J Mol Evol (1994) 39:62–70

Journal of Molecular Evolution

High Rates of Nucleotide Substitution in Nuclear Small-Subunit (18S) rDNA from Holoparasitic Flowering Plants

D.L. Nickrent, E.M. Starr

Department of Plant Biology, Southern Illinois University, Carbondale, IL 62901, USA

Received: 10 June 1993 / Revised: 8 November 1993

Abstract. Relative rate tests, using *Gnetum* as a reference taxon, were conducted on nuclear 18S rRNA sequences from 10 angiosperms including autotrophic nonparasites (*Arabidopsis, Asarum, Glycine, Malpighia,* and *Zea*), a chlorophyllous hemiparasite (*Arceuthobium*—Viscaceae), and achlorophyllous holoparasites (*Balanophora*—Balanophoraceae, *Prosopanche*—Hydnoraceae, and *Rafflesia* and *Rhizanthes*—Rafflesiaceae). Compared with *Glycine,* the mean number of substitutions per site (K) for five autotrophic angiosperms is 0.036 whereas for the holoparasites $K = 0.126$, i.e., 3.5 times higher. Comparisons of autotrophic species with short and long generation times showed no differences in K; hence, divergent rRNA sequences in the holoparasites are likely attributable to other mechanisms. These might include genetic bottlenecks, effective population size, and/or molecular drive. High substitution rates appear to be associated only with those parasitic angiosperms that have developed a highly modified haustorial system and extreme nutritional dependence upon the host. At present, high substitution rates in these parasites confound attempts to determine their phylogenetic position relative to other angiosperms.

Key words: Holoparasite — Balanophoraceae — Hydnoraceae — Rafflesiaceae — Viscaceae — Ribosomal RNA — Relative rates — Plant molecular evolution

Correspondence to: D.L. Nickrent

Introduction

Parasitism apparently arose independently at least eight times during angiosperm evolution (Kuijt 1969) and the majority of species occur in two orders, Scrophulariales and Santalales. Within these orders, both hemiparasitic and holoparasitic forms exist. The hemiparasites are photosynthetic and attach to adjacent host plants by means of modified roots called haustoria which are structurally diverse but effect the physiological link between the host and parasite. Hemiparasites obtain mainly water and dissolved nutrients via the haustorial bridge, whereas holoparasites must obtain photosynthates. All hemiparasites are chlorophyllous and both obligate and facultative modes exist. Holoparasites are either totally or nearly achlorophyllous and all are therefore obligately parasitic.

The evolution of holoparasitic flowering plants has been accompanied by physiological, anatomical, and morphological adaptations unparalleled in other plants. Determining the phylogenetic position of these plants has been difficult owing to pervasive loss of organs such as roots, stems, and leaves, as has occurred in members of the Balanophoraceae, Hydnoraceae, and Rafflesiaceae. Such reductive trends have progressed to such an extreme that assessing homology among remaining morphological structures is often difficult or impossible. Coevolutionary associations between parasites and their host plants include the isophasic development of the haustorium of *Arceuthobium* (Viscaceae) or *Pilostyles* (Rafflesiaceae) which grow in synchrony with their host apical meristems or the chimeric (composite) vascular bundles that are found in the "tuber" in *Balanophora* (Gedalovich-Shedletzky and Kuijt 1990).

Certainly the most spectacular among the holoparasites is *Rafflesia,* noted for having the largest flower of any angiosperm (1 m in diameter). Both *Rafflesia* and *Rhizanthes* are parasites of lianas of the genus *Tetrastigma* (Vitaceae) that occur in the tropical rain forests of Malaysia, Indonesia, and the Philippines. Following seed germination and invasion of the host vine, the vegetative portions consist only of myceliumlike strands (the endophyte) within the host cortex. No true stems, leaves, or roots are produced at any stage of the life cycle. Eventually, portions of the endophyte differentiate into a floral meristem which then erupts from the host as a bud. Full flowering requires several additional months of development. These unusual and specialized parasites are in danger of extinction owing to habit loss and flower bud abortion (Ismail 1988).

Intense interest has recently been focused on plant phylogenetic relationships, mainly because of results obtained from sequencing the chloroplast-encoded *rbc*L gene (Albert et al. 1992; Chase et al. 1993). Preliminary attempts using polymerase chain reaction (PCR—Mullis and Faloona 1987) and Southern blots failed to detect this gene in representatives of the Balanophoraceae, Hydnoraceae, and Rafflesiaceae (Nickrent and dePamphilis, unpubl.). For this reason, attention was focused on the use of 18S (small-subunit) rRNA for molecular phylogenetic analysis. Large- and small-subunit ribosomal RNAs have been extensively utilized for phylogeny estimation in bacteria and archaea (Woese 1987) as well as eukaryotes (Sogin et al. 1986). For higher plants, rRNA has been used to study wide divergences among cycads, gnetophytes, angiosperms (Hamby and Zimmer 1992), and parasitic and nonparasitic plants (Nickrent and Franchina 1990). Features of rRNA that make it attractive for phylogenetic studies are its conserved primary and secondary structure, the mosaic pattern of variable and conserved regions, and overall rate homogeneity across lineages (Wolfe et al. 1989b).

The rates of synonymous nucleotide substitution at silent codon positions in plant nuclear genes generally vary by a factor of two and are about four times higher (5×10^{-9} subst./site/year) than in chloroplast genes (Wolfe et al. 1989b). This value was obtained by making intergeneric comparison of cultivated grasses (ca. 9,000 cpDNA sites and 6,500 nuclear sites) as well as comparisons utilizing two dicot families, the Solanaceae and Brassicaceae (3,100 cpDNA sites and 5,900 nuclear sites). To date, few studies have examined rate variation in plant nuclear ribosomal genes. Wolfe et al. (1989a) used partial 18S and 26S rRNA sequences for four angiosperm species but reported no rate heterogeneity.

