Rampant Gene Loss in the Underground Orchid *Rhizanthella gardneri* Highlights Evolutionary Constraints on Plastid Genomes

Etienne Delannoy,*[,1,2] Sota Fujii,[1] Catherine Colas des Francs-Small,[1] Mark Brundrett,[3] and Ian Small[1]

[1]Australian Research Council Centre of Excellence in Plant Energy Biology, University of Western Australia, Perth, Australia

[2]Laboratoire de Biologie du Développement des Plantes, Commissariat à l'Energie Atomique, Cadarache, St. Paul Les Durances, France

[3]School of Plant Biology, University of Western Australia, Perth, Australia

***Corresponding author:** E-mail: delannoy@evry.inra.fr.

Associate editor: Charles Delwiche

Abstract

Since the endosymbiotic origin of chloroplasts from cyanobacteria 2 billion years ago, the evolution of plastids has been characterized by massive loss of genes. Most plants and algae depend on photosynthesis for energy and have retained ~110 genes in their chloroplast genome that encode components of the gene expression machinery and subunits of the photosystems. However, nonphotosynthetic parasitic plants have retained a reduced plastid genome, showing that plastids have other essential functions besides photosynthesis. We sequenced the complete plastid genome of the underground orchid, *Rhizanthella gardneri*. This remarkable parasitic subterranean orchid possesses the smallest organelle genome yet described in land plants. With only 20 proteins, 4 rRNAs, and 9 tRNAs encoded in 59,190 bp, it is the least gene-rich plastid genome known to date apart from the fragmented plastid genome of some dinoflagellates. Despite numerous differences, striking similarities with plastid genomes from unrelated parasitic plants identify a minimal set of protein-encoding and tRNA genes required to reside in plant plastids. This prime example of convergent evolution implies shared selective constraints on gene loss or transfer.

Key words: *Rhizanthella gardneri*, mycoheterotroph, chloroplast, tRNA import, gene loss.

Research article

Introduction

Plastid-based photosynthesis is the primary energy source for life on Earth providing oxygen and sugars as the basis of almost every "food chain" scaling up from photosynthetic cells to whole ecosystems. Because the endosymbiosis of cyanobacteria by a nonphotosynthetic eukaryote cell, the evolution of plastids have been characterized by massive gene loss and transfer of most of the remainder to the nucleus (Martin et al. 1998; Archibald 2009). However, nonphotosynthetic plants, algae, and apicomplexan parasites such as *Plasmodium* sp. and *Toxoplasma* sp. have retained a plastid and a plastid genome or "plastome" (Wolfe et al. 1992; Wilson et al. 1996; Funk et al. 2007) suggesting that the loss or transfer of DNA to the nucleus is limited by one or more essential nonphotosynthetic functions of plastids. Analyzing these functions in photosynthetic organisms is complex because perturbations of plastid metabolism have pleiotropic effects and are often deleterious or lethal (Ahlert et al. 2003; de Longevialle et al. 2008). The comparison of the plastid genomes from nonphotosynthetic plants should highlight these essential functions as well as the constraints preventing the transfer of genes from the plastid to the nucleus.

There are two distinctly different functional categories of plants that have partly or fully lost the capacity for photosynthesis: 1) parasites that exploit other plants via direct connections and 2) myco-heterotrophs that indirectly exploit other plants via mycorrhizal fungi (Brundrett 2009). Both categories are referred to as parasites here. In total, there are about 13 evolutionary lineages of parasitic plants and at least 30 of myco-heterotrophs, almost all the latter within the orchid family (Brundrett 2009; Westwood et al. 2010). To date, the analysis of complete plastomes from parasitic plants has been restricted to dicots which are direct parasites of either roots (*Orobanchaceae*; dePamphilis and Palmer 1990) or shoots (*Cuscuta*; Funk et al. 2007) of other plants. No similar analysis has been published on the plastid genomes from nonphotosynthetic orchids that have a distinct evolutionary origin (the last common ancestor of orchids and dicots lived at least 120 Ma; Ramirez et al. 2007), with a very different growth habit and a different mode of parasitism.

