and plastid-encoded 16S rDNA of holoparasites raises the question "what is happening in the mitochondrial genome?" The plant mitochondrial genome has been largely ignored as a source of data for phylogenetic and population genetic studies of plants (see Chapter 1), partly because the rate of synonymous nucleotide substitution in this genome is only one-third to one-quarter that of chloroplast DNA, despite the high frequency of genetic recombination (Palmer, 1990, 1992).

Small subunit sequences of mitochondrial rDNA (here designated 19S rDNA) exhibit especially low levels of divergence relative to their chloroplast and nuclear counterparts in angiosperms. The six published photosynthetic angiosperm sequences (Table 8.5) exhibit less than 1.3% sequence divergence. This value was determined from the "core" of the molecule, that is, excluding two variable (and unalignable) regions (helices 6 and 43). Six new mitochondrial 19S rDNA sequences have been generated (Duff and Nickrent, submitted) that include four holoparasites (*Cytinus, Hydnora, Corenaea, Epifagus*), and *Nicotiana.* These sequences were aligned with six published sequences and compared with *Glycine* (Table 8.5). As with the 18S and 16S rDNA data, *Cytinus* exhibited the smallest degree of sequence divergence (2.3%) followed by *Cornaea* (2.8%) and *Hydnora* (3.1%). The moderate divergence of the *Epifagus* sequence (1.5%) is comparable to divergence levels of its plastid 16S rDNA.

Formal relative rates tests using mitochondrial 19S rDNA (similar to those performed on the 18S and 16S rDNA sequences) showed that only the nonasterid holoparasites had significantly increased substitution rates (Duff and Nickrent, 1997). Borderline rate increase was seen for *Epifagus.* In contrast, rates of sequence evolution for a second mitochondrial gene, *coxI,* are not elevated for *Epifagus* or for *Cytinus* relative to nonparasitic relatives (C. dePamphilis et al., unpubl.). Unlike the 18S and 16S rDNA sequences of *Cytinus, Hydnora,* and *Corynaea,* which exhibit base composition biases, the mitochondrial sequences of these three genera show no such bias. In addition to sequence divergence, these sequences are characterized by a higher frequency of indels than sequences of nonparasitic plants. Models constructed from these sequences indicate that the majority of mutations do not disrupt the higher-order structure, thus providing

Table 8.5. Angiosperm mitochondrial 19S rDNA sequences compared with *Glycine.*

Species	GenBank Number	Substitutions[a]	Indels[b] (bp)	Helix 6 (bp)	Helix 43 (bp)
Secale cereale	Z14049	18	0	151	366
Triticum aestivum	Z14078	18	0	151	344
Zea mays	X60794	20	1(14)	144	350
Oenothera berteriana	X61277	10	3(3)	106	338
Glycine max	M16859	—	—	102	427
Lepidoceras chilense	U82641	9	1(1)	103	364
Lupinus luteus	Z11512	11	1(1)	103	461
Nicotiana tabacum	U82638	15	0	102	335
Cytinus ruber[c]	U82639	33	4(10)	99	330
Epifagus virginiana[c]	U82642	22	2(2)	87	not sequenced
Hydnora africana[c]	U82637	41	3(4)	89	449
Corynaea crassa[c]	U82636	40	3(39)	75	271

[a]Substitutions from comparisons excluding helix 6 and 43.

[b]Frequency and size of indels were determined by sequence alignment aided by secondary structure analysis based on a model for *Zea* (Gutell, 1993).

[c]Holoparasitic taxa.

strong indirect evidence that these genes are functional.

The presence of RNA editing in plant mitochondria has recently received increased attention (Arts and Benne, 1996; Bowe and dePamphilis, 1996; Sper-Whitis et al., 1996). Although RNA editing is recognized as a significant process in many mitochondrial genes, the holoparasite 19S rDNA sequences reported here are unlikely to be significantly affected as determined by reverse-transcriptase PCR experiments (R. Duff and D. Nickrent, unpubl.). Despite the numerous reports of editing among protein-coding genes in plant mitochondria, no cases of editing of mitochondrial 195 rRNA are known (Schuster et al., 1991). Neither the pattern of substitutions nor the presence of indels follows that seen in reported cases of RNA editing. For these reasons, we do not feel editing can be used to explain the observed pattern of substitutions in mitochondrial 19S rDNA.