As part of an ongoing effort to sequence ribosomal genes in parasitic angiosperms, we have obtained 18S rDNA sequences from over 50 hemiparasites and four holoparasites in the orders Santalales and Rafflesiales (sensu Cronquist 1988). We report here results of relative rate tests conducted on one chlorophyllous hemiparasite (*Arceuthobium*), four achlorophyllous holoparasites (*Balanophora, Prosopanche, Rafflesia,* and *Rhizanthes*) and selected nonparasites and discuss the factors that may contribute to rate differences. In this paper, substitution rate is defined as the number of nucleotide substitutions per site, not substitutions per site per year. The latter was not attempted since no independent information (e.g., fossil evidence) exists to calibrate the holoparasite divergences (see Discussion), although they are assumed to have arisen since the time of divergence of the Gnetales and angiosperms.

Materials and Methods

Acquisition of Plant Material. The plant species used for DNA extraction and associated collection information are listed in Table 1. Fresh tissue samples were used for all species except for *Rafflesia,* which was dried on silica gel. Voucher specimens are being maintained by D.L.N. or the original collector where noted. The 18S rRNA sequences for *Glycine max* (L.) Merr., *Arabidopsis thaliana* (L.) Hehnh., and *Zea mays* L. were previously determined (Eckenrode et al. 1985; Unfried et al. 1989; Messing et al. 1984, respectively).

DNA Extractions, PCR Amplification, and Sequencing. Plant tissues were ground to a powder on liquid nitrogen and total DNA was extracted following the CTAB method of Doyle and Doyle (1987). DNA resulting from extractions of some of the holoparasites were darkly pigmented and did not amplify initially. This was circumvented by running 20 µl of genomic DNA on a 2% TBE NuSieve agarose gel and excising the genomic DNA band. The agarose slice was diluted 1:1 with sterile water and heated to 65° C prior to PCR amplification.

A 100-µl PCR reaction mix consisted of 50 mM KCl, 10 mM Tris pH 8.2, 2.5 mM $MgCl_2$, 1.25 mM of each dNTP, 1.0 µM forward and reverse primers, and 2.5 units of Taq polymerase (Promega Corp.). The primers used for a 1.8-kb amplification were 25e Forward and 1769 Reverse (Table 2). Approximately 50 ng of genomic DNA was used as template for a symmetrical reaction. The reaction mix was overlaid with mineral oil and placed in a Perkin-Elmer thermal cycler. Cycling conditions consisted of an initial 3.0 min at 94° C followed by 35 cycles each consisting of 1.0 min at 94° C, 1.0 min at 50° C, and 2.5 min at 72° C with 2 s added to each extension step.

The double-stranded 1.8-kb PCR product was prepared for sequencing by means of agarose gel purification combined with binding to a DEAE membrane (Nickrent 1994). Approximately 80 µl of the amplified product was loaded with a Ficoll dye into a 33 mM acetate, 15 mM Tris (pH 8.3), 1% preparative agarose gel. During electrophoresis, the band was bound to a DEAE membrane (NA45, Schleicher & Schull, Keene, NH). Elution of the DNA from the membrane and subsequent precipitations followed the manufacturer's protocol. Normally, 400 µl of double-stranded product provided enough purified template DNA for subsequent sequencing reactions. Precipitated DNA pellets were resuspended in 30 µl of sterile water for every 100 µl of PCR product loaded on the gel; 7 µl of the template DNA was denatured in a boiling water bath for 1 min in the presence of 1 µl of 20 µM sequencing primer. The forward and reverse primers used for plant nuclear 18S rDNA sequencing are given in Table 2. The mixture was snap cooled in a −70° C ethanol bath and centrifuged briefly; 2 µl of 5× reaction buffer (Sequenase kit, U.S. Biochemical) was added. Primer annealing took place at 37° C for 20 min and the Sequenase protocol using [^{35}S-]dATP (New England Nuclear) was followed thereafter.

Sequence Alignment and Relative Rate Tests. The resulting sequences were aligned manually using Eyeball Sequence Editor (Cabot

Table 1. Voucher information for source material of 18S rDNA sequences

Species	Family	Tissue (accession No.)	Native to	Genbank Acc. No.
Rafflesia keithii Meijer	Rafflesiaceae	Perigone lobe (2944)[a]	Sabah, Malaysia	L24041
Rhizanthes zipelli (Bl.) Spach	Rafflesiaceae	Ovary (2844)[b]	Batu Ampar, W. Sumatra	L24048
Balanophora fungosa Bl.	Balanophoraceae	Bracts (2825)	Queensland Australia	L24044
Prosopanche americana (R. Br.) Baillon	Hydnoraceae	Root (2765)[c]	Argentina, S. America	L24047
Arceuthobium verticilliflorum Engl.	Viscaceae	Shoot (2065)	Durango, Mexico	L24042
Asarum canadense L.	Aristolochiaceae	Leaves (2888)	S. Illinois, U.S.A.	L24043
Malpighia coccigera L.	Malpighiaceae	Leaves (2903)	W. Indies, Cultivated SIUC	L24046
Gnetum leyboldii Tul. var. *woodsonianum* Markg.	Gnetaceae	Leaves (2946)[d]	Panama	L24045

[a] Tissue collected by W. Meijer (no number). Author's accession number in parentheses refers to voucher present at SIU
[b] Tissue collected by Satar, obtained via Futami W.S. & J. Trice (no number)
[c] Tissue collected by L.J. Musselman (no number)
[d] Plant cultivated at University of Illinois (Urbana/Champaign) accession no. 85181 (J. Kramer)

