Plant Molecular Biology **34:** 731–743, 1997.

Structural analyses of plastid-derived 16S rRNAs in holoparasitic angiosperms

Daniel L. Nickrent[1,*], R. Joel Duff[1] and D.A.M. Konings[2]
[1]*Department of Plant Biology, Southern Illinois University, Carbondale, IL 62901-6509, USA (*author for correspondence); [2]Department of Microbiology, Southern Illinois University, Carbondale, IL 62901-6508, USA*

Received 9 October 1996; accepted in revised form 16 April 1997

Key words: nonphotosynthetic plant, small-subunit ribosomal RNA structure, substitution rate acceleration, composition bias

Abstract

Higher-order structures have been constructed for plastid-encoded small-subunit (SSU, 16S), rRNAs from representatives of seven nonphotosynthetic holoparasitic angiosperm families: Apodanthaceae, Cynomoriaceae, Cytinaceae, Balanophoraceae, Hydnoraceae, Mitrastemonaceae, and Rafflesiaceae. Whereas most pairwise comparisons among angiosperms differ by 2–3% in substitutions, the 16S rRNAs of the 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%). Despite this high level of sequence variation, SSU structures constructed for all species except *Pilostyles* possess the typical complement of 50 helices (that contain numerous compensatory mutations) thereby providing indirect evidence supporting their functionality. *Pilostyles*, likely with the most unusual plastid 16S rRNA yet documented, lacks four major helices and contains lengthy insertions for four others. Sequences of products generated via RT-PCR show that these structural modifications are present on a mature (transcribed) rRNA. The trend toward increasing numbers of base substitutions in the holoparasites is accompanied by a marked increase in A+U content of the rRNA. This 'A/T drift' phenomenon of rDNA is especially apparent in *Corynaea* whose SSU rDNA sequence is 72% A+T. A comparison of *Cytinus* to tobacco showed that substitution rates appear to be dependent upon the composition of neighboring bases. Transversions represented 26% of the mutations when flanking bases were G or C whereas transversions increased to 36% when the flanking bases were A to T. The underlying molecular mechanism associated with these high substitution rates is presently unknown, however, relaxation of selection pressure on ribosome function resulting in altered DNA replication and/or repair systems may be involved.

Introduction

Small-subunit (16S and 16S-like) rRNA sequences are likely the largest category of macromolecular sequence in existence. About 3000 are being maintained by the Ribosomal Database Project [25]. One reason for the continued generation of 16S rRNA sequences is their tremendous utility as markers of phylogenetic change among all known life forms. To correctly extract information from these rRNA sequences, phylogenetic analyses must incorporate structural and functional information [6, 31]. In recent years, tremendous advances have been made in understanding the role rRNA plays in the translational process [39]. Further understanding of the mechanism of translation has been gained by studies of the higher-order structure of ribosomal RNAs [20].

Two distinct but complementary approaches to predicting RNA higher-order structure exist: energy minimization and comparative sequence analysis. Steady improvement has been made in the development of algorithms that use thermodynamic energy values for secondary structure predictions [21, 53]. While refine-

The nucleotide sequence data reported will appear in the EMBL, GenBank and DDBJ Nucleotide Sequence Databases under the accession numbers U67740, U67741, U67742 and U67743.

ment of this approach promises to further improve the prediction of structural models of 16S rRNA, at present only a portion of the existing helical elements are predicted, i.e. in particular base pairs interacting over short distances (<100 bases). Predictions for archaeal and eubacterial 16S rRNA sequences are better than for chloroplasts and significantly better than for mitochondrial and eukaryotic sequences [22]. The comparative sequence approach has been successful in providing detailed structural information for a wide variety of organisms [15–17, 49]. In addition, this method can be used to predict secondary structure and identify tertiary interactions using information on compensatory changes and covariance [15]. A requirement of the comparative approach, however, is the availability of sufficiently diverse sequences, especially in relatively conservative regions, to allow structural predictions to be made.

Results from phylogenetic analysis of plastid gene sequences and overall genome organization generally confirm that plastids arose once from cyanobacteria via an endosymbiotic event [7, 28, 33, 41]. Chloroplast genomes are highly conserved for gene size and arrangement and their primary mode of evolution is via substitutional mutations [40]. When 13 plastid 16S rRNA sequences of land plants are compared to *Marchantia* (a liverwort), the mean percent nucleotide similarity is 95.2%. This value increases to 97.6% when ten angiosperm sequences are compared, thus demonstrating the highly conserved nature of this gene. Given the low rate of substitution, plastid 16S rRNA have only rarely been used in phylogenetic studies such as those examining the origin of plastids or deep divergences among algae and land plants [7, 26, 27].

Among flowering plants, dramatic reorganization of the plastid genome (plastome) has taken place in parasitic members that have lost most photosynthetic genes (reviewed in [3, 35]). The 70 kb plastome of the holoparasite *Epifagus virginiana* (beechdrops, Scrophulariaceae) has been mapped and sequenced [4, 51] and it contains functional ribosomal cistrons that are present on inverted repeats [29]. Although the beechdrops 16S rDNA sequence differs little from that of tobacco (2.7% of the sites), Wolfe *et al.* [50] showed that the number of substitutions on the branch leading to *Epifagus* was higher than in comparisons involving most other angiosperms. Work with *Epifagus* was the first to show that genome reorganization in holoparasitic flowering plants also affects nucleotide substitutions on plastid ribosomal genes. In addition to Scrophulariaceae, holoparasitism has evolved at least five more times among several 'nonasterid' angiosperm lineages [23, 35, 36]. Considerable disagreement exists on the classification of these plants, but here a modification of the system proposed by Takhtajan [44] will be followed (Table 1). Little structural information is available concerning the presence of plastids in these holoparasites [37]; however, results from at least one species (*Cytinus ruber*, Cytinaceae) suggests the retention of a vestigial plastome.

The comparative approach to predicting rRNA secondary structure requires the presence of sufficient sequence variation among the taxa. In addition to exploiting wide phylogenetic diversity to obtain such variation, an alternative approach is to utilize organisms whose ribosomal loci are evolving at faster rates. Analyses of nuclear 18S rDNA sequences of these nonasterid holoparasites showed they had nucleotide substitution rates, on average, 3.5 times higher than nonparasitic plants [38]. The present study shows that substitutional changes in plastid 16S rDNA of these plants are also increased but to a greater degree than with nuclear 18S rDNA. The specific goals of this study are to construct higher-order structural models of 16S rRNAs from representatives of all the lineages of nonasterid holoparasites and compare these models to those derived from photosynthetic plants. The evolutionary progression from moderate to extreme structural reorganization in these rRNAs will be discussed in a comparative framework.

Materials and methods

The species shown in Table 1 are representatives of all major lineages of nonasterid holoparasites. Both fresh (*Cytinus*, *Pilostyles*, *Corynaea* and *Hydnora*) and silica dried (*Bdallophyton*, *Mitrastema* and *Cynomorium*) tissues were used for DNA extraction. Total genomic DNA was obtained by grinding the tissue to a powder on liquid nitrogen and extracting using 2× CTAB as described [34]. PCR reactions, gel purification of resulting products, and sequencing conditions were according to Nickrent [34]. The primers used for PCR amplification and sequencing of plastid SSU rDNA are given in the accompanying paper [37].