RATE VARIATION AMONG GENOMES AND LINEAGES

The patterns of nucleotide substitution across lineages and organellar gene loci provide a framework for addressing the underlying causes of rate variation. Relatively few studies exist in which rate comparisons are made between different organellar genes across lineages. An exception is recent work showing that synonymous (but not nonsynonymous) substitution rates are slower for palms than grasses for both nuclear-encoded *Adh* and plastid-encoded *rbcL* (Gaut et al., 1996). Three basic models of rate variation can be defined: (1) organismal, (2) genome-specific, and (3) gene-specific. The first model requires that DNA replication rate for each subcellular genome be correlated; hence if one locus is accelerated, all loci will be similarly affected. In the absence of selection, comparisons of molecular phylogenies using (say) two different organellar loci would show each taxon to have the same relative branch lengths on the respective trees. In the genome-specific model, there is no assumption that a correlation exists between DNA replication and/or substitution rates between organellar genomes. Such a model could result from differences in DNA replication and/or repair mechanisms in the distinct organellar genomes. Here, comparisons of molecular phylogenies derived from the different organellar loci would not be expected to show the same distribution of rate variation among taxa. Finally, the gene-specific model implies that each subcellular genome, and possibly individual genes within each genome, are constrained by their own selectional environment. This model would be manifest as molecular phylogenies with patterns of rate variation that are not correlated between gene loci.

Sequences of parasitic plant genes derived from all three subcellular genomes can be used to address the above models and to examine the association between rate changes and loss of photosynthesis. As shown in Table 8.6, the patterns of rate variation are complex; some ex-

Table 8.6. Rate comparisons for parasitic plants across genomes.

Taxon	Nuclear 18S rDNA	Plastid 16S rDNA	Plastid *rbcL*	Plastid *rps2*	Mitochondrial 19S rDNA
Nonasterid holoparasites	++[a]	++	absent	?	++
Santalales (minus Viscaceae)	−	−	−	?	−
Arceuthobium	++	+	+	?	?
Epifagus	+/−	+	ψ	+	+/−
Conopholis	+/−	+	absent	+	?
Hyobanche	+	−	ψ	−	?
Cuscuta	+[b]	−[c]	+[d]	?	−

[a]Molecular evolutionary rates: ++ = very fast, + = fast, +/− = borderline fast, − = not accelerated in relative rate tests, ψ = pseudogene.

[b]From *Cuscuta gronovii* (A. Colwell and D. Nickrent, unpubl.).

[c]From *Cuscuta reflexa* (Haberhausen and Zetsche, 1994).

[d]Gene from *Cuscuta reflexa* present but with low transcription levels (Haberhausen et al., 1992).

amples appear to support one model, whereas others do not. Consider first the nonasterid holoparasites. Formal relative rate tests, performed using sequences from nuclear, plastid, and mitochondrial small subunit rDNA, revealed that all three genomes are accelerated relative to autotrophic plants. These results support the organismal model of rate variation.

In other instances, the available data do not support the organismal model. Considering a particular gene locus, different substitution rates can be observed among different parasites of the same lineage. For 18S rDNA and *rbcL,* rates are elevated for Viscaceae, but not for other families of Santalales (Nickrent and Soltis, 1995; Nickrent and Duff, 1996), demonstrating that rate acceleration is not a general evolutionary feature of the order but has occurred only following the divergence of Viscaceae (see Fig. 8.5). Similarly, rate acceleration for *rps2* can be seen in *Epifagus, Conopholis,* and *Orobanche,* but not in other holoparasitic Scrophulariaceae such as *Hyobanche* (dePamphilis et al., 1997). Rate elevations may occur differentially among genes within a single genome. For example, *Arceuthobium* shows elevated rates for *rbcL,* but not for 16S rDNA, which indicates that elevated substitution rates can exist in genes of fully functional chloroplasts, but that conservative loci may lag behind more variable ones. In Scrophulariaceae, the pattern of rate increase (or pseudogene formation) for *rbcL* does not parallel that of 16S rDNA, again, with the latter showing more conservative rates of change. Finally, differences in substitution rates may be observed between different subcellular genomes within the same species. For *Arceuthobium, Hyobanche,* and *Cuscuta,* rates for nuclear but not plastid small subunit rDNA are elevated. These data do not tend to support either the organismal or the genome-specific model of rate variation, thus leaving only the gene-specific model as an explanation.