Table 2. Plant nuclear 18S rDNA primers

Primer sequence (5′ → 3′)	Position on *Glycine* (length)	Reverse primer name	Primer sequence (5′ → 3′)	Position on *Glycine* (length)	Forward primer name
CTAGYATTRCTACGGTTATCC	141–161 (21)	145	CTGGTTGAT CCTGCCAG	4–24 (17)	25e
CCATCGAAAGTTGATAGGGCA	310–330 (20)	310	GTCTGCCCT ATCAACT	307–322 (16)	366
TTTCTCAGGCTCCCTCTCCGG	385–405 (21)	428	GTGCCAGCMGCCGCGG	566–581 (16)	530
AATTACCGCGGCTGCTGGCA	567–586 (20)	567	YAGAGGTGA AATTCTTGG	901–918 (18)	700e
TCCAACTACGAGCTT	626–640 (15)	626	GAAACTTAA AKGAATTG	1134–1150 (17)	922
ATCATTACTCCGATCCCG	847–864 (18)	854	TAACGAACG AGACCTCAGCCT	1322-1343 (21)	1322
AATCCAAGAATTTCACCTCT	905–925 (20)	905	ARCTTCTTA GAGGGAC	1378–1393 (16)	1453
CAATTCCTTTAAGTTTCAGCC	1130–1150 (21)	1131	TTTGTACAC ACYGCCCGTCG	1631–1650 (20)	1631
AAGAACGGCCATGCACCACC	1269–1288 (20)	1324			
ATCTAAGGGCATS ACAGACC	1433–1452 (20)	1433			
CTCGTTGAAGACCAACAA	1562–1579 (18)	1575			
CACCTACGGAAACCTTGTT	1769–1787 (19)	1769			

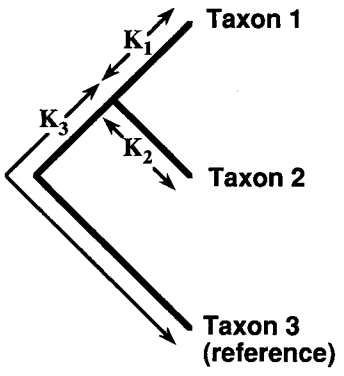

Fig. 1. The three-taxon tree used in the relative rate test. K_1 = number of nucleotide substitutions per site for taxon 1, K_2 for taxon 2 and K_3 for taxon 3 (the reference).

and Beckenbach 1989) using secondary structural features as guidelines (Neefs et al. 1993; Nickrent and Sargent 1991). The complete alignment can be obtained from the first author upon request.

To test whether the number of nucleotide substitutions was the same or different in two lineages, the relative rate test (Sarich and Wilson 1967; Wu and Li 1985) was used. As pointed out by Wu and Li (1985), the choice of the reference taxon is critical to minimize the errors in the estimations of substitutions per site (K). Although all of the parasitic plants used in this study are considered dicotyledonous by Cronquist (1988), Tahktajan (1980, 1987), and Thorne (1992), the assumption that they arose after the monocot/dicot divergence is uncertain. For this reason, *Gnetum leyboldii* was chosen as the reference species for relative rate tests since this genus (and other Gnetales) has been shown by multiple sources of evidence to be the sister to the angiosperms (Crane 1985; Hamby and Zimmer 1992; Chase et al. 1993).

Given that rDNA is not a protein coding gene, the nucleotide substitutions per site were not classified as synonymous or nonsynonymous. The total number of substitutions, K, is therefore the sum of the number of transitional and transversional substitutions per site. The null hypothesis of a relative rate test (Fig. 1) is that K from taxon 1 to taxon 3 (the reference) equals K from taxon 2 to the reference, i.e. H_o: $K_{13} = K_{23}$ or H_o: $K_{13} - K_{23} = 0$. The numbers of transitions and transversions for all pairwise comparisons of the 10 ingroup species and *Gnetum* were calculated. For rate calculations, only base substitutions were used, not length mutations (insertions and deletions). Since the number of sites differed slightly (less than 0.5%), a mean

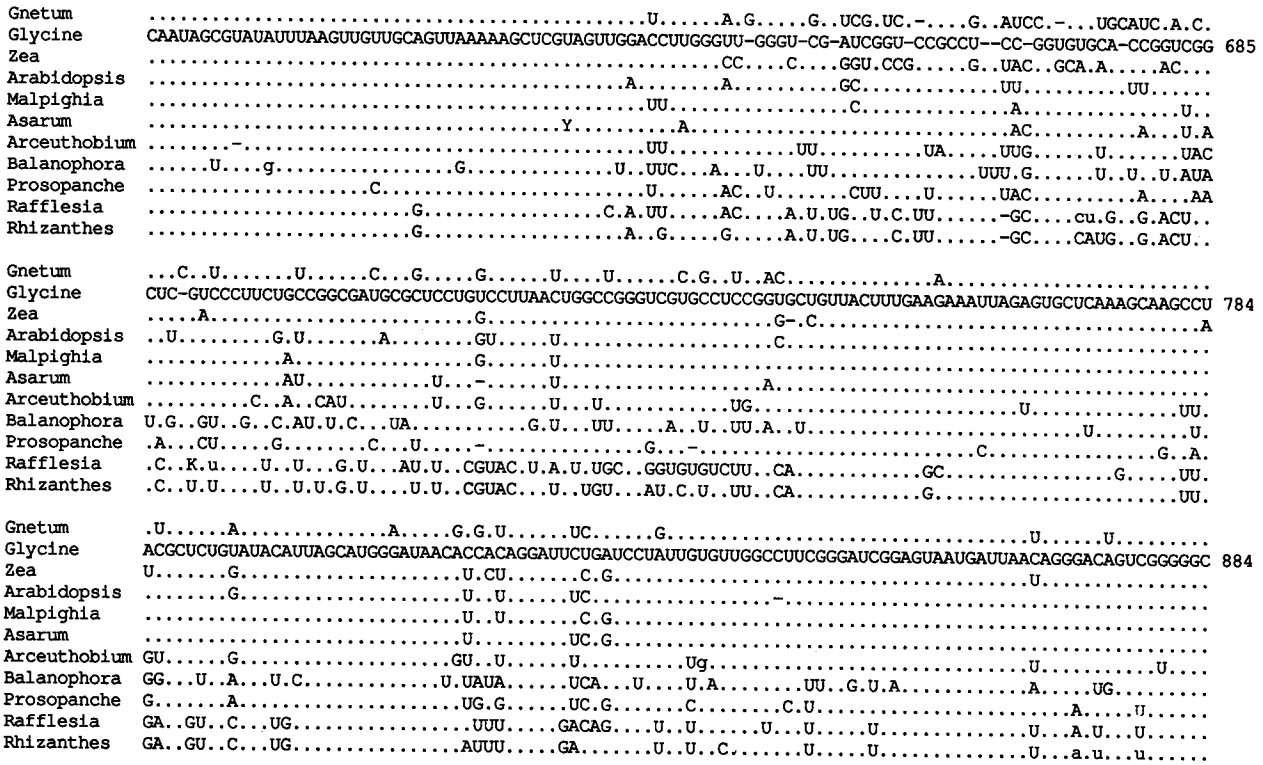

Fig. 2. A portion of the alignment of nuclear 18S rRNA from parasitic and nonparasitic angiosperms and a gnetophyte (*Gnetum*). The region shown, from site 594 to 884 on the soybean sequence, encompasses the V4 domain which contains 26% of the substitutions present in most angiosperms despite representing only 12.5% of the total length.