For this study, we chose the Western Underground Orchid (*Rhizanthella gardneri*; fig. 1A and B) an iconic West Australian species. Remarkably, and unlike land plants from any other genus, the entire life cycle of *R. gardneri* occurs underground, with the flowers opening several centimeters below the soil surface. Its subterranean lifestyle ensures that it is incapable of photosynthesis; instead, it gains its energy and nutrients as a myco-heterotroph via *Ceratobasidium* fungi that form ectomycorrhizas with roots of broom bush, *Melaleuca uncinata* (fig. 1C; Bougoure et al. 2009, 2010). *R. gardneri* is critically endangered and known from only five locations in Western Australia (fig. 1D), but related

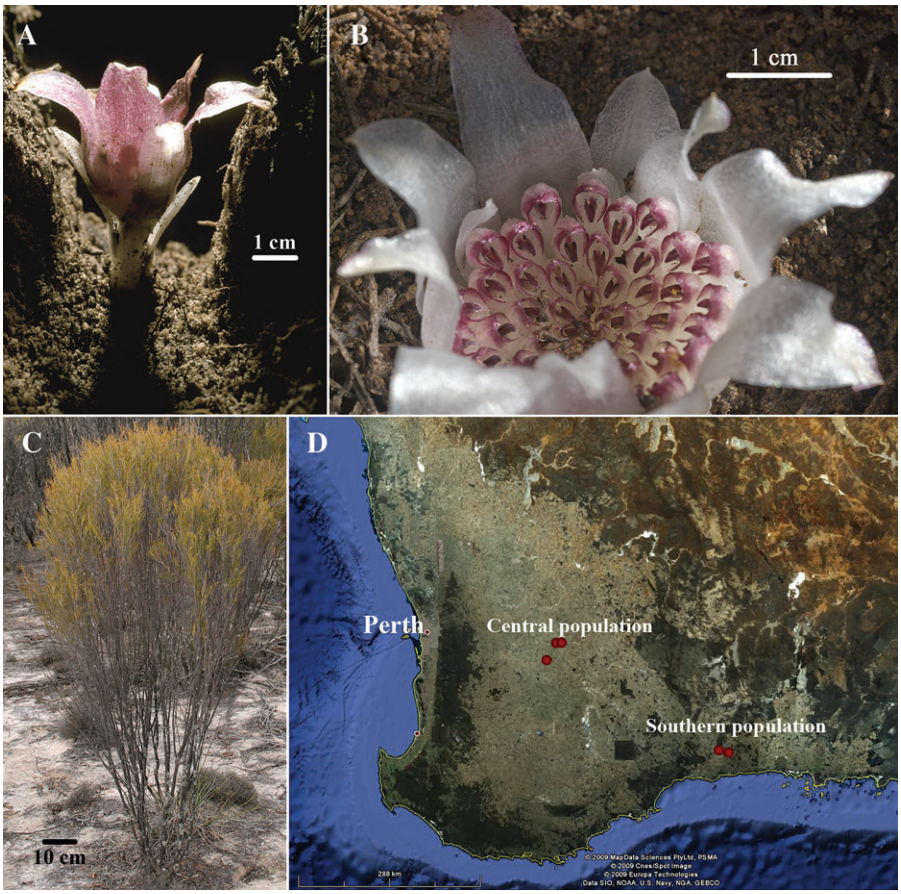

FIG. 1. Description of *Rhizanthella gardneri*. (*A*) Uncovered capitulum of *R. gardneri*. Picture courtesy of Susumu Yamaguchi. (*B*) Each capitulum encloses up to ~50 flowers. (*C*) Typical habitat of *R. gardneri*, showing the host plant, *Melaleuca uncinata*. (*D*) Approximate known sites for *R. gardneri* in south-western Australia. Map from Google Earth.