The relationship between the loss of photosynthesis and rate accelerations in nuclear and plastid genes can be interpreted using the data presented in Table 8.6. As shown by *Arceuthobium,* not all plants with rate increases (for the nuclear or plastid genomes) are nonphotosynthetic. Conversely, not all plants that lose photosynthesis show increased rates, as evidenced by *rps2* in *Alectra orobanchoides* or *Hyobanche.* Further support for the latter concept is seen following relative rate tests of nuclear rDNA in ericaceous mycotrophs (e.g., *Monotropa*) or nonphotosynthetic orchids (e.g., *Corallorhiza*) where no rate increases were observed (A. Colwell and D. Nickrent, unpubl.). Results from *rps2* analyses of Scrophulariaceae indicate that rate increases can occur in hemiparasites prior to the loss of photosynthesis (dePamphilis et al., 1997), thus providing evidence supporting the recency and rapidity of this trophic change.

The relationship between rate accelerations in plastomes and the nuclear genome is not clear at present. One might predict that changes in the plastome occur first and that the comparatively more conservative nuclear ribosomal loci follow. That the two events are linked is evidenced by the paucity of cases, at least in parasitic plants, where plastome genes are clearly accelerated but nuclear rDNA genes are not. Such correlations suggest a greater degree of signal transduction from the plastid to the nucleus than has previously been suspected (Susek and Chory, 1992).

CONCLUSIONS

Parasitic plants represent some of the most unusual organisms on earth, and each independent lineage represents a natural genetic experiment whose organelles and genes have evolved under relaxed functional constraints. Although *Cuscuta* is amenable to culture and direct experimental manipulation, very few holoparasitic parasitic plants have been established as model systems for biochemical and genetic research. Certainly, attempts to culture holoparasites would be hindered by lack of knowledge of their life histories. To date, studies of Santalales, nonasterid holoparasites, Schrophulariaceae s. l., and *Cuscuta* have yielded a wealth of new insights into molecular phylogenetic and molecular evolutionary processes. Despite this progress, a number of questions about the evolutionary process still remain, which will hopefully stimulate continued study of these fascinating plants.

LITERATURE CITED

Adam, Z. 1995. A mutation in the small subunit of ribulose-1,5-biphosphate carboxylase/oxygenase that reduces the rate of its incorporation into holoenzyme. Photosynthesis Research **43:**143–147.

An-Ming, L. 1990. A preliminary cladistic study of the families of the superorder Lamiiflorae. Botanical Journal of the Linnean Society **103:**39–57.

Arts, G. J., and R. Benne. 1996. Mechanism and evolution of RNA editing in kinetoplastida. Biochimica Biophysica Acta **1307:**39–54.

Barlow, B. A. 1964. Classification of the Loranthaceae and Viscaceae. Proceedings of the Linnaean Society of New South Wales **89:**268–272.

Barlow, B. A. 1983. Biogeography of Loranthaceae and Viscaceae. **In** The Biology of Mistletoes, eds. M. Calder and P. Bernhardt, pp. 19–45. Academic Press, New York.

Barlow, B. A., and D. Wiens. 1971. The cytogeography of the loranthaceous mistletoes. Taxon **20:**291–312.

Bhandari, N. N., and S. C. A. Vohra. 1983. Embryology and affinities of Viscaceae. **In** The Biology of Mistletoes, eds., M. Calder and P. Bernhardt, pp. 69–86. Academic Press, New York.

Boeshore, I. 1920. The morphological continuity of Scrophulariaceae and Orobanchaceae. Contributions of the Botanical Laboratory of the Morris Arboretum **5:**139–177.

Bömmer, D., G. Haberhausen, and K. Zetsche. 1993. A large deletion in the plastid DNA of the holoparasitic flowering plant *Cuscuta reflexa* concerning two ribosomal proteins (*rpl2, rpl23*), one transfer RNA (*trnI*) and an ORF 2280 homologue. Current Genetics **24:**171–176.

