Studies of plastid 16S rDNA from nonphotosynthetic plants suggest that the accelerated rate is not a ubiquitous phenomenon

Shin-Yi Shyu and Jer-Ming Hu
jmhu@ntu.edu.tw
Institute of Ecology and Evolutionary Biology, National Taiwan University, Taipei, Taiwan

Abstract

Nonphotosynthetic plants usually retain a reduced plastid genome, although most of the photosynthesis-related genes are lost or become pseudogenes. In order to elucidate the plastid genome evolution in non-photosynthetic plants, we examined plastid-derived 16S rDNAs of nine nonphotosynthetic plants, including some with partial photosynthetic ability. The 16S rDNA sequences from all examined taxa were included in a phylogenetic analysis with other land plants, algae, and cyanobacteria to show their relationships. To evaluate rate heterogeneity among various 16S rDNA sequences, the relative rate tests were conducted. Pinus was chosen as the reference in this study. Compared with other angiosperms, the 16S rDNA sequences show an increasing substitution rates in four nonphotosynthetic species: Balanophora laxiflora, Mitrastemon kanehirai, Cheilotheca humilis and Cheilotheca macrocarpa. However the other five species, Aeginetia indica, Cassytha filiformi, Cuscuta australis, Galeola lindleyana, Orobanche coerulescens, exhibited only slightly higher rates relative to most angiosperms. Some of the high sequence divergence are accompanied by an increase in A+T content of the sequences, especially in Mitrastemon kanehirai, which has the highest rate among the examined taxa. In species with higher evolutionary rates of 16S rDNAs, Balanophora laxiflora, Mitrastemon kanehirai are holoparasites and Cheilotheca humilis, Cheilotheca macrocarpa are mycotrophic plants. Theses plants have completely lost their photosynthetic ability. However, among the other five species examined, Cassytha filiformi is a hemiparasitic plant and can photosynthesize itself; Galeola lindleyana and young stage of Cuscuta australis are slight green and maybe retain some photosynthetic ability. Although Aeginetia indica and Orobanche coerulescens are holoparasite plants, Orobanchaceae parasites lost their photosynthetic ability recently, and may not have enough time to accumulate variations.

Keywords: parasitic plants, mycotrophic plants, rDNA and plastid sequence data
**Introduction**

Plastids are the organelles mainly responsible for photosynthesis in plant cells, and were endosymbionts derived from previously free-living cyanobacteria. However, plastid genomes encode only 5~10% as many genes as those of cyanobacteria in the present day. The plastid genomes of photosynthetic plants are circular molecules ranging in size from about 150~200kb (Martin et al. 2002, Palmer & Delwiche 2000). The genes in plastids can be divided into three categories: transcription, translation related genes, photosynthesis related genes and genes related other biosynthesis (Race et al. 1999, Timmis et al. 2004).

The nonphotosynthetic plants are heterotrophic angiosperms that completely lost their photosynthetic ability, and can be classified as either mycotrophs or haustorial parasites (Nickrent et al. 2002). The former obtain nutrients via a symbiotic relationship with mycorrhizal fungi, whereas haustorial parasites directly obtain water and nutrients from host via haustoria.

The cells of nonphotosynthetic plants retain remnant plastids and the plastid genome is highly reduced (dePamphilis & Palmer 1990, Nickrent et al. 1997a). Nonphotosynthetic plants, therefore, provide an ideal model to study evolution and function of plastid genome and genes without selection pressures. Many of the plastid genes are lost or become pseudogenes in nonphotosynthetic plants, especially the photosynthesis-related genes such as *rbcL* (Wolfe et al. 1992). On the other hand the genes responsible for plastid gene expression like rDNA and ribosomal proteins are most often retained. However, there are some studies that show these translation-related genes such as *rps2* and 16S rDNA have experienced rate accelerations in several parasitic plants (dePamphilis et al. 1997, Nickrent et al. 1997b).