(and even rarer) species have been collected in Queensland and New South Wales, thousands of kilometers to the east, suggesting that these are ancient relictual species from a time before the center of Australia became arid, a process that started ~15 Ma (Byrne et al. 2008). Here, we report the full plastome sequence of *R. gardneri*. It is the smallest and most derived plastome described in land plants but shows strong evolutionary convergences with plastomes of other unrelated parasitic organisms.

Materials and Methods

DNA was extracted from a capitulum of *R. gardneri* collected in the Munglinup area on the south coast of Western Australia. Due to the status of this endangered species, precise locations cannot be published. It was purified as described (Lang and Burger 2007) except that plastid DNA could not be isolated from nuclear DNA by ultracentrifugation. Shotgun sequencing of total DNA was carried out on a Roche FLX system using the titanium kit (Roche Diagnostics, Castle Hill, NSW, Australia) at the Lotterywest State Biomedical Facility (Perth, WA, Australia). De novo assembly was performed with wgs (Miller et al. 2008), which yielded four large plastid DNA contigs. After alignment on the *Phalaenopsis aphrodite* subsp. *formosana* plastid genome (AY916449) and manual editing of the assembly using the Seqman module of Lasergene 8 (DNASTAR, Madison, WI), the assembly of the contigs, gaps, and low coverage areas were verified by polymerase chain reaction (PCR) using primer pairs described in supplementary table 1, Supplementary Material online, and PrimeSTAR DNA polymerase (Takara, Japan). tRNA detection was performed with Aragorn (Laslett and Canback 2004) and by systematic alignment of *P. aphrodite* subsp. *formosona* plastid tRNA exons.

Total RNA from an aliquot of the sample used for DNA extraction was purified with acid phenol bromochloropropane (1/1) and precipitated with ethanol. RNAs were precipitated with LiCl before treatment with DNAfree DNAse (Ambion, Austin, TX). One microgram of total RNA was reverse transcribed with random primers and Superscript III (Invitrogen, Mount Waverley, VIC, Australia). *rpl2*, *rpl16*, and *clpP* cDNAs were amplified with the primers pairs described in supplementary table 1, Supplementary Material online, and PrimeSTAR DNA polymerase (Takara, Japan).

To assess the variability of the plastid genome, DNA was extracted from a bract of *R. gardneri* collected in the Corrigin area (WA, Australia) and purified as described above. Fragments covering 32.9% of the plastid genome (accession numbers GU066242–GU066252) were sequenced using the