Total RNA was isolated from frozen plant tissue of *Pilostyles* (collection number 2994) following the method described in Ems *et al.* [10]. The total RNA sample was treated overnight with RNase-free DNase (Promega). cDNA were synthesized by reverse transcription and amplification by PCR (RT-PCR). Reverse

Table 1. Holoparasites used to obtain plastid-encoded 16S rDNA sequences.

Species	Family	Collection number[a]	Native to	GenBank accession number	%GC
Cytinus ruber Fritsch	Cytinaceae[b]	2738	Mediterranean	U47845	50
Bdallophyton americanum R. Br.	Rafflesiaceae	3072	Mexico	U67740	46.4
Pilostyles thurberi Gray	Apodanthaceae[b]	2994	Texas, USA	U67741	34
Mitrastema yamamotoi Makino	Mitrastemonaceae[b]	2941	Borneo	U67742	46.5
Cynomorium coccineum L.	Cynomoriaceae[c]	4000	Israel	U67743	52
Corynaea crassa Hook f.	Balanophoraceae[c]	3011	Costa Rica	U67744	28
Hydnora africana Thunb.	Hydnoraceae	2767	South Africa	U67745	43.1

[a]Collection numbers (of DLN) refer to voucher specimens present at SIU.
[b]Classification according to Takhtajan [44]. Traditionally placed in a broadly defined Rafflesiaceae.
[c]Classification according to Takhtajan [44]. Traditionally placed in a broadly defined Balanophoraceae.

transcription took place in a reaction mixture containing 10 ng of antisense primer (1289 rev), 1 mM for each dNTP (Pharmacia) and PCR buffer (10 mM Tris-HCl pH 8.5, 6 mM $MgCl_2$, 50 mM KCl, 10 mM DTT). Total cellular RNA was denatured at 90 °C for 1 min and allowed to cool to room temperature to anneal the primer. First-strand cDNA synthesis was performed with 1000 units of AMV-reverse transcriptase (USB) for 20 min at 37 °C. PCR amplification of cDNAs was performed with 250 ng of each primer (323 for and 1289 rev) and one unit of *Taq* polymerase. Amplification was performed with a Perkin Elmer thermal cycler with the same conditions as described for DNA amplifications above. The absence of genomic DNA contamination in RNA samples was verified by attempted amplification from the non-reverse transcribed RNA samples with several combinations of PCR primers used on the cDNA sample.

The sequences were entered into the computer program SeqApp [12] where a manual multiple sequence alignment was conducted. The 16S rDNA alignment contained 35 land plant, 11 algal and two cyanobacterial sequences. Concurrent with alignment, secondary structural models were constructed for all taxa and this information was used to guide the alignment of variable regions. The structural models were constructed following that of *Nicotiana* [13] using Microsoft ClarisDraw Version 1.0. At present, three landplant plastid 16S rRNA secondary structures are available for downloading as postscript files via the 16S rRNA Comparative Structure Database [13]: *Marchantia polymorpha* (X04465), *Zea mays* (Z00028), and *Nicotiana tabacum* (Z00044). The tobacco structure, possibly based upon a previous sequence (V00165), contained six errors that required correction (positions 357, 787, 879, 935, 1167 and 1403). Since these 16S rRNA models contain information on which base pairs are supported by covariance data [15], attempts were made to strictly adhere to these helix constraints. In doing so, some sites required noncanonical pairing (discussed below). In this study, G•U is not considered a noncanonical base pair. For reference, we have followed the helix numbering system proposed for prokaryotic SSU rRNA by Dewachter and associates [32, 47]. The nomenclature used to describe structural motifs in rRNA generally follows Woese *et al.* [49] with additional details proposed by Ehresmann *et al.* [9].

Results and discussion

Plastid SSU rRNAs for most plants are remarkably similar in overall length (ca. 1490 bp) and contain comparatively little structural variation in the 50 major helices. Conversely, the 16S rRNAs of the seven holoparasitic angiosperms reported herein differ markedly in primary and secondary structure. As shown in Table 1, the G+C content of holoparasite SSU sequences ranges from 28–52%, all lower than the mean for more typical land plants (56%). The number of substitutional mutations between tobacco and liverwort is 55, whereas 109 mutations (7.3%) separate the phylogenetically closer tobacco and *Cynomorium*. The other holoparasites show an increasingly greater number of mutational differences: *Cytinus* (121, 8.0%), *Bdallophyton* (193, 12.7%), *Mitrastema* (219, 14.9%), *Hydnora* (292, 19.4%), *Pilostyles* (ca. 465, 30.4%) and *Corynaea* (ca. 513, 35.9%). Of these mutations, transitions outnumber transversion by at least 1.9:1 which is lower than the ratio for nonparasitic plants (2.6:1).

The trend toward increasing transversions is especially apparent in *Pilostyles* and *Corynaea* where the TS/TV ratio approaches parity. The frequency of noncanonical base pairs in plastid SSU sequences ranges from 8 to 12% in the holoparasites as compared with 7.5% in a sample set of photosynthetic plants [22].

Structural features of plastid 16S rRNA in holoparasitic angiosperms

Despite containing a high number of substitutions, the plastid 16S rRNA structure of *Cynomorium* is very similar to SSU rRNAs from representatives of land plant and algal groups (12 plastid SSU structures available on the 16S rRNA Comparative Structure Database). Being the least altered species among the holoparasites discussed here, *Cynomorium* will be used to illustrate a 'typical' 16S rRNA molecule (Figure 1). In the following discussion, the structural variations between the holoparasite 16S rRNAs and those of other plants will be compared and illustrated. The observed variation includes structural conformation changes for helical and loop regions, noncanonical base pair frequencies and composition bias. Complete structural diagrams for all holoparasites have been constructed (available upon request from DLN), however, a composite diagram (Figure 2) will be used to illustrate only those helical regions that show significant structural differences from the norm. To demonstrate that these divergent rRNAs can be folded into secondary structures and to aid discussion, two additional complete structures will be shown for *Hydnora* (Figure 3) and *Pilostyles* (Figure 4), the most unusual plastid SSU sequence reported to date (R. Gutell, pers. com.). The base pairing schemes presented in these figures are according to the covariance model of *Nicotiana* by Gutell [14].

Helix length variation

The major type of structural variation seen in the holoparasite 16S rRNAs is variation in the length of helical structures. This feature is found in all sequences reported here but is most extreme in *Pilostyles* (Figure 4). Long extensions of helices are seen in *Pilostyles* for helices 6, 23-1 and 29 that have, respectively, 22, 34 and 92 additional bases (Figure 4). One criterion typically used to recognize helices is the presence of conserved apex loops and flanking regions [49]. Using this criterion for helix 23-1, a pentaloop of CUYAA (UCAAA in *Corynaea*) occurs in 90% of all eubacteria and plastid sequences. Despite containing this conserved motif in the 34 bp extension, the flanking sequences in *Pilostyles* do not appear to engage in canonical pairing, hence a structure for these additional bases cannot be proposed. For the remaining extensions, some have potential to form secondary structures, yet no comparative data exist for confirmation. *Pilostyles* shows additional examples where helix lengths differ from the norm such as helix 15 (6 vs. 4 bp), 17 (5 vs. 8 bp), 21 (8 vs. 7 bp), the internal portion of helix 24 (additional 3 bp), and the external portion of helix 46 (truncated to 9 nucleotides). Helix 48 is typically composed of 8 base pairs but differs in *Pilostyles* by the addition of 27 nucleotides. To determine the location of the insertion(s), the conserved apical motif (UCAGYHAU) must be identified; however, this signature sequence is not apparent in *Pilostyles*. If the sequence ACGGUGAA is used as the apex (50% similar to signature), an 8 base pair helix with an 8 nucleotide loop can be drawn, thus indicating two insertions in the unpaired (loop) portion at the 5′ and 3′ base of the helix.