length of 1,805 was used for all 18S rDNA sequences to simplify calculations of standard errors.

To test the possible effect of generation time on *K*, sequences of plants with different life-history characteristics were chosen. Alternative definitions of generation time have been advanced such as one involving the number of mitotic cell divisions in the germline; however, unlike animals, plants do not possess a germline and sex cells are continually formed from vegetative cells (Klekowski 1988). For the purpose of general comparison, generation time is here defined as the time for the species to develop from a seed to its first production of fruit. The herbaceous annual *Arabidopsis thalliana* has a generation time as short as 6 weeks. The herbaceous perennial, *Asarum canadense,* assumed to have a generation time of at least 1 year, was also included to test a possible relationship between the Aristolochiaceae and Rafflesiaceae. The woody perennial *Malpighia coccigera* is assumed to have a generation time of at least 3 years.

Results

Alignments and Secondary Structure

In agreement with other eukaryotic 18S rDNA sequences, those reported here show regions of high conservation interspersed with regions of high variability. The latter can be demonstrated by the V4 domain, which contains ca. 26% of the total variation despite representing only 12.5% of the total length (227/1,813 bp). An alignment of this domain for the 10 angiosperm species and the outgroup *Gnetum* is shown in Fig. 2. The number of substitutions in the holoparasites (relative to *Glycine*) often exceeds the number seen in *Gnetum,* as can be seen upon inspection of more conserved regions (e.g., sites 594–641 in Fig. 2).

The number of nucleotide substitutions per site (*K*) differs widely among the plants examined in this study. The mean value for the four nonparasitic angiosperms (*Zea, Asarum, Malpighia,* and *Arabidopsis*) compared with *Glycine* is $K = 0.036$. With the addition of *Gnetum,* the mean *K* value increases to 0.045. In contrast, mean *K* for all the parasites is 0.115 and 0.126 when the four holoparasites are considered separately. These values show that the number of substitutions is 3.5 times higher in the holoparasites than in nonparasitic plants.

The large number of nucleotide substitutions in the holoparasites suggested that the sequence of the PCR product may be that of a pseudogene. That such modifications exist in holoparasitic angiosperms is evidenced by the presence of two tRNA pseudogenes in the *Epifagus virginiana* chloroplast genome (Wolfe et al. 1992). Despite the large number of substitutions, however, the 18S rRNA molecule of *Rafflesii keithii,* and the other parasite sequences reported here, can be folded into a secondary structure (Fig. 3) that conforms to those reported for *Saccharomyces* (Gutell 1993) and other eukaryotes (Neefs et al. 1993). Also, additional PCR amplifications resulted in identical sequences. Indirect but compelling evidence that these 18S rRNA molecules are functional is provided by the frequency of compensating mutations in helices where cannonical and/or non-

Table 3. Differences in number of nucleotide substitutions per site $K_1 - K_2$ ($=K_{13} - K_{23}$) for nuclear 18S rRNA genes using *Gnetum leyboldii* as reference (taxon 3)

	Taxon 2			
	Nonparasites			
Taxon 1 (nonparasites)	*Asarum*	*Malpighia*	*Glycine*	*Arabidopsis*
Zea	0.0171 ± 0.006[a]	0.0190 ± 0.006[a]	0.0220 ± 0.006[a]	0.0215 ± 0.006[a]
Asarum		0.0019 ± 0.004	0.0049 ± 0.005	0.0044 ± 0.005
Malpighia			0.0030 ± 0.004	0.0025 ± 0.005
Glycine				−0.005 ± 0.005
Arabidopsis				

[a] >3 standard errors
[b] >5 standard errors

cannonical pairing is maintained despite mutations of both members of the base pair. Length mutations are infrequent; however, one insertion of $(U)_4$ at position 133 characterizes both *Rafflesia* and *Rhizanthes*. A transitional bias (C to U) is evidenced by the large number of G • U pairs present in many helices (see positions 703 to 751 on helix 21-3).

Relative Rate Tests

Table 3 shows the results of 35 relative rate tests comparing five nonparasitic plants (taxon 1) with each other and five parasites (taxon 2). With the exception of the *Glycine* with *Arabidopsis* comparison, all tests involving nonparasites showed that the number of substitutions for taxon 1 was higher than taxon 2, although six of these nine differences were less than three times their standard errors. Three comparisons involving *Zea* showed differences slightly above 3 standard errors. All 25 tests using hemi- and holoparasites for taxon 2 were negative, indicating higher numbers of substitution in the parasites. Two of the tests using *Zea* (compared with *Arceuthobium* and *Prosopanche*) were not significantly different (less than 2 standard errors). The remaining differences were at least three and often greater than five times their standard errors, indicating significantly higher substitution rates in the parasites, especially in the holoparasitic Balanophoraceae, Hydnoraceae, and Rafflesiaceae.