In this study, we examined plastid-derived 16S rDNAs of eight nonphotosynthetic plants, including some with partial photosynthetic ability: *Aeginetia indica*, *Balanophora laxiflora*, *Cheilotheca humilis*, *Cheilotheca macrocarpa*, *Cuscuta australis*, *Galeola lindleyana*, *Mitrastemon kanehirai* and *Orobanche coerulescens*, and one hemiparasite *Cassytha filiformi*. The 16S rDNA sequences from selected taxa were used to perform phylogenetic analyses and examine their evolutionary rates via relative rate tests.

**Materials and methods**

*Plant Materials.* The plant species for DNA extraction and associated information are listed in Table 1. Nine species, including one hemiparasite, five holoparasities and three mycotrophic plants, native to Taiwan were collected and store in -20°C.
DNA Extraction, PCR Amplification, and Sequencing. Fresh or frozen tissues were used for DNA extraction. Total genomic DNAs were extracted by using the modified CTAB methods (Barnwell et al. 1998, Doyle & Doyle 1987, Nickrent 1994). The primers used for PCR amplification and sequencing of plastid 16S rDNA are listed in Table 2.

Phylogenetic analyses. The plastid 16S rDNA sequences of other land plants, algae, and cyanobacteria were downloaded from NCBI Genbank. Alignments were conducted by ClustalX 1.83 (Thompson et al. 1997), and modified in MacClade 4.06 (Maddison & Maddison 2000). Phylogenetic analyses for each matrix were conducted using neighbor-joining (NJ) and maximum parsimony (MP) methods were performed by PAUP*4.0.b10 (Swofford 2002). For neighbor joining analyses, the HKY85 nucleotide substitution model was used. For parsimony analyses, heuristic searches were conducted with 1,000 random addition replicates, tree bisection-reconnection (TBR) branch swapping, and ten trees were saved from each replicate.

Relative rate tests. To evaluate rate heterogeneity among various 16S rDNA sequences, the relative rate 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 parasitic angiosperms relative to monocots and dicots is uncertain, Pinus was chosen as the reference (Nickrent et al. 2000). 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 are the sum of the number of transitional and transversional substitutions per site.

Results and discussions

Phylogenetic relationship
Plastid 16S rDNA sequences have seldom been used in phylogenetic studies of angiosperms given their overall sequence conservation. The 16S rDNA sequences from all examined taxa were included in a phylogenetic analysis with other land plants. The MP tree shows their 16S rDNA sequences are most closely related to other angiosperm plastid 16S rDNAs and graphically demonstrates rate increases (Fig. 1). The topology of the tree has the major features coincides with previous study (Nickrent et al. 2000). However the NJ analysis included algae and cyanobacteria had similar result with Nickrent’s study (Nickrent & Starr 1994) analyses using nuclear 18S rDNA sequences of holoparasites resulted in the migration of these long-branch taxa near the base of the tree. For this character-based analysis (parsimony) of the
nonphotosynthetic plants, the types of change, not simply the number of changes, apparently determined the placement of the clade. On the other hand, the interrelationships among the nonphotosynthetic plants and the results of the analyses only included angiosperms are likely artificial because the topological position simply reflect an increasing number of substitutions (Nickrent et al. 2000). These wrong relationships probably come of the long-branch attraction. Long-branch attraction, a bias in certain phylogenetic inference methods in which similarity due to convergent or parallel changes produces an erroneous phylogenetic grouping of taxa. It has been suggested that some data sets with marked among-lineage rate heterogeneity cannot be applied to particular phylogenetic problems owing to hypothesized long-branch attraction artifacts (Soltis et al. 1999). The previous studies indicate analyze data using model-based methods (e.g., maximum likelihood or Bayesian inference methods) are less likely to be misled by long-branch attraction, adding taxa also can “break up” long branches to allow parsimony recovered the correct topology (Nickrent et al. 2004).