Bdallophyton is atypical in its length of helix 10. Whereas eubacteria rRNAs show quite some length variation for this helix (3, 10 or 11 bp), 3 bp are more common in chloroplasts. For this reason, the 9 bp extension in *Bdallophyton* is quite unusual (Figure 2). Helix 29 is 10 or 11 bp in length in eubacteria; however, significant length variation occurs in plastid 16S rRNAs such as *Chlamydomonas* and *Euglena* (8 bp), *Marchantia* (10 bp), *Nicotiana* and most other plants (11 bp), *Cyanidium* (12 bp) and *Astasia* (28 bp). In addition to *Pilostyles* (12 bp plus the additional 92 nucleotides), variation in the length of this helix can be seen in *Corynaea*, *Cynomorium*, and *Mitrastema* (8, 10 and 11 bp, respectively; Figure 2). The irregular helix 43 contains an asymmetrical interior loop and frequent noncanonicals. For most species, the helix is 12 bp long, but variation is seen in *Cynomorium*, *Mitrastema* (both with 11 bp) and *Corynaea* (10 bp). *Pilostyles* has several insertions (Us) and a longer helix (by 3 bp). The external portion of helix 49 (U_{1371} to A_{1403} in *Cynomorium*) is more variable in sequence and length among the holoparasites. As shown in Figure 2, truncations of this region occur in several species making the length of entire helix variable: *Cynomorium* (103 bp), *Cytinus* (104 bp), *Bdallophyton* (102 bp), *Corynaea* (90 bp), *Hydnora* and *Pilostyles* (both with 87 bp), and *Mitrastema* (81 bp). High sequence variation among the holoparasites makes the task of constructing comparable structures for the exterior portion of helix 49 difficult.

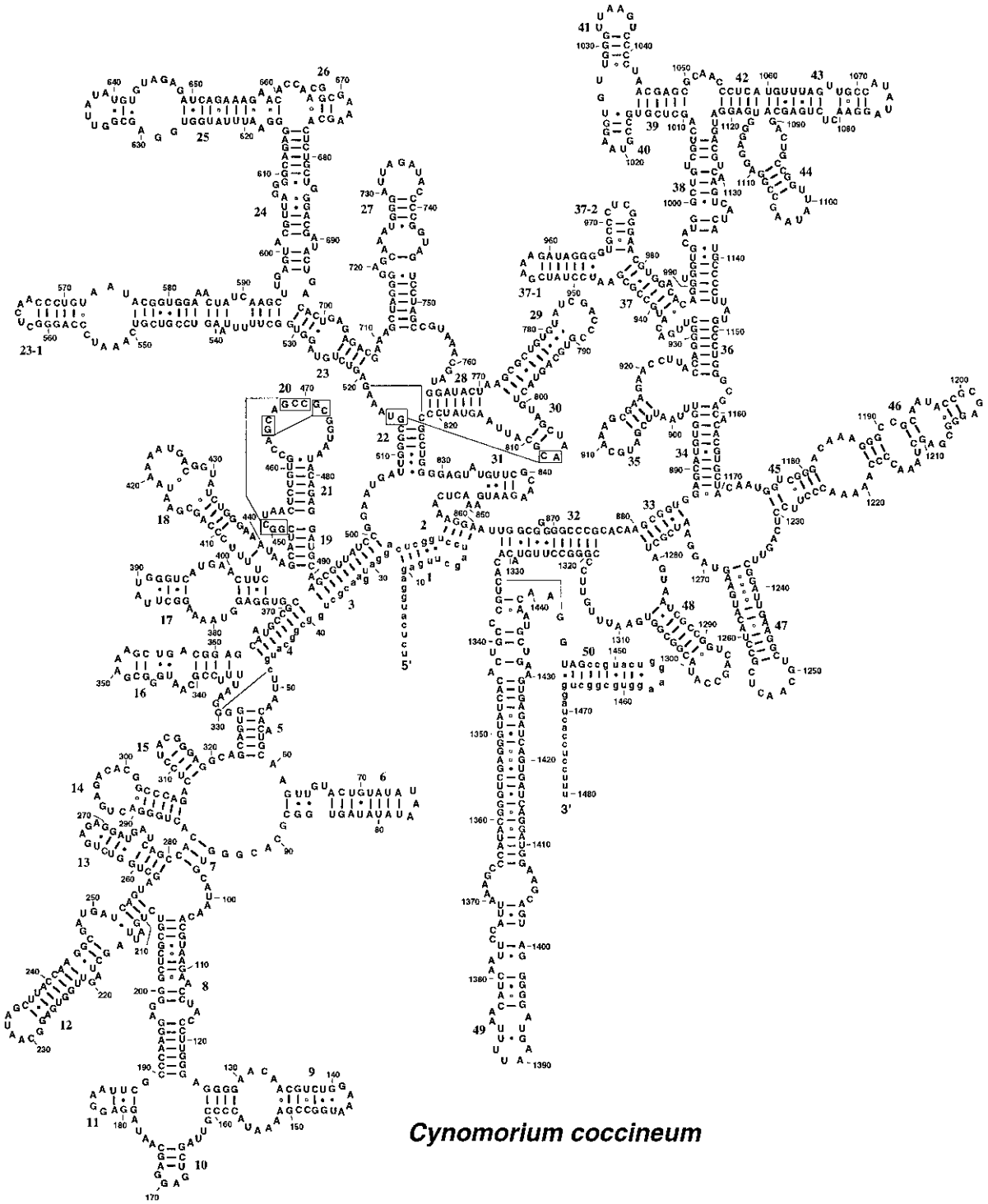

Figure 1. Higher-order structural model for plastid-encoded small-subunit (SSU, 16S) ribosomal RNA for *Cynomorium coccineum* (Cynomoriaceae). Base pairing follows the covariance model proposed for *Nicotiana* by Gutell [15]. Lowercase bases at the 5′ and 3′ ends of the molecule were not determined because they are near priming sites; the sequence of *Nicotiana* is shown for clarity.