Discussion

The first evidence that 18S rRNA in parasitic angiosperms displayed high evolutionary rates was presented by Nickrent and Franchina (1990). Parsimony analysis resulted in phylograms with longer branches leading to two members of the mistletoe family Viscaceae. Further sequencing of members of holoparasitic families (this paper) has documented an unprecedented degree of molecular divergence in the 18S rRNA in these plants relative to hemiparasites and autotrophs. The unusual evolutionary dynamic of these parasites is characterized by a high proportion of transitional changes which outnumber transversions by a factor of 2:1. The C-to-U transition is especially frequent—for example, in the 18S rRNA of *Rhizanthes*, the G/C ratio is 510/344 (28%/19% of total) whereas the A/U ratio is 426/526 (23%/29%). By comparison, *Glycine* has a G/C ratio of 490/397 (27%/22%) and A/U of 452/469 (25%/26%). Such transitional biases can also be observed in organellar genomes, which suggests a possible relationship with extreme or highly specialized types of interorganismal associations (including endosymbiosis, parasitism, mutualism, etc.).

Evidence that major structural rearrangements can occur in the plastid genomes of holoparasitic angiosperms came from work on achlorophyllous, nonphotosynthetic parasites such as *Epifagus* (Orobanchaceae), whose plastid genome is reduced from 156 kb to 71 kb, mainly as a result of loss of genes associated with photosynthesis (dePamphilis and Palmer 1990). Given their nutritional dependence upon another plant, such losses were explained by a relaxation of selectional constraints. The loss of photosynthetic genes can be rationalized in parasitic plants that obtain photosynthates from the host; however, it is not immediately clear why their nuclear-encoded ribosomal genes show higher mutation fixation rates since rRNA is essential to the translational function of the ribosome. Additional evidence linking nutritional mode with accelerated evolutionary rates come from mycotrophic plants. Mycotrophs utilize a fungal intermediate to obtain nutrients from host roots and are therefore not strictly parasitic since they lack direct physical connection to the host via a haustorium. Elevated substitution rates in nuclear 28S rRNA were noted in the mycotrophic Monotropoideae as compared with other members of the Ericales (Cullings and Bruns 1992; T. Bruns, personal communication). These results suggest that highly derived nutritional strategies, not parasitism per se, may be associated with this phenomenon.

Higher rates of nucleotide substitution for *rbc*L were

Table 3. Continued

Hemiparasite	Holoparasites			
Arceuthobium	*Prosopanche*	*Balanophora*	*Rhizanthes*	*Rafflesia*
−0.0055 ± 0.008	−0.0144 ± 0.008	−0.0521 ± 0.010[b]	−0.0518 ± 0.010[b]	−0.0554 ± 0.010[b]
−0.0226 ± 0.007[a]	−0.0315 ± 0.007[a]	−0.0692 ± 0.010[b]	−0.0688 ± 0.010[b]	−0.0725 ± 0.010[b]
−0.0245 ± 0.007[a]	−0.0334 ± 0.007[a]	−0.0711 ± 0.010[b]	−0.0708 ± 0.010[b]	−0.0744 ± 0.010[b]
−0.0275 ± 0.007[a]	−0.0364 ± 0.007[b]	−0.0741 ± 0.010[b]	−0.0738 ± 0.010[b]	−0.0774 ± 0.010[b]
−0.0270 ± 0.007[a]	−0.0359 ± 0.008[b]	−0.0736 ± 0.010[b]	−0.0732 ± 0.010[b]	−0.0769 ± 0.010[b]

observed in the grass family (Poaceae) compared to other monocots such as palms (Gaut et al. 1992). That investigation implicated generation time as the cause of rate heterogeneity, as had been suggested by previous studies of other organisms (Wu and Li 1985), but this explanation was rejected by Bousquet et al. (1992). The present study does not support a relationship between generation time and substitution rate in plant nuclear 18S rRNA. The relative rate test comparing fast-cycling *Arabidopsis* with a long-generation-time woody perennial (*Malpighia*) showed no significant differences in *K*. Differences in *K* were seen in comparisons involving *Zea;* however, a new sequence may be required since the existing one is relatively old (Messing et al. 1984). Sequencing errors in the first maize *rbc*L sequence resulted in perceived rate differences between it and other grasses that disappeared following determination of the correct sequence (Gaut et al. 1992). To accept generation time as an explanation of increased substitution rates in the holoparasites would imply that they complete their life cycles in a shorter time span than any of the other plants studied, which is clearly not the case. The amount of time required for *Rafflesia* to reach sexual maturity is approximately 4–5 years (Meijer 1958).

The underlying causes of unequal rates of mutation fixation are unknown, but a number of hypotheses have been advanced. Small effective population size (Wu and Li 1985) resulting in genetic bottlenecks may be implicated since some of the holoparasites described here have small, often widely scattered populations. The populational dynamics of parasites are characterized by patchy resources (hosts), low probabilities of colonization, and high probabilities of extinction (Price 1980). These conditions apply to *Rafflesia, Rhizanthes,* and possibly *Balanophora* and *Prosopanche.* They do not apply, however, to the hemiparasite *Arceuthobium,* which also shows an accelerated rate yet is present in large, widespread populations containing thousands or millions of individuals. It is therefore unlikely that effective population size can be unequivocally invoked as an explanation of higher substitution rates in all parasitic plants. Other hypotheses that merit further investigation include organellar location of the gene (Vawter and Brown 1986; Wolfe et al. 1987), polymerase fidelity (Kunkel and Loeb 1982), efficiency of DNA replication or repair (Topal and Fresco 1976), choice of organism (Wu and Li 1985; Gaut et al. 1992), and selective constraints imposed by the molecular environment (Kimura 1987).

Complete 18S rDNA sequences have been obtained for over 50 hemiparasitic members of the Santalales. Although these plants are parasitic, they do not show the increased rates seen in the holoparasites (and Viscaceae). Higher rates are only associated with those plants that have evolved highly specialized haustoria (including internal endophytes as well as the balanophoraceous tuber). This evidence implies a complex interaction between the host and parasite tissues at the cellular and biochemical level. The holoparasites may have developed an increased tolerance of mutations on their 18S rRNA due to the "selection buffered" host environment, but only within the limits of functional constraints of the molecule. It is unlikely that these parasites have commandeered their host's ribosomes; it is only known that the number of functionally constrained sites on their rRNA is lower than expected. Wilson et al. (1977) made a direct correlation between the rates of (protein) evolution and the probability that a substitution will be compatible with the biochemical function of the molecule and its dispensability. Although it is accepted that ribosomes are indispensable, the evidence presented here does not exclude the possibility that holoparasitic rRNAs (and hence their ribosomes) are less efficient than those of their hosts. Comparative in vitro assays are required to test this hypothesis.