Bousquet, J., S. H. Strauss, A. H. Doerksen, and R. A. Price. 1992. Extensive variation in evolutionary rate of *rbcL* gene sequences among seed plants. Proceedings of the National Academy of Sciences U.S.A. **89:**7844–7848.

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

Bricaud, C. H., P. Thalouarn, and S. Renaudin. 1986. Ribulose 1,5-bisphosphate carboxylase activity in the holoparasite *Lathraea clandestina* L. Journal of Plant Physiology **125:**367–370.

Burtt, B. L. 1965. The transfer of *Cyrtandromoea* from Gesneriaceae to Scrophulariaceae, with notes on the classification of that family. Bulletin of the Botanical Survey **7:**73–88.

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. Hedrén, 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, and V. A. Albert. 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.

Cronquist, A. 1981. An Integrated System of Classification of Flowering Plants. Columbia University Press, New York.

Cronquist, A. 1988. The Evolution and Classification of Flowering Plants. New York Botanical Garden, New York.

Dayhoff, M. O., R. M. Schwartz, and B. C. Orcutt. 1978. A model of evolutionary change in proteins. **In** Atlas of Protein Sequence and Structure, ed. M. O. Dayhoff, pp. 345–352. National Biomedical Research Foundation, Silver Springs, Maryland.

Delavault, P., V. Sakanyan, and P. Thalouarn. 1995. Divergent evolution of two plastid genes, *rbcL* and *atpB,* in a non-photosynthetic parasitic plant. Plant Molecular Biology **29:**1071–1079.

Delavault, P. M., N. M. Russo, N. A. Lusson, and P. A. Thalouarn. 1996. Organization of the reduced plastid genome of *Lathraea clandestina,* an achlorophyllous parasitic plant. Physiologia Plantarum **96:** 674–682.

dePamphilis, C. W. 1995. Genes and genomes. **In** Parasitic Plants, eds. M. C. Press and J. D. Graves, pp. 176–205. Chapman & Hall, London.

dePamphilis, C. W., and J. D. Palmer. 1989. Evolution and function of plastid DNA: a review with special reference to nonphotosynthetic plants. **In** Physiology, Biochemistry, and Genetics of Nongreen Plastids, eds. C. D. Boyer, J. C. Shannon, and R. C. Hardison, pp. 182–202. American Society of Plant Physiologists, Rockville, Maryland.

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

dePamphilis, C. W., N. D. Young, and A. D. Wolfe. 1997. Evolution of plastid gene *rps2* in a lineage of hemiparasitic and holoparasitic plants: many losses of photosynthesis and complex patterns of rate variation. Proceedings of the National Academy of Sciences U.S.A. **94:**7367–7372.

Dodge, J. D. and G. B. Lawes. 1974. Plastid ultrastructure in some parasitic and semi-parasitic plants. Cytobiology **9:**1–9.

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

Downie, S. R., and J. D. Palmer. 1992. Use of chloroplast DNA rearrangements in reconstructing plant phylogeny. **In** Molecular Plant Systematics, eds. P. S.

Soltis, D. E. Soltis, and J. J. Doyle, pp. 14–35. Chapman & Hall, New York.

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

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

Eplee, R. E. 1981. Striga's status as a plant parasite in the United States. Plant Disease **56:**951–954.

Escalante, A. A., and F. J. Ayala. 1995. Evolutionary origin of *Plasmodium* and other Apicomplexa based on rRNA genes. Proceedings of the National Academy of Sciences U.S.A. **92:**5793–5797.

Fagerlind, F. 1948. Beitrage zür 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.

Freyer, R., K. Neckermann, R. M. Maier, and H. Kössel. 1995. Structural and functional analysis of plastid genomes from parasitic plants: loss of an intron within the genus *Cuscuta.* Current Genetics **27:**580–586.

Gaut, B., S. Muse, W. Clark, and M. Clegg. 1992. Relative rates of nucleotide substitution at the *rbcL* locus of monocotyledonous plants. Journal of Molecular Evolution **35:**292–303.

Gaut, B. S., B. R. Morton, B. C. McCaig, and M. T. Clegg. 1996. Substitution rate comparisons between grasses and palms: synonymous rate differences at the nuclear gene *Adh* parallel rate differences at the plastid gene *rbcL.* Proceedings of the National Academy of Sciences U.S.A. **93:**10274–10279.