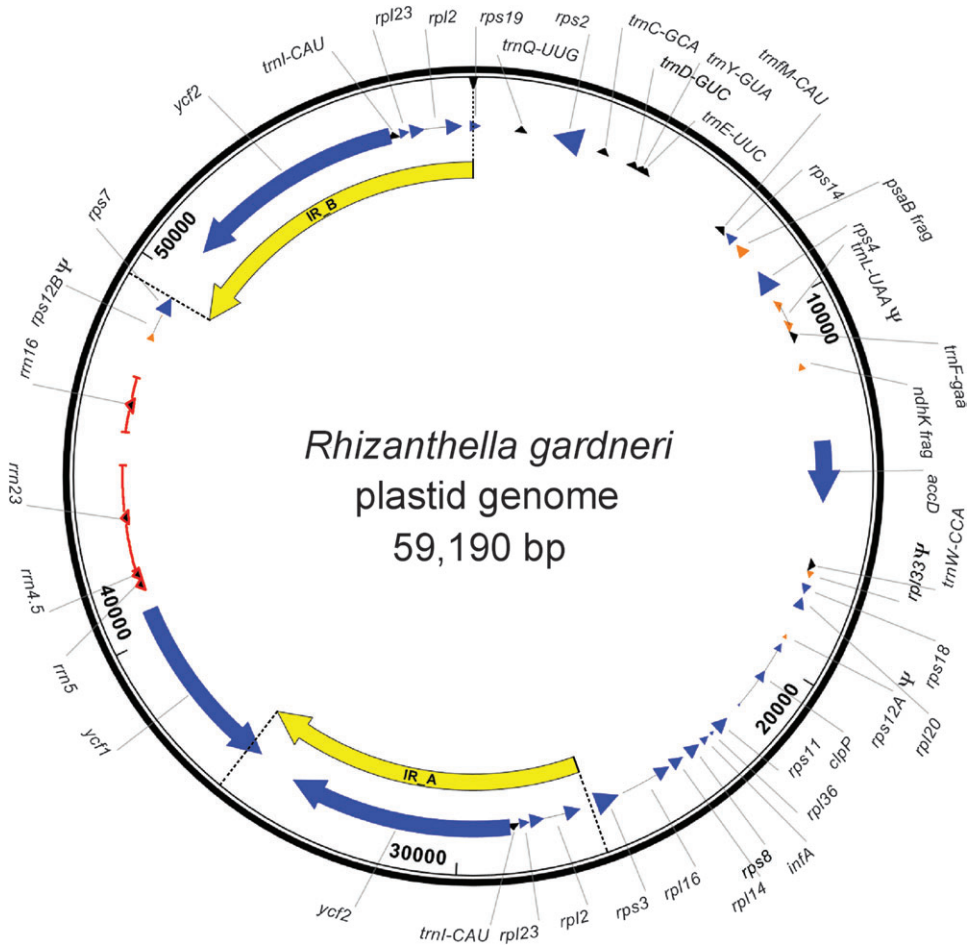

Fig. 2. *Rhizanthella gardneri* plastid genome. Exons are displayed as blue arrows, introns as lines joining exons, tRNAs as black triangles, rRNAs in red, IRs as yellow arrows, and fragments or pseudogenes in orange. Ψ: pseudogenes. frag: fragment.

primers pairs described in supplementary table 1, Supplementary Material online, and the PrimeSTAR DNA polymerase (Takara, Japan). *R. gardneri* mitochondrial sequences among the library of reads from the FLX system were identified by similarity to rice mitochondrial DNA. For each open reading frame (ORF), the matching reads were assembled with Seqman (Lasergene 8, DNASTAR, Madison, WI). Because of low coverage, only partial sequences could be recovered for most coding sequences (CDS) except *atp4*, *cox3*, and *nad9* genes. The corresponding sequences from samples collected in the Munglinup area (accession numbers GU066236, GU066239, and GU066241) and the Corrigin area (accession numbers GU066237, GU066238, and GU066240) were produced as described above using primer pairs described in supplementary table 1, Supplementary Material online. Sequence alignments were performed using MUSCLE (Edgar 2004) and manually edited. The gene lists of 172 complete plastid genomes were collected from National Center for Biotechnology Information and manually curated (supplementary table 2, Supplementary Material online). Duplicates within each genome were discarded as well as hypothetical ORFs without homologs. Using this matrix, the genomes were clustered by Bayesian inference using MrBayes 3.1 (Ronquist and Huelsenbeck 2003).

Results

Given the rarity of *R. gardneri*, biological material is extremely limited and insufficient for isolation of plastids. We estimated the proportion of plastid DNA in a total *R. gardneri* DNA preparation by quantitative PCR and calculated that it should be possible to achieve sufficient coverage of the plastid genome from a single sequencing run on total DNA. This prediction turned out to be well founded, and we assembled the 59,190 bp genome sequence (GenBank accession GQ413967) with an average of 15.7-fold coverage by shotgun deep sequencing of total DNA and targeted PCR amplifications to fill three short gaps and verify homopolymer runs.

The *R. gardneri* plastome (fig. 2) contains two inverted repeats (IR) of 9,767 bp, a large single-copy region of 26,361 bp, and a small single-copy region of 13,295 bp. This is the smallest known land plant organelle genome, being smaller than the plastid genome of the nonphotosynthetic dicot parasite *Epifagus virginiana* (Wolfe et al. 1992). The synteny with the plastome of the photosynthetic orchid *P. aphrodite* (Chang et al. 2006) is perfect (supplementary fig. 1, Supplementary Material online) despite rampant gene loss. Unlike any other plastome with large IRs (Strauss et al. 1988; Cai et al. 2008), the ribosomal RNA (rRNA) genes

Table 1. Global Features of Selected Plastomes.