Processes other than mutation, natural selection, and genetic drift may influence mutation fixation rates in gene families. These stochastic and directional intragenomic mechanisms are referred to as molecular drive and include unequal gene exchange, gene conversion, slippage, transposition, and RNA-mediated changes (Dover 1982, 1987). These processes can result in variant molecules being spread throughout a sexual pop-

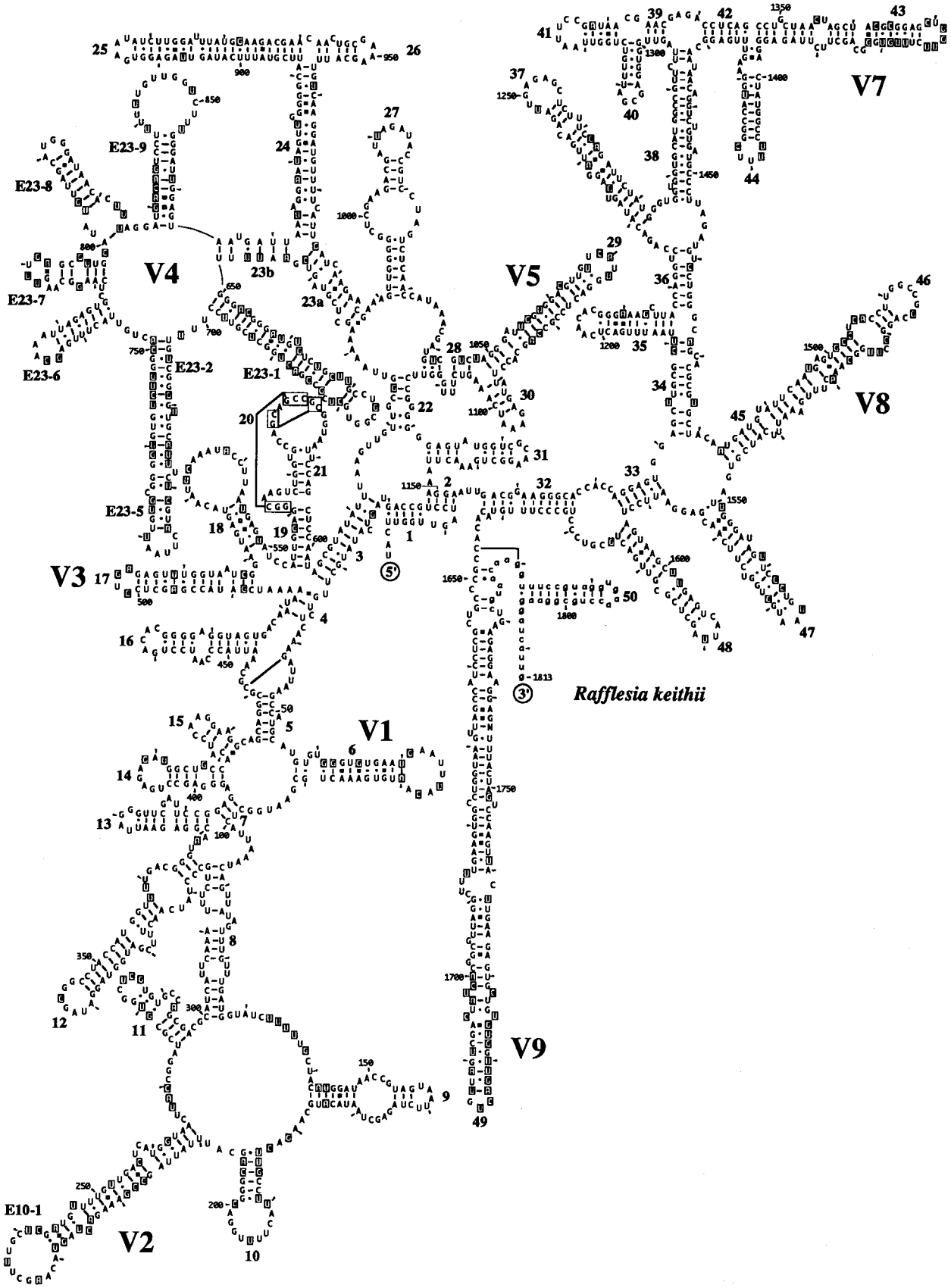

Fig. 3. Higher-order structure for nuclear 18S rRNA of *Rafflesia keithii* (Rafflesiaceae) allowing both cannonical and noncannonical base pairing. The general structural configuration follows *Saccharomyces cerevisiae* (Gutell 1993) and Nickrent and Sargent (1991) for the V4 region. The V6 region is found only in prokaryotes. Helix numbering corresponds to Neefs et al. (1993). Bases in *reverse font* indicate mutations relative to *Glycine max* (soybean). The 3′ terminal 42 bases (in *lowercase*) were at or beyond the 1830 reverse priming site and therefore were not determined. The consensus of all angiosperm sequences is given here for completeness. Tertiary interactions are indicated by *thick lines*.

ulation and have been cited as significant factors in accelerating the rates of DNA divergence between populations. The extreme homogeneity of ribosomal DNA cistrons is thought to be a result of horizontal or concerted evolution (Arnheim et al. 1980; Brown et al. 1972). This replacement process can result in the rapid establishment of taxa that exhibit ribosomal sequences unique to their particular phyletic line but that differ substantially from relatives (Dover and Coen 1981). Heritable changes in the number of ribosomal RNA genes within an individual can be environmentally induced in flax (Cullis 1984). Such "genotrophs" illustrate the dynamic nature of ribosomal genes and suggest that variations can arise rapidly, be amplified within a genome via molecular drive, and proliferate via sexual reproduction, especially where effective population size is small.