16S rDNA sequences variation and accelerated rate
We obtained five full length 16S rDNA sequences, four partial sequences from the selected nonphotosynthetic and partial photosynthetic plants (Table 1). All examined taxa have polymorphism in their 16S rDNA sequences, but only the 16S rDNA sequence of *A. indica* has a large insertion about 14bp (Fig. 2). Most angiosperm 16S rDNA sequences differ by 2-3% when compared with tobacco, whereas the four nonphotosynthetic species in this study show significantly higher substitution: *B. laxiflora* (7%), *M. kanehirai* (15%), *C. humilis* (8%) and *C. macrocarpa* (8%). However the other five species exhibited low substitution rates as in angiosperms: *A. indica* (2%), *C. filiformi* (3%), *C. australis* (2%), *G. lindleyana* (3%), *O. coerulescens* (2%). Figure 3 shows the results of the relative rate tests using plastid 16S rDNA sequences. The magnitude of rate increase among the 9 species matches the substitutions data described above, and these results suggest the accelerated rate is not a ubiquitous phenomenon among nonphotosynthetic plants. The G+C content of the nonphotosynthetic plants 16S rDNAs ranges from 44.4-56.3% (Table 1), 6 out of 9 lower than the mean for more typical land plants (56%). A departure from the typical base composition in favor of higher A+T content in the rDNA, will be referred to here as the ‘A/T drift’ phenomenon (Nickrent et al. 1997b). However, the substitution rate of the 16S sequences didn’t coincide with ‘A/T drift’ phenomenon. These indicate these two phenomena are driven by different evolutionary mechanism. As for whether the high substitution rate of the sequences affect the structure of 16S rDNA, it needs to proceed further structure analysis of sequences.
Conclusion
The 16S plastid rDNAs of some nonphotosynthetic angiosperms show an increased base substitution rate that is accompanied by an ‘A/T drift’ phenomenon. Rate increases for nuclear, mitochondrial and plastid SSU rDNA have documented in other nonphotosynthetic plants. However the results indicate that the accelerated rate is not a ubiquitous phenomenon in nonphotosynthetic plants. Among the 9 species examined, Cassytha filiformi is a hemiparasitic plant and can photosynthesize itself; Galeola lindleyana and young stage of Cuscuta australis are slight green and maybe retain some photosynthetic ability. The substitution rates of the 16S plastid rDNAs in these three species show no difference compared with other angiosperms. The other 5 species are holoparasites or mycotrophic plants, and they have completely lost their photosynthetic ability. But Aeginetia indica and Orobanche coerulescens didn’t exhibit accelerated rates of 16S rDNAs as Balanophora laxiflora, Mitrastemon kanehirai, Cheilotheca humilis and Cheilotheca macrocarpa. This probably because Orobanchaceae parasites lost their photosynthetic ability recently, and may not have enough time to accumulate variations. The underlying molecular mechanism associated with the high substitution rates is presently unclear and may differ among various organisms, genes and nutritional modes.
### Table 1. Nonphotosynthetic plants used to obtain plastid 16S rDNA sequences.

<table>
<thead>
<tr>
<th>Species</th>
<th>Family</th>
<th>Nutrient mode</th>
<th>%GC</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Aeginetia indica</em> L.</td>
<td>Orobanchaceae</td>
<td>Holoparasite</td>
<td>54.4</td>
</tr>
<tr>
<td><em>Balanophora laxiflora</em> Hems.</td>
<td>Balanophoraceae</td>
<td>Holoparasite</td>
<td>56.1</td>
</tr>
<tr>
<td><em>Cassytha filiformis</em> L.</td>
<td>Lauraceae</td>
<td>Hemiparasite</td>
<td>52.1</td>
</tr>
<tr>
<td><em>Cheilotheca humilis</em></td>
<td>Pyrolaceae</td>
<td>Mycotroph</td>
<td>52.3</td>
</tr>
<tr>
<td><em>Cheilotheca macrocarpa</em></td>
<td>Pyrolaceae</td>
<td>Mycotroph</td>
<td>52.3</td>
</tr>
<tr>
<td><em>Cuscuta australis</em> R. Brown</td>
<td>Convolvulaceae</td>
<td>Holoparasite</td>
<td>56.4</td>
</tr>
<tr>
<td><em>Galeola lindleyana</em></td>
<td>Orchidaceae</td>
<td>Mycotroph</td>
<td>55.4</td>
</tr>
<tr>
<td><em>Mitrastron kanehirai</em> Yamamotoa</td>
<td>Mitrastemonaceae</td>
<td>Holoparasite</td>
<td>44.4</td>
</tr>
<tr>
<td><em>Orobanche coerulescens</em> Stephan</td>
<td>Orobanchaceae</td>
<td>Holoparasite</td>
<td>56.3</td>
</tr>
</tbody>
</table>

*a:* partial sequences, size about 1000 bp.