Gockel, G., W. Hachtel, S. Baier, C. Fliss, and M. Henke. 1994. Genes for components of the chloroplast translational apparatus are conserved in the reduced 73-kb plastid DNA of the nonphotosynthetic euglenoid flagellate *Astasia longa.* Current Genetics **26:** 256–262.

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

Haberhausen, G., and K. Zetsche. 1992. Nucleotide sequences of the *rbcL* gene and the intergenic promoter region between the divergently transcribed *rbcL* and *atpB* genes in *Ipomoea purpurea* (L.). Plant Molecular Biology **18:**823–825.

Haberhausen, G., and K. Zetsche. 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.

Haberhausen, G. K., Valentin, and K. Zetsche. 1992. Organization and sequence of photosynthetic genes from the plastid genome of the holoparasitic flowering plant *Cuscuta reflexa.* Molecular and General Genetics **232:**154–161.

Harms, H. 1935. Hydnoraceae. In Die Natürlichen Pflanzenfamilien, eds. A. Engler and H. Harms, pp. 282–295. W. Engelmann, Leipzig.

Hong, D.-Y. 1984. Taxonomy and evolution of the Veroniceae (Scrophulariaceae) with special reference to palynology. Opera Botanica **75:**1–60.

Hull, R. J., and O. A. Leonard. 1964. Physiological aspects of parasitism in mistletoes (*Arceuthobium* and *Phoradendron*). 2. The photosynthetic capacity of mistletoes. Plant Physiology **39:**1008–1017.

Kellogg, E. A., and N. D. Juliano. 1997. The structure and function of RuBisCo and their implications for systematic studies. American Journal of Botany **84:**413–428.

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.

Kumar, S., K. Tamura, and M. Nei. 1993. MEGA: Molecular Evolutionary Genetics Analysis, version 1.01. The Pennsylvania State University, University Park.

Li, W.-H. 1983. Evolution of duplicate genes and pseudogenes. **In** Evolution of Genes and Proteins, eds. M. Nei and R. K. Koehn, pp. 14–37. Sinauer Associates, Sunderland, Massachusetts.

Machado, M. A., and K. Zetsche. 1990. A structural, functional and molecular analysis of plastids of the holoparasites *Cuscuta reflexa* and *Cuscuta europaea.* Planta **181:**91–96.

Maddison, W. P. 1989. Reconstructing character evolution on polytomous cladograms. Cladistics **5:**365–377.

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

Mayfield, S. P., C. B. Yohn, A. Cohen, and A. Danon. 1995. Regulation of chloroplast gene expression. Annual Review of Plant Physiology and Plant Molecular Biology **46:**147–166.

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

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

Muller, J. 1981. Fossil pollen records of extant angiosperms. Botanical Review **47:**1–142.

Musselman, L. J. 1980. The biology of *Striga, Orobanche,* and other root-parasitic weeds. Annual Review of Phytopathology **18:**463–489.

Nelissen, B., Y. Van de Peer, A. Wilmotte, and R. DeWachter. 1995. An early origin of plastids within the cyanobacterial divergence is suggested by evolutionary trees based on complete 16S rRNA sequences. Molecular Biology and Evolution **12:**1166–1173.

Nickrent, D. L., and R. J. Duff. 1996. Molecular studies of parasitic plants using ribosomal RNA. **In** Advances in Parasitic Plant Research, eds. M. T. Moreno, J. I. Cubero, D. Berner, D. Joel, L. J. Musselman, and

C. Parker, pp. 28–52. Junta de Andalucia, Dirección General de Investigación Agraria, Cordoba, Spain.

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

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

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

Nickrent, D. L., R. J. Duff, and D. A. M. Konings. 1997a. Structural analyses of plastid-encoded 16S rRNAs in holoparasitic angiosperms. Plant Molecular Biology **74:**731–743.

Nickrent, D. L., Y. Ouyang, R. J. Duff, and C. W. dePamphilis. 1997b. Do nonasterid holoparasitic flowering plants have plastid genomes? Plant Molecular Biology **34:**717–729.

Olmstead, R. G., and P. A. Reeves. 1995. Evidence for the polyphyly of the Scrophulariaceae based on chloroplast *rbcL* and *ndhF* sequences. Annals of the Missouri Botanical Garden **82:**176–193.