	Size (kb)	% Noncoding DNA[a]	Coding Capacity[b]	Number of Pseudogenes	Number of tRNAs[c]	Accession Number
Angiosperms						
Rhizanthella gardneri	59.2	45.1	33	5	9	GQ413967
Phalaenopsis aphrodite	149	49.8	95	7	27	NC_007499
Epifagus virginiana	70	41.2	40	18	15	NC_001568
Cuscuta obtusiflora	85.3	29.2	85	1	23	NC_009949
C. exaltata	125.4	44.8	96	7	27	NC_009963
Nicotiana tabacum	155.9	53.5	110	1	28	NC_001879
Chlorophyta						
Helicosporidium sp.	37.5	5.1	53	0	24	NC_008100
Euglena longa	73.3	47.7	56	1	26	NC_002652
E. gracilis	143.2	57.4	86	7	26	NC_001603
Containing secondarily acquired red algal plastids						
Eimeria tenella	34.8	8.4	40	7	23	NC_004823
Theileria parva	39.6	38.3	43	0	21	NC_007758
Toxoplasma gondii	35	19.3	44	0	21	NC_001799
Chromera velia	119.8	31.4	84	5	25	NC_014340
Alveolata sp. CCMP3155	85.5	13.6	99	2	25	NC_014345
Porphyra purpurea	191	24.2	218	0	35	NC_000925
Peridin-containing dinoflagellates[d]						
Amphidinium operculatum	34.2	64.2	18	0	1	Barbrook, Santucci, et al. (2006)

NOTE.—*P. aphrodite*, *N. tabacum*, *E. gracilis*, *C. velia*, *Alveolata* sp., *P. purpurea*, and *A. operculatum* are photosynthetic organisms while others are parasitic or heterotrophic.

[a] Including structural RNAs but not introns. Only ORFs conserved in at least two plastid genomes were included.

[b] Number of genes including structural RNAs and ORFs conserved in at least two plastid genomes. Duplicated genes were counted as one.

[c] Duplicated tRNAs were counted as one.

[d] The plastome of these organisms is fragmented into several plasmids.

are not included within them (fig. 2). The *R. gardneri* plastome has a high proportion of noncoding DNA (table 1). This is not primarily due to the presence of pseudogenes as there are only two obvious pseudogenes (*rpl33* and *trnL*) and two gene fragments (*psaB* and *ndhK*) covering 1.9% of the genome. Although easily detected by sequence alignments, the three *rps12* exons predict an RPS12 protein lacking some highly conserved amino acids at the C-terminus and we failed to detect any transspliced *rps12* mRNA, leading us to consider *rps12* as a third pseudogene.

The *R. gardneri* plastome contains only 37 genes (including duplicates within the repeats) encoding 20 proteins, 4 rRNAs, and 9 transfer RNAs (table 2). In comparison, the *P. aphrodite* plastome contains 110 genes (Chang et al. 2006), the plastome of *E. virginiana* contains 53 (Wolfe et al. 1992) and even the highly reduced genomes of *Toxoplasma gondii* (35 kb; NC 001799) or the parasitic green alga *Helicosporidium* sp. (37.5 kb; de Koning and Keeling 2006) contain 65 and 54, respectively. With the exception of the very peculiar plastid genomes of the peridin-containing dinoflagellates that are fragmented into several plasmids (Barbrook, Santucci, et al. 2006; Howe et al. 2008), the *R. gardneri* plastid is the most gene-poor plastid genome characterized so far (table 1, supplementary table 2, Supplementary Material online). It contains fewer genes than any other characterized genetic system in land plants. In comparison with the plastome of *P. aphrodite*, which can be taken to resemble that of the photosynthetic ancestor of *R. gardneri*, an estimated 70% of the original genes were lost or transferred to the nucleus after the switch to a parasitic nonphotosynthetic lifestyle. These missing genes include those coding for the plastid-encoded RNA polymerase (PEP), the maturase-like protein MatK, all the genes required for photosynthesis (encoding subunits of photosystem I, photosystem II, cytochrome b_6f complex, and ATP synthase), as well as 6 genes encoding ribosomal proteins and 27 genes encoding tRNAs. Some of these missing genes may have been transferred to the nucleus. Although no unambiguous examples of this could be found in the sequencing data set, we cannot rule out the eventuality because of the low coverage of the nuclear genome.