*b:* partial sequences, size about 1470 bp.

### Table 2. Primers used to amplify plant plastid 16S rDNA sequences.

<table>
<thead>
<tr>
<th>Primer name</th>
<th>Primer sequence</th>
<th>Position on <em>Nicotiana</em></th>
</tr>
</thead>
<tbody>
<tr>
<td>Forward</td>
<td></td>
<td></td>
</tr>
<tr>
<td>16S -32F</td>
<td>AACAGGAAGCTATAAGTAATGCAA</td>
<td>102730-102754</td>
</tr>
<tr>
<td>16S 8F</td>
<td>GGAGAGTTTCGATCTGCTGCTAG</td>
<td>102768-102789</td>
</tr>
<tr>
<td>16S 323F</td>
<td>CAGCAGTGGGAAATTTCCG</td>
<td>103084-103103</td>
</tr>
<tr>
<td>16S 856F</td>
<td>AACTCAAAGGAAATTG</td>
<td>103617-103631</td>
</tr>
<tr>
<td>Reverse</td>
<td></td>
<td></td>
</tr>
<tr>
<td>16S 627R</td>
<td>GAAATCCCTCTGCCCCCTAC</td>
<td>103389-103370</td>
</tr>
<tr>
<td>16S 878R</td>
<td>GCCCCCCGTYAATCTCCCT</td>
<td>103640-103625</td>
</tr>
<tr>
<td>16S 1300R</td>
<td>GCGATTACTAGCCGATT</td>
<td>104062-104047</td>
</tr>
<tr>
<td>16S 1525R</td>
<td>AAGGAGGGTATCCAGGCC</td>
<td>104251-104235</td>
</tr>
<tr>
<td>16S +34R</td>
<td>ACCAAAATACCCCAACGCA</td>
<td>104286-104267</td>
</tr>
</tbody>
</table>
Figure 1. One of 3781 MP trees constructed by plastid 16S rDNA sequences. The blue scripts are the 9 species examined in this study, the red ones are the other nonphotosynthetic plants.
Figure 2. A portion of the alignment of plastide 16S rDNA sequences from parasitic and nonparasitic angiosperms. *A. indica* has a large insert about 14bp, *B. laxiflora* and *M. kanehirai* are much more divergent than other species.

Figure 3. The results of relative rate tests using plastid 16S rDNA sequences. The 14 land plants indicated on the abscissa are the test organisms (taxon 1), *Nicotiana* was taxon 2, and *Pinus* 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 purpose. The fastest rates are observed in the four nonphotosynthetic plants.
References


Nickrent, D. L., R. O. Duff, and D. A. M. Konings (1997b) Structural analyses of plastid-derived 16S rRNAs in holoparasitic angiosperms. Plant Mol. Biol. 34:
Nickrent, D. L., R. O. Duff, A. E. Colwell, A. D. Wolfe, N. D. Young, K. E. Steiner,
and C. W. dePamphilis (2000) Molecular phylogenetic and evolutionary studies of
parasitic plants. In Molecular systematics of plants II: DNA sequencing, D. E.
Norwell, MA, USA.
Peninsula and Balearic Islands, J. A. López-Sáez, P. Catalán and L. Sáez (eds.),
Chapter 2, pp. 7-27. Mundi-Prensa Libros, S. A., Madrid.
Phylogenetic inference in Rafflesiales: the influence of rate heterogeneity and
genomes. In Molecular systematics of plants II: DNA sequencing, D. E. Soltis, P. S.
MA, USA.
Race, H. L., R. G. Herrmann, and W. Martin (1999) Why have organelles retained
Sinauer Associates Inc., Sunderland, MA, USA.
Thompson, J. D., T. J. Gilbson, F. Plewniak, F. Jeanmougin, and D. G. Higgins, (1997)
The CLUSTAL-X windows interface: Flexible strategies for multiple sequence
Genet. 5: 123-135.