High substitution rates appear to be affecting phylogeny reconstruction as seen during analyses that included the holoparasite 18S rDNA sequences in a dataset with over 80 other angiosperms. Following parsimony analysis, artifactual attractions (Felsenstein 1978) were seen as clades composed of long branches leading to the holoparasites that have high substitution rates. In addition, the holoparasites tended to cluster at the base of the tree between *Gnetum* (the outgroup) and the nonparasitic angiosperms. Similar results were obtained using maximum likelihood distance methods that do not assume a molecular clock such as FITCH (Fitch and Margoliash 1967) and neighbor-joining methods (Saitou and Nei 1987). These results suggest that the holoparasite families are related and that they diverged very early from all other angiosperms. At present this is considered unlikely since rate heterogeneity is clearly influencing phylogeny estimation, as demonstrated below. The relative rate test requires that the three-taxon topology be a correct representation of the phylogeny of those taxa. If this is not the case, perceived differences in K may be artifactual. As mentioned above, many lines of independent evidence support the concept that the gnetophytes are sister to the angiosperms. For this reason, the results from the above phylogeny inference methods should not be viewed as compromising the validity of the relative rate tests.

When using fewer taxa (such as the eight angiosperms shown in Table 1 plus *Glycine, Arabidopsis,* and *Zea*) in parsimony, distance, and maximum likelihood analyses, all the parasites cluster to the base of the tree, usually in the topology: (*Gnetum*(*Prosopanche*((*Arceuthobium*(*Balanophora*(*Rhizanthes, Rafflesia*)))other angiosperms))). This topology also results when the sequence data are truncated by excluding all eight variable domains (resulting in a dataset of ca. 1,100 bp), which indicates that the number of substitutions is high, even in the conserved "core" of the molecule. The above placement of *Arceuthobium* is clearly anomalous, for it clusters with other mistletoes in its family (Viscaceae) when additional members of the order Santalales are included.

As has been demonstrated by other phylogenetic studies of angiosperms (e.g., Chase et al. 1993), it is essential to obtain sufficient taxon density to properly polarize character state changes and recover the correct tree. Since which plants represent the closest relatives of the holoparasites is disputed, further sequencing both within and outside of the group is required. Determining their relationships to other angiosperms may also be aided by alternate analytical procedures that are designed to measure evolutionary distances among taxa with greatly different substitution rates. These include operator metrics and evolutionary parsimony (Lake 1987) and variability calibration algorithms (Van de Peer et al. 1993).

The divergent rRNA sequences reported here for holoparasitic angiosperms likely resulted from several mechanisms acting simultaneously. These include (but are not restricted to) small effective population size, relaxed constraints on function and structure of rRNA manifested as an increase in the proportion of sites free to vary, and molecular drive. Whatever their causes, these divergent sequences present a challenge to molecular phylogenetic analysis and highlight yet another unusual aspect of the evolution of parasitic flowering plants. It is hoped that this type of information on the unique features of plants such as *Rafflesia* can be used to further justify measures aimed at their conservation.

Acknowledgments. We gratefully acknowledge the collectors listed in Table 1 without whose assistance samples of these unusual plants could not have been obtained. Thanks go to Joseph Agan and Brent Beilschmidt, who provided laboratory assistance, and to Carey Krajewski, who helped with the relative rate calculations and provided many insightful comments. The manuscript was improved by a critical review by Gary Olsen, who also assisted in the maximum likelihood analyses. This project was supported by a grant (BSR-8918385) from the National Science Foundation.

References

Albert VA, Williams SE, Chase MW (1992) Carnivorous plants: phylogeny and structural evolution. Science 257:1491–1495

Arnheim ND Krystal M, Schmickel R, Wilson G, Ryder O, Zimmer E (1980) Molecular evidence for genetic exchanges among ribosomal genes on nonhomologous chromosomes in man and apes. Proc Natl Acad Sci USA 77:7323–7327

Bosquet J, Strauss SH, Doerksen AH, Price RA (1992) Extensive variation in evolutionary rate of *rbc*L gene sequences among seed plants. Proc Natl Acad Sci USA 89:7844–7848

Brown DD, Wensink PC, Jordan E (1972) A comparison of the ribosomal DNA's of *Xenopus laevis* and *Xenopus mulleri:* The evolution of tandem genes. J Mol Biol 63:57–73

Cabot EL, Beckenbach AT (1989) Simultaneous editing of multiple nucleic acid and protein sequences with ESEE. Comp Appl Biosci 5:233–234

Chase MW, Soltis DE, Olmstead RG, Morgan D, Les DH, Mishler BD, Duvall MR, Price RA, Hills HG, Qiu Y-L, Kron KA, Rettig JH, Conti E, Palmer JD, Manhart JR, Sytsma KJ, Michaels HJ, Kress WJ, Karol KG, Clark WD, Hedrén M, Gaut BS, Jansen RK, Kim K-J, Wimpee CF, Smith JF, Furnier GR, Strauss SH, Xiang QY, Plunkett GM, Soltis PS, Swensen SM, Williams SE, Gadek PA, Quinn CJ, Eguiarte LE, Golenberg E, Learn GH Jr, Graham SW, Barrett SCH, Dayanandan S, Albert VA (1993) Phylogenetics of seed plants: an analysis of nucleotide sequences from the plastid gene *rbc*L. Ann Mo Bot Gard 80:528–580

Crane PR (1985) Phylogenetic analysis of seed plants and the origin of angiosperms. Ann Mo Bot Gard 72:716–793

Cronquist A (1988) The evolution and classification of flowering plants, 2nd ed. The New York Botanical Garden, Bronx, NY

Cullings KW, Bruns TD (1992) Phylogenetic origin of the Monotropoideae inferred from partial 28S ribosomal RNA gene sequences. Can J Bot 70:1703–1708

Cullis CA (1984) Quantitative variation of the ribosomal RNA genes. In: Jordan EG, Cullis CA (eds) The nucleolus. Cambridge University Press, Cambridge, pp 103–112

dePamphilis CW, Palmer JD (1990) Loss of photosynthetic and chlororespiratory genes from the plastid genome of a parasitic flowering plant. Nature 348:337–339