The loss of the *rpo* genes encoding the major RNA polymerase, PEP, is associated with sequence divergence in some of the remaining promoters. Among the *R. gardneri* plastid genes, some are transcribed both from nucleus-encoded polymerase (NEP) promoters and PEP promoters in *Arabidopsis thaliana* (Swiatecka-Hagenbruch et al. 2007). The analysis of the corresponding upstream sequences in *R. gardneri* (supplementary fig. 2, Supplementary Material online) showed that the putative PEP promoters are not conserved in the sequences upstream of the *rps4* and *rrn16* genes but are apparently still present upstream of *ycf1*. NEP promoter motifs are conserved in the promoters of *ycf1*, *rps4*, and *clpP* but have diverged upstream of *rrn16*.

Sixteen of the *R. gardneri* genes encode proteins of the translation machinery (6 *rpl* genes, 9 *rps* genes, and an initiation factor); the other four protein-encoding genes are *accD*, *ycf1*, *ycf2*, and *clpP*, all easily identified by homology to plastid genes from other plants. Despite the reduced tRNA set, there is no significant change in codon usage compared with *P. aphrodite* (supplementary table 3, Supplementary Material online).

The clustering of 172 complete plastid genomes based on their gene content (fig. 3) shows that plastomes

Table 2. Gene Contents of Selected Plastomes.

	Photosynthesis					Protein Synthesis					Transfer RNA Genes			
	Pa	*Rg*	*Cg*	*Ev*		*Pa*	*Rg*	*Cg*	*Ev*		*Pa*	*Rg*	*Cg*	*Ev*
atpA	+	−	+	Ψ	*infA*	+	+	−	+	*trnA-ugc*	+	−	Ψ	Ψ
atpB	+	−	+	Ψ	*rpl2*	+	+	+	+	*trnC-gca*	+	+	+	Ψ
atpE	+	−	+	−	*rpl14*	+	+	+	Ψ	*trnD-guc*	+	+	+	+
atpF	+	−	+	−	*rpl16*	+	+	+	+	*trnE-uuc*	+	+	+	+
atpH	+	−	+	−	*rpl20*	+	+	+	+	*trnF-gaa*	+	+	+	+
atpI	+	−	+	−	*rpl22*	+	−	+	−	*trnfM-cau*	+	+	+	+
ccsA	+	−	+	−	*rpl23*	+	+	−	Ψ	*trnG-gcc*	+	−	+	−
cemA	+	−	+	−	*rpl32*	+	−	−	−	*trnG-ucc*	+	−	Ψ	−
petA	+	−	+	−	*rpl33*	+	Ψ	+	+	*trnH-gug*	+	−	+	+
petB	+	−	+	−	*rpl36*	+	+	+	+	*trnI-cau*	+	+	+	+
petD	+	−	+	−	*rps2*	+	+	+	+	*trnI-gau*	+	−	Ψ	Ψ
petG	+	−	+	−	*rps3*	+	+	+	+	*trnK-uuu*	+	−	−	−
petL	+	−	+	−	*rps4*	+	+	+	+	*trnL-caa*	+	−	+	+
petN	+	−	+	−	*rps7*	+	+	+	+	*trnL-uaa*	+	Ψ	+	−
psaA	+	−	+	−	*rps8*	+	+	+	+	*trnL-uag*	+	−	+	+
psaB	+	Ψ	+	−	*rps11*	+	+	+	+	*trnM-cau*	+	−	+	+
psaC	+	−	+	−	*rps12*	+	Ψ	+	+	*trnN-guu*	+	−	+	+
psaI	+	−	−	−	*rps14*	+	+	−	+	*trnP-ugg*	+	−	+	+
psaJ	+	−	+	−	*rps15*	+	−	+	−	*trnQ-uug*	+	+	+	+
psbA	+	−	+	Ψ	*rps16*	+	−	−	−	*trnR-acg*	+	−	Ψ	+
psbB	+	−	+	Ψ	*rps18*	+	+	+	+	*trnR-ucu*	+	−	+	Ψ
psbC	+	−	+	−	*rps19*	+	+	+	+	*trnS-gcu*	+	−	+	+
psbD	+	−	+	−	**RNA Metabolism**					*trnS-gga*	+	−	+	Ψ
psbE	+	−	+	−		*Pa*	*Rg*	*Cg*	*Ev*	*trnS-uga*	+	−	+	+
psbF	+	−	+	−	*matK*	+	−	−	+	*trnT-ggu*	+	−	+	−
psbH	+	−	+	−	*rpoA*	+	−	−	Ψ	*trnT-ugu*	+	−	+	−
psbI	+	−	+	−	*rpoB*	+	−	−	−	*trnV-gac*	+	−	+	−
psbJ	+	−	+	−	*rpoC1*	+	−	−	−	*trnV-uac*	+	−	−	−
psbK	+	−	+	−	*rpoC2*	+	−	−	−	*trnW-cca*	+	+	+	+
psbL	+	−	+	−	**Essential Genes**					*trnY-gua*	+	+	+	+
psbM	+	−	+	−		*Pa*	*Rg*	*Cg*	*Ev*	**Ribosomal RNA Genes**				
psbN	+	−	+	−	*clpP*	+	+	+	+		*Pa*	*Rg*	*Cg*	*Ev*
psbT	+	−	+	−	*accD*	+	+	+	+	*rrn16*	+	+	+	+
rbcL	+	−	+	Ψ	*ycf1*	+	+	+	+	*rrn23*	+	+	+	+
psbZ	+	−	+	−	*ycf2*	+	+	+	+	*rrn4.5*	+	+	+	+
ycf3	+	−	+	−						*rrn5*	+	+	+	+
ycf4	+	−	+	−										