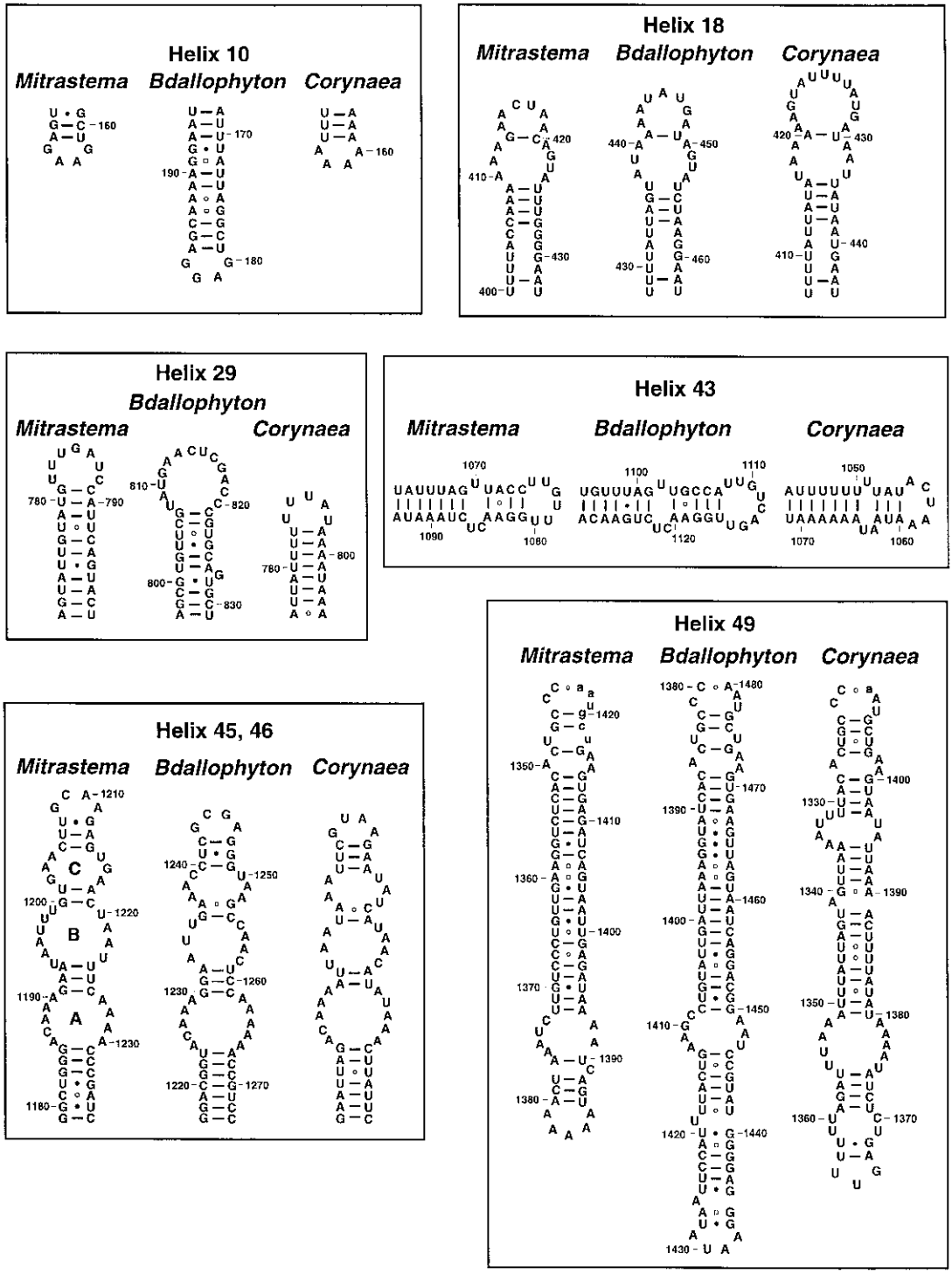

Figure 2. Structural models of selected variable helices for three holoparasitic angiosperms. The shaded regions in the loop of helix 18 represent proposed insertions in the conserved AAURA apical sequence. Noncanonicals are proposed based upon base pairs seen in models of other SSU rRNAs.

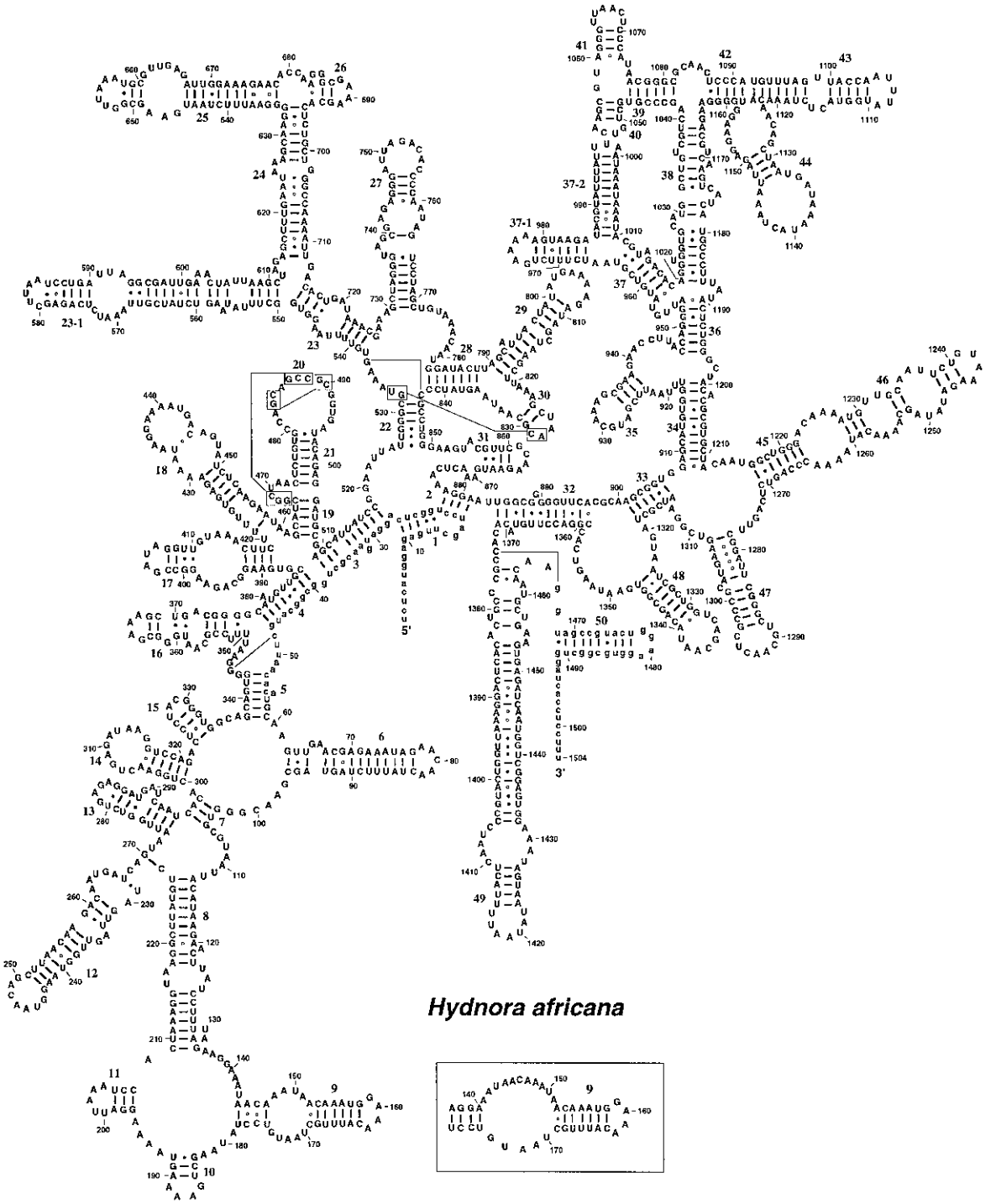

Figure 3. Higher-order structural model for plastid-encoded SSU ribosomal RNA for *Hydnora africana* (Hydnoraceae). Base pairing follows the covariance model proposed for *Nicotiana* by Gutell [15]. Lowercase bases at the 5′ and 3′ ends of the molecule were not determined because they are near priming sites; the sequence of *Nicotiana* is shown for clarity. Given our uncertainty as to the placement of the novel insertion at position 136, two alternative structural models are presented.