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

Palmer, J. D. 1990. Contrasting modes and tempos of genome evolution in land plant organelles. Trends in Genetics **6:**115–120.

Palmer, J. D. 1992. Mitochondrial DNA in plant systematics: applications and limitations. **In** Molecular Plant Systematics, eds. P. S. Soltis, D. E. Soltis, and J. J. Doyle, pp. 36–49. Chapman & Hall, New York.

Pazy, B., U. Plitmann, and O. Cohen. 1996. Bimodal karyotype in *Cynomorium coccineum* L. and its systematic implications. Journal of the Linnaean Society, London **120:**279–281.

Pennell, F. W. 1935. The Scrophulariaceae of Eastern Temperate North America. Academy of Natural Sciences, Philadelphia.

Press, M. C., and J. D. Graves. 1995. Parasitic Plants. Chapman & Hall, London.

Press, M. C., N. Shah, and G. R. Stewart. 1986. The parasitic habit: trends in metabolic reductionism. **In** Biology and Control of *Orobanche,* ed. S. J. ter Borg, pp. 96–106. LH/VPO, Wageningen, Netherlands.

Press, M. C., S. Smith, and G. R. Stewart. 1991. Carbon acquisition and assimilation in parasitic plants. Functional Ecology **5:**278–283.

Price, P. 1980. Evolutionary Biology of Parasites. Monographs in Population Biology No. 15. Princeton University Press, Princeton, New Jersey.

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

Roy, H., and S. A. Nierzwicki-Bauer. 1991. RuBisCO: genes, structure, assembly, and evolution. **In** The Photosynthetic Apparatus: Molecular Biology and Operation, eds. L. Bogorad and I. K. Vasil, pp. 347–364. Academic Press, San Diego.

Schreuder, H. A., S. Knight, P. M. G. Curmi, I. Andersson, D. Cascio, C.-I. Brändén, and D. Eisenberg. 1993. Formation of the active site of ribulose-1,5-bisphosphate carboxylase/oxygenase by a disorder-order transition from the unactivated to the activated form. Proceedings of the National Academy of Sciences U.S.A. **90:**9968–9972.

Schuster, W., R. Ternes, V. Knoop, R. Hiesel, B. Wissinger, and A. Brennicke. 1991. Distribution of RNA editing sites in *Oenothera* mitochondrial mRNAs and rRNAs. Current Genetics **20:**397–404.

Shimada, H., and M. Sugiura. 1991. Fine structural features of the chloroplast genome: comparison of the sequenced chloroplast genomes. Nucleic Acids Research **19:**983–995.

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

Solms-Laubach, H. 1894. Hydnoraceae. **In** Die Natürlichen Planzenfamilien, Part III, eds. A. Engler and K. Prantl, pp. 282–285. W. Engelmann, Leipzig.

Soltis, D. E., P. S. Soltis, D. L. Nickrent, L. A. Johnson, W. J. Hahn, S. B. Hoot, J. A. Sweere, R. K. Kuzoff, K. A. Kron, M. W. Chase, S. M. Swensen, E. A. Zimmer, S.-M. Chaw, L. J. Gillespie, W. J. Kress, and K. J. Sytsma. 1997. Angiosperm phylogeny inferred from 18S ribosomal DNA sequences. Annals of the Missouri Botanical Garden **84:**1–49.

Sper-Whitis, G. L., J. L. Moody, and J. C. Vaughn. 1996. Universality of mitochondrial RNA editing in cytochrome-c oxidase subunit I (*coxI*) among the land plants. Biochimica Biophysica Acta **1307:**301–308.

Stewart, G. R., and M. C. Press. 1990. The physiology and biochemistry of parasitic angiosperms. Annual Review of Plant Physiology and Plant Molecular Biology **41:**127–151.

Susek, R. E., and J. Chory. 1992. A tale of two genomes: role of a chloroplast signal in coordinating nuclear and plastid genome expression. Australian Journal of Plant Physiology **19:**387–399.

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

Takhtajan, A. L., N. R. Meyer, and V. N. Kosenko. 1985. Morfologiya pyl'tsy i klassifikatsiya semeystva Rafflesiaceae s. 1. Botanicheskii Zhurnal **70:**153–162.