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

Dover GA (1987) DNA turnover and the molecular clock. J Mol Evol 26:47–58

Dover G, Coen E (1981) Springcleaning ribosomal DNA: a model for multgene evolution? Nature 290:731–732

Doyle JJ, Doyle JL (1987) A rapid DNA isolation procedure for small quantities of fresh leaf tissue. Phytochem Bull 19:11–15

Eckenrode VK, Arnold J, Meagher RB (1985) Comparison of the nucleotide sequence of soybean 18S rRNA with the sequences of other small-subunit rRNAs. J Mol Evol 21:259–269

Felsenstein J (1978) Cases in which parsimony or compatibility will be positively misleading. Syst Zool 27:410–410

Fitch WM, Margoliash E (1967) Construction of phylogenetic trees. Science 155:279–284

Gaut BS, Muse SV, Clark WD, Clegg MT (1992) Relative rates of nucleotide substitution at the *rbc*L locus of monocotyledonous plants. J Mol Evol 35:292–303

Gedalovich-Shedletzky E, Kuijt J (1990) An ultrastructural study of the tuber stands of *Balanophora* (Balanophoraceae). Can J Bot 68: 1271–1279

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

Hamby KR, Zimmer EA (1992) Ribosomal RNA as a phylogenetic tool in plant systematics. In: Soltis PS, Soltis DE, Doyle JJ (eds) Molecular systematics of plants. Chapman and Hall, New York, pp 50–91

Ismail G (1988) Conservation of the giant *Rafflesia* in Sabah, Malaysia. TREE 3:316–317

Kimura M (1987) Molecular evolutionary clock and the neutral theory. J Mol Evol 26:24–33

Klekowski EJ (1988) Mechanisms that maintain the genetic integrity of plants. In: Greuter W, Zimmer B (eds) Proceedings of the XIV International Botanical Congress. Koeltz, Königstein/Taunus, pp 137–152

Kuijt J (1969) The biology of parasitic flowering plants. University of California Press, Berkeley, p 246

Kunkel TA, Loeb LA (1982) Fidelity of mammalian DNA polymerases. Science 213:765–767

Lake JA (1987) Determining evolutionary distances from highly diverged nucleic acid sequences: operator metrics. J Mol Evol 26: 59–73

Meijer W (1958) A contribution to the taxonomy and biology of *Rafflesia arnoldi* in West Sumatra. Ann Bogorienses 3:33–44

Messing J, Carlson J, Hagen G, Rubenstein I, Oleson A (1984) Cloning and sequencing of the ribosomal RNA genes in maize: the 17S region. DNA 3:31–40

Mullis KB, Faloona FA (1987) Specific synthesis of DNA *in vitro* via a polymerase-catalyzed chain reaction. Methods Enzymol 155: 335–350

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

Nickrent DL, Sargent ML (1991) An overview of the secondary structure of the V4 region of eukaryotic small-subunit ribosomal RNA. Nucleic Acids Res 19:227–235

Nickrent DL, Franchina CR (1990) Phylogenetic relationships of the Santalales and relatives. J Mol Evol 31:294–301

Nickrent DL (1994) From field to film: rapid sequencing methods for field-collected plant species. Biotechniques 16:470–475

Price PW (1980) Evolutionary biology of parasites. Monographs in Population Biology No. 15. Princeton University Press, Princeton, p 237

Saitou N, Nei M (1987) The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol Biol Evol 4:406–425

Sarich VM, Wilson AC (1967) Rates of albumin evolution in primates. Proc Natl Acad Sci USA 58:142–148

Sogin ML, Elwood HH, Gunderson HH (1986) Evolutionary diversity of eukaryotic small-subunit rRNA genes. Proc Natl Acad Sci USA 83:1383–1387

Takhtajan AL (1980) Outline of the classification of flowering plants (Magnoliophyta). Bot Rev 46:225–359

Takhtajan AL (1987) Systema Magnoliophytorum. Nauka, Leningrad (in Russian)

Thorne RF (1992) An updated classification of the flowering plants. Aliso 13:365–389

Topal MD, Fresco JR (1976) Complementary base pairing and the origin of substitution mutations. Nature 263:285–289

Unfried I, Stocker U, Gruendler P (1989) Nucleotide sequence of the 18S rRNA gene from *Arabidopsis thaliana* Col0. Nucleic Acids Res 17:7513

Van de Peer Y, Neefs J-M, De Rijk P, De Wachter R (1993) Reconstructing evolution from eukaryotic small-ribosomal-subunit RNA sequences: calibration of the molecular clock. J Mol Evol 37:221–232

Vawter L, Brown WM (1986) Nuclear and mitochondrial DNA comparisons reveal extreme rate variation in the molecular clock. Science 234:194–196

Wilson AC, Carlson SS, White TJ (1977) Biochemical evolution. Annu Rev Biochem 46:573–639

Woese CR (1987) Bacterial evolution. Microbiol Rev 51:221–271

Wolfe KH, Li W-H, Sharp PM (1987) Rates of nucleotide substitution vary greatly among plant mitochondrial, chloroplast, and nuclear DNAs. Proc Natl Acad Sci USA 84:9054–9058

Wolfe KH, Katz-Downie DS, Morden CW, Palmer JD (1992) Evolution of the plastid ribosomal RNA operon in a nongreen parasitic plant: accelerated evolution, altered promoter structure, and tRNA pseudogenes. Plant Mol Biol 18:1037–1048

Wolfe KH, Gouy M, Yang Y-W, Sharp PM, Li W-H (1989a) Date of the monocot-dicot divergence estimated from chloroplast DNA sequence data. Proc Natl Acad Sci USA 86:6201–6205

Wolfe KH, Sharp PM, Li W-H (1989b) Rates of synonymous substitution in plant nuclear genes. J Mol Evol 29:208–211

Wu C-I, Li W-H (1985) Evidence for higher rates of nucleotide substitution in rodents than in man. Proc Natl Acad Sci USA 82:1741–1745