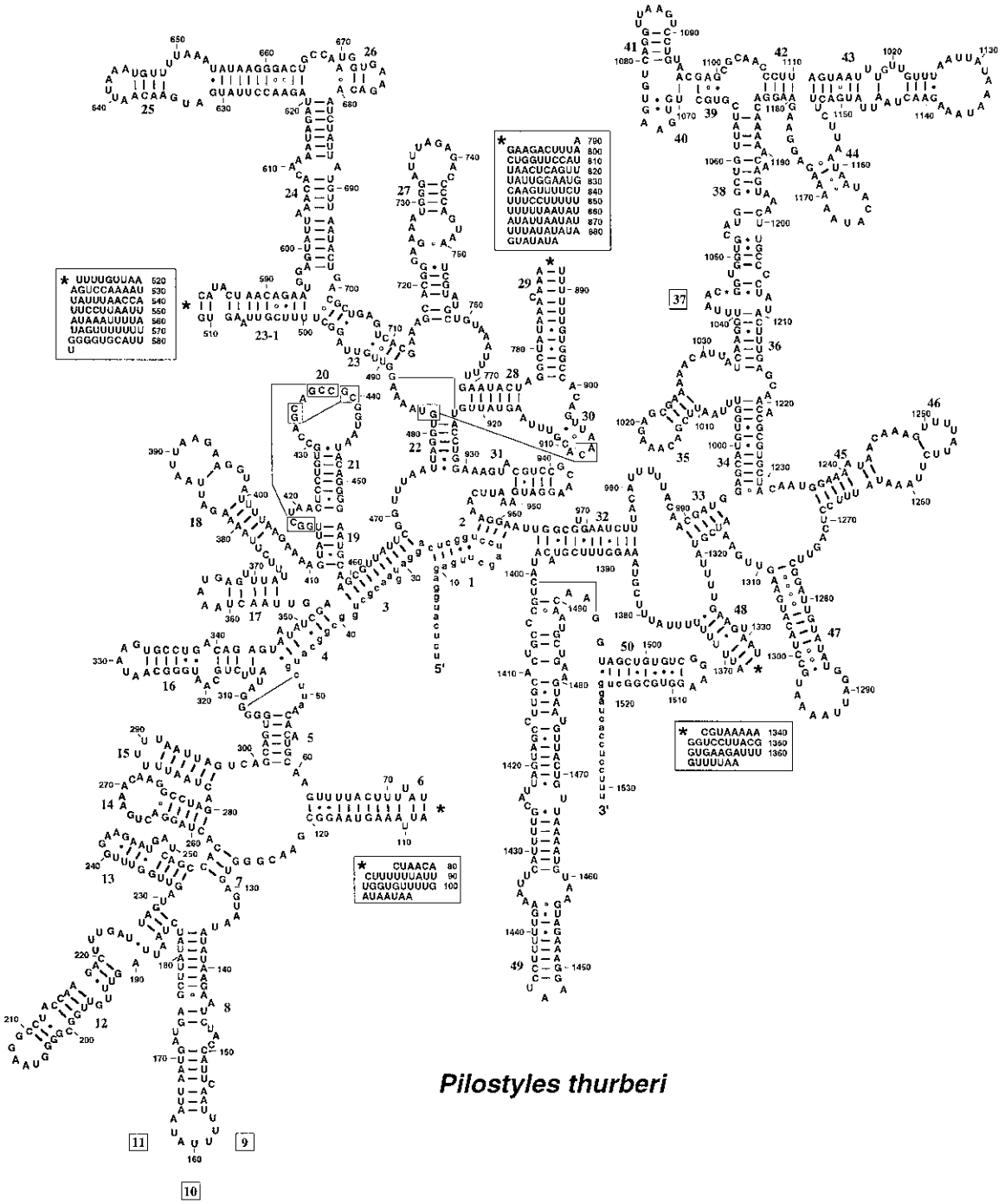

Figure 4. Higher-order structural model for plastid-encoded SSU ribosomal RNA for *Pilostyles thurberi* (Apodanthaceae). Base pairing follows the covariance model proposed for *Nicotiana* by Gutell [15]. Lowercase bases at the 5′ and 3′ ends of the molecule were not determined because they are near priming sites; the sequence of *Nicotiana* is shown for clarity. Helices 9, 10, 11 and 37 are absent from this rRNA. Sequences shown in boxes adjacent to helices 6, 23-1, 29, and 48 represent extensions not observed in other plastid SSU rRNAs, hence no structure is shown.

Loop length and sequence variation

Another characteristic that shows variation in the holoparasite rRNA structures is the length and sequence composition of loop regions. *Hydnora* shows an additional 7 nucleotides at ca. position 136 that could be accomodated in the region between helix 8 and 9 or within the interior loop of helix 9 (two alternatives shown in Figure 3). Helix 11 consists of 3 pairs in all species and is terminated in most plastid 16S rRNAs by a tetraloop of GRAA. Sequence and length variations for this apex occur in *Mitrastema*, *Bdallophyton*, *Hydnora*, *Cytinus*, and *Cornynaea*. Most (97%) green plant chloroplasts have an apical tetraloop for helix 15 of UACG [22] and UAYG is sufficient for all holoparasites except *Pilostyles* (U_4) and *Corynaea* (U_8). The apex of helix 16 is typically a tetraloop of GAAA but is AUAAG in *Pilostyles*. The motif AAUGA can be found in the terminus of helix 18; however, insertions commonly interrupt this motif: *Mitrastema* (C), *Bdallophyton* (AU), *Pilostyles* (UAA) and *Corynaea* (U_4). The length of the apical loop of helix 29 is variable: 5 nucleotides in *Cornynaea*, 7 in *Mitrastema*, 8 in *Cynomorium* (typical length) and 13 in *Bdallophyton* (Figure 2). The apical loop of helix 43 is variable in length among all plastid 16S rRNAs, as is the case for the holoparasites: *Hydnora*, *Cynomorium*, *Corynaea* (all with 5 nucleotides), *Mitrastema* (6 nt), and *Bdallophyton* (8 nt). No apparent consensus sequence can be ascertained for this apical loop. A highly unusual feature of the *Pilostyles* SSU rRNA is the presence of additional bases in the normally conserved loop region separating helix 32 and 33. The conserved CACA motif apparently changed to UACA, duplicated, and added a U_4 insertion. The hairpin loop of helix 44 is generally composed of 7 nucleotides but is expanded to 14 nucleotides in *Hydnora*. The apical loop of helix 47 consists of 8 nucleotides with a signature of UGMAACUC in plastids. A similar signature (GCAACUC) is found in 88% of eubacteria [49]. Despite this extensive conservation, *Pilostyles* contains a 10 nucleotide apical loop (GGAUUAAAAU).