Taylor, G. W., K. H. Wolfe, C. W. Morden, C. W. dePamphilis, and J. D. Palmer. 1991. Lack of a functional plastid $tRNA^{cys}$ gene is associated with loss of photosynthesis in a lineage of parasitic plants. Current Genetics **20:**515–518.

Thalouarn, P., M.-C. Arnaud, and S. Renaudin. 1989. Evidence of ribulose-bisphosphate carboxylase in the Scrophulariaceae holoparasite *Lathraea clandestina* L.

Comptes Rendu Academy of Science Paris **309:** 275–280.

Thalouarn, P., C. Theodet, and S. Renaudin. 1991. Evidence of plastid and nuclear genes for the large and small subunits of Rubisco in the Scrophulariaceae holoparasite *Lathraea clandestina* L. Comparison with the autotroph *Digitalis purpurea* L. and hemiparasite *Melampyrum pratense* L. Comptes Rendu Academy of Science Paris **312:**1–6.

Thalouarn, P., C. Theodet, N. Russo, and P. Delavault. 1994. The reduced plastid genome of a nonphotosynthetic angiosperm *Orobanche hederae* has retained the *rbcL* gene. Plant Physiology and Biochemistry **32:**233–242.

Thieret, J. W. 1967. Supraspecific classification in the Scrophulariceae: a review. Sida **3**87–106.

Thorne, R. F. 1992. An updated phylogenetic classification of the flowering plants. Aliso **13:**365–389.

Tocher, R. D., S. W. Gustafson, and D. M. Knutson. 1984. Water metabolism and seedling photosynthesis in dwarf mistletoes. **In** Biology of Dwarf Mistletoes: Proceedings of the Symposium, eds. F. G. Hawksworth and R. F. Scharpf, pp. 62–69. USDA, Rocky Mt. Forest & Range Experimental Station, Ft. Collins, Colorado.

van der Vies, S. M., P. V. Viitanen, A. A. Gatenby, G. H. Lorimer, and R. Jaenicke. 1992. Conformational states of ribulosebisphosphate carboxylase and their interaction with chaperonin 60. Biochemistry **31:**3635–3644.

Walsh, M. A., E. A. Rechel and T. M. Popovich. 1980. Observations of plastid fine structure in the holoparasitic angiosperm *Epifagus virginiana.* American Journal of Botany **67:**833–837.

Weber, H. C. 1980. Evolution of parasitism in Scrophulariceae and Orobanchaceae. Plant Systematics and Evolution **136:**217–232.

Wettstein, R. 1897. Scrophulariceae. **In** Die Natürlichen Pflanzenfamilien Nachtrage I–IV, eds. A. Engler and K. Prantl, pp. 293–299. W Engelmann, Leipzig.

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

Wilson, R. J. M., P. W. Denny, P. R. Preiser, K. Rangachari, K. Roberts, A. Roy, A. Whyte, M. Strath, D. J. Moore, P. W. Moore, and D. H. Williamson. 1996. Complete gene map of the plastid-like DNA of the malaria parasite *Plasmodium falciparum.* Journal of Molecular Biology **261:**155–172.

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

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

Wimpee, C. F., R. Morgan, and R. L. Wronel. 1992b. Loss of transfer RNA genes from the plastid 16S–23S ribosomal RNA gene spacer in a parasitic plant. Current Genetics **21:**417–422.

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

Wolfe, A. D., and C. W. dePamphilis. 1997. Alternate paths of evolution for the photosynthetic gene *rbcL* in four nonphotosynthetic species of *Orobanche.* Plant Molecular Biology **33:**965–977.

Wolfe, K. H., D. S. Katz-Downie, C. W. Morden, and J. D. Palmer. 1992a. 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.

Wolfe, K. H., C. W. Morden, S. C. Ems, and J. D. Palmer. 1992b. Rapid evolution of the plastid translational apparatus in a nonphotosynthetic plant: loss or accelerated sequence evolution of tRNA and ribosomal protein genes. Journal of Molecular Evolution **35:** 304–317.

Wolfe, K. H., C. W. Morden, and J. D. Palmer. 1992c. Function and evolution of a minimal plastid genome from a nonphotosynthetic parasitic plant. Proceedings of the National Academy of Sciences U.S.A. **89:** 10648–10652.

Wu, C.-I., and W.-H. Li. 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.