Absence of helices

Unlike all other plastid-encoded 16S rRNAs, *Pilostyles* shows deletions of entire helical regions. The multistem structure that includes helix 9, 10 and 11 is missing in this holoparasite (Figure 4). Also entirely lacking are helices 37, 37-1, and 37-2. These deletions, as well as the substantial insertions (above) invite the question as to whether the *Pilostyles* 16S rRNA is functional in plastid translation. Reverse transcriptase PCR resulted in a product whose sequence was identical to the one previous generated directly from genomic DNA. These results indicate that the unusual insertions and deletions are present on a ribosomal RNA. Although these data provide evidence that this SSU rRNA is transcribed, no information exists as to the functionality of these rRNAs in translation.

Increase of noncanonical base pairs

Compared to other plastid 16S rRNA sequences, helices of the holoparasites show an increase in noncanonical base pairs. For example, in addition to the A•C pair typically adjacent to the interior loop of helix 9, several additional noncanonicals are found: U•C (*Cynomorium*), U•U (*Cytinus*), A•C and C•C in *Hydnora*, and A•C and U•C (*Mitrastema*). Other targets for noncanonicals are helices 8, 23 (*Pilostyles*), 37-1, 43, 45, 46 and the internal portion of 49. For the latter helix, noncanonical G•A pairs appear to be phylogenetically constrained.

A/U composition bias

Compared with plastid 16S rRNA in general, the holoparasites 16S RNAs show an increased frequency of A and U (Table 1). The holoparasite rRNAs that show higher substitution rates and more extreme structural modifications also show the most extreme composition bias. Like the noncanonicals, this feature is particularly prevalent in a subset of helices, including helices 6, 9, 15, 18, 23, 24, 29, 37-1 and 48. In particular, for helix 6 and 18, most holoparasites show an A+U content at least 10% higher than the mean obtained from 17 of the longer helices on the molecule. In *Corynaea*, helix 9, 10 and 11 consist entirely of A and U. The same bias is shown in some helices in *Pilostyles*: 23-1 (75%), 24 (77%), 29 (77%), and 48 (84%). The increase in A+U content of helix 29 is shown in *Mitrastema* and *Hydnora* as well as in the euglenophyte *Astasia longa* (76%), a nonphotosynthetic relative of *Euglena*. That a similar composition bias exists for this helix in both nonphotosynthetic flagellates and holoparasitic angiosperms suggests a relationship between this region and relaxation of selection pressure on plastid-encoded rRNA accompanying the loss of photosynthesis.

Identity, structure and function of holoparasite plastid 16S rRNA

Given the high degree of structural modification seen in the holoparasite 16S rRNAs reported herein, questions arise as to their identity and functionality. That these sequences are indeed of plastid origin (as opposed to mitochondrial or bacterial) is indicated by their possession of a number of plastid-specific signature sequences [37]. All holoparasite 16S rDNA sequences are shown to be most similar to other plastid SSU sequences following BLAST searches and phylogenetic analyses using distance and character-based algorithms unambiguously place these sequences on long branches within an angiosperm clade (data not shown). These data provide evidence for the existence of plastid SSU sequence in these plants; however, the actual subcellular location of these genes awaits further analyses. Since 16S rDNA sequences have been obtained from representatives of all lineages of nonasterid holoparasites, our working hypothesis is that all nonphotosynthetic plants retain a vestigial plastid genome [37]. Given the RT-PCR results for *Pilostyles*, we do not believe the SSU sequence represents a primary pseudogene. Although pseudogenes for two tRNA genes are present in *Epifagus* [29, 45, 52], no reports exist of a landplant plastid SSU pseudogene (note, however, the presence of a 'truncated' copy in *rrnC* in the euglenophyte *Astasia* [42]).

Are the losses of helices in *Pilostyles* suggestive of nonfunctionality of the rRNA? Similar losses of helices in variable regions can be seen in functional SSU rRNAs of eukaryotes and mitochondria, thus structural modifications of the degree reported for the holoparasites is not prima facie evidence for nonfunctionality. With the exception of *Pilostyles*, none of the holoparasites have lost any of the 50 helices that are present in all plastid and bacterial 16S rRNAs. Despite high mutation pressure and lengthening and shortening of particular helices, selection has acted to maintain these structural elements. Without direct experimental evidence, it is difficult to determine the functional significance of rare structural changes, such as the presence of novel extensions to helices as found in *Pilostyles*, *Bdallophyton*, or *Hydnora*. More subtle but equally rare changes, such as the lengthening of helix 21 by one C-G base pair in *Pilostyles*, requires two precisely placed insertions, thus further suggesting a functional constraint. For all 16S rRNAs, mutations are not random but occur in a mosaic pattern in regions that are typically variable in other SSU sequences. Specific regions identified to be critical to ribosome function (such as the interior of helix 49) have been conserved in sequence and structure in all holoparasites. If these were pseudogenes, the fixation of random mutations in these regions would be as likely as other regions. It is therefore proposed that these unusual SSU rRNAs have been the targets of mutation rate increases that have arisen following a relaxation of selection pressure (loss of photosynthesis). In contrast to the holoparasite *Epifagus*, this process has become exaggerated (and/or has had a longer time to act) in the nonasterid holoparasites.

The effect of compositional bias and neighboring bases on mutation rates

The G+C content of plastid SSU sequences of photosynthetic landplants varies little and has a mean of 56% (25 plants compared). In contrast to all other landplants, the G+C content of holoparasite SSU rDNA sequences ranges widely from near 50% to as low as 28% in *Corynaea* (with a mean a 42.9% for seven holoparasites; Table 1). A departure from the typical base composition in favor of higher A+T content in the rDNA, (hence higher A+U in the rRNA) will be referred to here as the 'A/T drift' phenomenon. A composition bias can also be seen in nuclear-encoded 18S rRNA sequences of nonasterid holoparasites; however, here, C to T transitions outnumber G to A transitions approximately three to one ([38] and Nickrent and Colwell, unpublished). Complete mitochondrial SSU sequences have been obtained from *Cytinus*, *Hydnora* and *Corynaea* [8]. Despite the fact that these sequences show increased rates of substitution relative to other angiosperms, they do not show the A/T drift phenomenon. Taken together, these data indicate that this phenomenon is not organism-wide and that different processes are occurring in the subcellular compartments and among varying gene regions.

The A/T drift phenomenon does not appear to be restricted to parasitic plants but can also be seen in other parasitic and/or nonphotosynthetic organisms. Endocellular parasitic bacteria such as *Borrelia*, *Rickettsia* and *Mycoplasma* that have comparatively small genomes also show extreme A+T bias [1]. The A+T content of the vestigial plastid genome (52% of the sequence determined) present in the malarial parasite *Plasmodium falciparum* is 86.9% [48]. Finally, the 73 kb plastome of the nonphotosynthetic euglenoid flagellate *Astasia longa* is 80.2% A+T as compared with its photosynthetic relative *Euglena gracilis* at 73.8%

[18, 42]. The SSU rRNA of *Astasia* is similarly biased (59.5% vs. 52.9%). These examples suggest that the A/T drift phenomenon may be pervasive in organisms that adopt highly specialized trophic modes.

Studies of the dynamics of nucleotide substitutions have often utilized noncoding spacer regions or pseudogenes since they are assumed to be under no selective constraints and would therefore represent neutral models of spontaneous mutation [24]. Analyses of human gene/pseudogene pairs have shown that substitution rates are highly dependent upon the composition of neighboring bases [19]. Transversion rates in mitochondrial SSU sequences of mushrooms have also been shown to be strongly affected by A+T bias [2]. For chloroplast genomes, a comparison of 19 noncoding regions in rice and maize showed that the transversion/transition ratio increased as the A+T content increased [30]. That study showed that when the 5′- and 3′-flanking nucleotides are G or C, 25% of the observed substitutions are transversions whereas when the flanking nucleotides are A to T, transversions constitute 57% of the substitutions. To determine whether neighboring base composition biases the substitution type in rDNA of holoparasites, the sequence of *Cytinus* was compared to tobacco. For simplicity, tandem mutations and indels were excluded thereby leaving 90 of the possible 121 substitutional changes for this analysis. In close agreement with [30], transversions represent 26% of the mutations when flanking bases were G or C. The percentage of transversions increased to 36% when the flanking bases were A or T, thus showing that a similar trend exists in plastid rDNA as exists in noncoding spacers. The lower value (36% vs. 57%) likely stems from higher functional constraint on rDNA as compared with noncoding regions.

Proofreading and mismatch repair

Because rate increases and compositional effects do not always occur coordinately in the subcellular genomes of parasites, the underlying cause may stem from variations in their respective DNA replication and repair systems [5, 38, 46]. Error correction is accomplished by the 3′ to 5′ exonuclease activity of DNA polymerase I and III whereas the postreplicative repair system excises mismatches by recognizing the methylation state of template and newly synthesized strands of DNA. A recently described system from *Escherichia coli* results in increased levels of transversions, especially A-T to T-A and G-C to T-A [43]. Here, mutations in the *mutA* and *mutC* alleles result in a changed anticodon of glycine tRNAs that mistranslate polymerase III thereby increasing its error rate. Higher transversion rates have also been documented in studies of plastid DNA noncoding spacers [30] and *Alu* sequences [11], both of which suggest that mismatch repair is influenced by neighboring base composition. A frequent and spontaneous mutation is the deamination of cytosine to uracil which results in G-C pairs changing to A-U in the first round of replication and then A-T in the second round. Thus, without error correction, DNA tends towards higher proportions of A-T. The excision repair process that corrects uracil back to cytosine involves a suite of enzymes (uracil N-glycosylase, AP endonuclease, exonuclease, etc.) that must function properly to correct errors. Mutations that negatively affect such repair systems or failure to import these enzymes into the chloroplast could result in the A/T drift phenomenon.

Conclusions

The higher-order structures constructed for plastid-encoded SSU rRNAs of nonasterid holoparasitic angiosperms show an increased base substitution rate that is accompanied by an 'A/T drift' phenomenon. These sequences are the most divergent plastid SSU sequences documented for plants and therefore serve as valuable models for the study a number of molecular evolutionary processes. Rate increases for nuclear [38] and mitochondrial [8] SSU rRNA have also occurred in these plants; however, the magnitude and types of changes are not equivalent to those observed in the plastid 16S sequences. The novel helix extensions and losses shown for *Pilostyles* do not appear to be introns as ascertained by RT-PCR experiments. A number of causes have been proposed to account for increased mutation and fixation rates such as generation time, population size, organellar location of the gene, polymerase fidelity, DNA replication and repair, etc. (see discussion in [38]) The underlying molecular mechanism associated with these high substitution rates is presently unknown and may differ among various organisms, genes and nutritional modes. At present, we favor a hypothesis whereby mutations in DNA replication and/or repair systems have not been eliminated in the relaxed selectional environment of a nonphotosynthetic plastid.

Acknowledgements

This research was supported by a grant from NSF (DEB 94-07894) and a Special Research grant from the Office of Research Development and Administration, SIUC. We are indebted to the following people who provided tissue samples of these unusual parasitic plants: S. Carlquist (*Hydnora*), W. Barthlott (*Cytinus*), D. Joel and O. Cohen (*Cynomorium*), W. Meijer (*Mitrastema*), and D. Seigler (*Bdallophyton*, No. 14099). Thanks go to R. Gutell for helpful comments about the higher-order structures.

References

1. Andersson SGE, Kurland CG: Genomic evolution drives the evolution of the translational system. Biochem Cell Biol 73: 775–787 (1995).
2. Bruns TD, Szaro TM: Rate and mode differences between nuclear and mitochondrial small-subunit rRNA genes in mushrooms. Mol Biol Evol 9: 836–855 (1992).
3. dePamphilis CW: Genes and genomes. In: Press MC, Graves JD (eds) Parasitic Plants, pp. 176–205. Chapman and Hall, London (1995).
4. dePamphilis CW, Palmer JD: Loss of photosynthetic and chlororespiratory genes from the plastid genome of a parasitic flowering plant. Nature 348: 337–339 (1990).
5. dePamphilis CW, Young ND, Wolfe AD: Evolution of plastid gene *rps*2 in a lineage of hemiparasitic and holoparasitic plants: many losses of photosynthesis and complex patterns of rate variation. Proc Natl Acad Sci USA, in press (1997).
6. Dixon MT, Hillis DM: Ribosomal RNA secondary structure: compensatory mutations and implications for phylogenetic analysis. Mol Biol Evol 10: 256–267 (1993).
7. Douglas SE, Turner S: Molecular evidence for the origin of plastids from a cyanobacterium-like ancestor. J Mol Evol 33: 266–273 (1991).
8. Duff RJ, Nickrent DL: Characterization of mitochondrial SSU (18S) rDNA from holoparasitic plants. J Mol Evol (submitted).
9. Ehresmann B, Ehresmann C, Romby P, Mougel M, Baudin F, Westhof E, Ebel J-P: Detailed Structures of rRNAs: New Approaches. American Society for Microbiology, Washington, DC (1990).
10. Ems SC, Morden CW, Dixon CK, Wolfe KH, dePamphilis CW, Palmer JD: Transcription, splicing and editing of plastid RNAs in the nonphotosynthetic plant *Epifagus virginiana*. Plant Mol Biol 29: 721–733 (1995).
11. Filipski J, Salinas J, Rodier F: Chromosome localization-dependent compositional bias of point mutations in *Alu* repetive sequences. J Mol Biol 206: 563–566 (1989).
12. Gilbert DG: SeqApp, version 1.9a157. Biocomputing Office, Biology Dept., Indiana University, Bloomington, IN 47405 (1993).
13. Gutell RR: Collection of small subunit (16S- and 16S-like) ribosomal RNA structures: 1994. Nucl Acids Res 22: 3502–3507 (1994).
14. Gutell RR: Collection of small subunit (16S- and 16S-like) ribosomal RNA structures. Nucl Acids Res 21: 3051–3054 (1993).
15. Gutell RR: Comparative studies of RNA: inferring higher-order structure from patterns of sequence variation. Curr Opin Struct Biol 3: 313–322 (1993).
16. Gutell RR, Weiser B, Woese CR, Noller HF: Comparative anatomy of 16S-like ribosomal RNA. Progress Nucl Acids Res Mol Biol 32: 155–216 (1985).
17. Gutell RR, Woese CR: Higher order structural elements in ribosomal RNAs: pseudo-knots and the use of noncanonical pairs. Proc Natl Acad Sci USA 87: 663–667 (1990).
18. Hallick RB, Hong L, Drager RG, Favreau MR, Monfort A, Orsat B, Speilmann A, Stutz E: Complete sequence of *Euglena gracilis* chloroplast DNA. Nucl Acids Res 21: 3537–3544 (1993).
19. Hess ST, Blake JD, Blake RD: Wide variations in neighbor-dependent substitution rates. J Mol Biol 236: 1022–1033 (1994).
20. Hill WA, Dahlberg A, Garrett RA, Moore PB, Schlessinger D, Warner JR: The Ribosome: Structure, Function and Evolution. American Society for Microbiology, Washington, DC (1990).
21. Jaeger JA, Turner DH, Zuker M: Predicting optimal and suboptimal secondary structure for RNA. Meth Enzymol 183: 281–306 (1990).
22. Konings DAM, Gutell RR: A comparison of thermodynamic foldings with comparatively derived structures of 16S and 16S-like rRNAs. RNA 1: 559–574 (1995).
23. Kuijt J: The Biology of Parasitic Flowering Plants. University of California Press, Berkeley, CA (1969).
24. Li W-H, Wu C-I, Luo C-C: Nonrandomness of point mutation as reflected in nucleotide substitutions in pseudogenes and its evolutionary implications. J Mol Evol 21: 58–71 (1984).
25. Maidak BL, Larsen N, McCaughey MJ, Overbeek R, Olsen GJ, Fogel K, Blandy J, Woese CR: The ribosomal database project. Nucl Acids Res 22: 3485–3487 (1994).
26. Manhart J: Chloroplast 16S rDNA sequences and phylogenetic relationships of fern allies and ferns. Am Fern J 85: 182–192 (1995).
27. Mishler BD, Thrall PH, Hopple JSJ, De Luna E, Vilgalys R: A molecular approach to the phylogeny of bryophytes: cladistic analysis of chloroplast-encoded 16S and 23S ribosomal RNA genes. Bryologist 95: 172–180 (1992).
28. Morden CW, Delwiche CF, Kuhsel M, Palmer JD: Gene phylogenies and the endosymbiotic origin of plastids. BioSystems 28: 75–90 (1992).
29. Morden CW, Wolfe KH, dePamphilis CW, Palmer JD: Plastid translation and transcription genes in a nonphotosynthetic plant: intact, missing and pseudo genes. EMBO J 10: 3281–3288 (1991).
30. Morton BR: Neighboring base composition and transversion/transition bias in a comparison of rice and maize chloroplast noncoding regions. Proc Natl Acad Sci USA USA 92: 9717–9721 (1995).
31. Muse SV: Evolutionary analysis of DNA sequences subject to constraints on secondary structure. Genetics 139: 1429–1439 (1995).
32. Neefs J-M, van de Peer Y, de Rijk P, Chapelle S, de Wachter R: Compilation of small subunit RNA structures. Nucl Acids Res 21: 3025–3049 (1993).
33. Nelissen B, van de Peer Y, Wilmotte A, de Wachter R: An early origin of plastids within the cyanobacterial divergence is suggested by evolutionary trees based on complete 16S rRNA sequences. Mol Biol Evol 12: 1166–1173 (1995).
34. Nickrent DL: From field to film: rapid sequencing methods for field collected plant species. BioTechniques 16: 470–475 (1994).

35. Nickrent DL, Duff JR, Colwell AE, Wolfe AD, Young ND, Steiner KE, dePamphilis CW: Molecular phylogenetic and evolutionary studies of parasitic plants. In Soltis DE, Soltis PS, Doyle JJ (ed) Molecular Systematics of Plants, 2nd. ed., Chapman and Hall, New York, in press (1997).
36. Nickrent DL, Duff RJ: Molecular studies of parasitic plants using ribosomal RNA. In: Moreno MT, Cubero JI, Berner D, Joel D, Musselman LJ, Parker C (eds) Advances in Parasitic Plant Research, pp. 28–52. Junta de Andalucia, Dirección General de Investigación Agraria, Cordoba, Spain (1996).
37. Nickrent DL, Ouyang Y, Duff RJ, dePamphilis CW: Do nonasterid holoparasitic flowering plants have plastid genomes? Plant Mol Biol 34: 717–729 (1997).
38. Nickrent DL, Starr EM: High rates of nucleotide substitution in nuclear small-subunit (18S) rDNA from holoparasitic flowering plants. J Mol Evol 39: 62–70 (1994).
39. Noller HF, Hoffarth V, Zimniak L: Unusual resistance of peptidyl transferase to protein extraction procedures. Science 256: 1416–1419 (1992).
40. Palmer JD: Contrasting modes and tempos of genome evolution in land plant organelles. Trends Genet 6: 115–120 (1990).
41. Palmer JD: A genetic rainbow of plastids. Nature 364: 762–763 (1993).
42. Siemeister G, Hachtel W: Organization and nucleotide sequence of ribosomal RNA genes on a circular 73 kbp DNA from the colourless flagellate *Astasia longa*. Curr Genet 17: 433–438 (1990).
43. Slupska MM, Baikalov C, Lloyd R, Miller JH: Mutator tRNAs are encoded by the *Escherichia coli* mutator genes *mutA* and *mutC*: a novel pathway for mutagenesis. Proc Natl Acad Sci USA USA 93: 4380–4385 (1996).
44. Takhtajan AL: Sistema magnoliofitov [in Russian]. Nauka, Leningrad (1987).
45. Taylor GW, Wolfe KH, Morden CW, dePamphilis CW, Palmer JD: Lack of a functional plastid $tRNA^{cys}$ gene is associated with loss of photosynthesis in a lineage of parasitic plants. Curr Genet 20: 515–518 (1991).
46. Topal M, Fresco J: Complementary base pairing and the origin of substitution mutations. Nature 263: 285–289 (1976).
47. van de Peer Y, van den Broeck I, de Rijk P, de Wachter R: Database on the structure of small ribosomal subunit RNA. Nucl Acids Res 22: 3488–3494 (1994).
48. Wilson RJM, Denny PW, Preiser PR, Rangachari K, Roberts K, Roy A, Whyte A, Strath M, Moore DJ, Moore PW, Williamson DH: Complete gene map of the plastid-like DNA of the malaria parasite *Plasmodium falciparum*. J Mol Biol 261: 155–172 (1996).
49. Woese CR, Gutell R, Gupta R, Noller HF: Detailed analysis of the higher-order structure of 16S-like ribosomal ribonucleic acids. Microbiol Rev 47: 621–669 (1983).
50. Wolfe KH, Katz-Downie DS, Morden CW, Palmer JD: Evolution of the plastid ribosomal RNA operon in a nongreen parasitic plant: accelerated sequence evolution, altered promoter structure, and tRNA pseudogenes. Plant Mol Biol 18: 1037–1048 (1992).
51. Wolfe KH, Morden CW, Palmer JD: Function and evolution of a minimal plastid genome from a nonphotosynthetic parasitic plant. Proc Natl Acad Sci USA USA 89: 10648–10652 (1992).
52. Wolfe KH, Morden CW, Palmer JD: Small single-copy region of plastid DNA in the non-photosynthetic angiosperm *Epifagus virginiana* contains only two genes. J Mol Biol 223: 95–104 (1992).
53. Zuker M, Jaeger JA, Turner DH: A comparison of optimal and suboptimal RNA secondary structures predicted by free energy minimization with structures determined by phylogenetic comparison. Nucl Acids Res 19: 2707–2714 (1991).