NOTE.—*Pa*, *Phalaenopsis aphrodite*; *Rg*, *Rhizanthella gardneri*; *Cg*, *Cuscuta gronovii*; *Ev*, *Epifagus virginiana*; Ψ, pseudogene; +, present; −, missing. Genes present in all four plastomes are indicated in bold.

deriving from green or red algae form two distinct groups, apart from some heterotrophs. These latter species derive either from green algae (*Euglena longa*, *Helicosporidium* sp., *E. virginiana*, and *R. gardneri*) or red algae (*Eimeria tenella*, *T. gondii*, and *Theileria parva*; Janouskovec et al. 2010) but cluster together. Some parasitic plants or algae (*Cuscuta* species, *Aneura mirabilis* and *Cryptomonas paramecium*) are not part of this cluster. However, these have either not completely switched to a nonphotosynthetic lifestyle (Hibberd et al. 1998) or did so only recently (Wickett et al. 2008; Donaher et al. 2009).

Comparison of DNA sequences from the central and southern populations of *R. gardneri* showed that the plastome is accumulating mutations at a very high rate, at least 15-fold faster than the mitochondrial genome (supplementary table 4, Supplementary Material online). This result is in line with the increased rates of fixation of plastid DNA mutations described in other parasitic plants (Young and dePamphilis 2005). The sequence divergence observed between the two populations of *R. gardneri* are similar to those observed between different species of *Cuscuta* (supplementary table 5, Supplementary Material online). The relative rate of divergence of the *R. gardneri rrn23* gene is higher than that of the *rps8* and *rpl36* genes when compared with the equivalent rates in the *Cuscuta* species (supplementary table 5, Supplementary Material online).

Despite this high rate of sequence divergence in the *R. gardneri* plastome, preferential conservation of the coding sequences (supplementary table 4, Supplementary Material online) suggests that the genes are expressed and functional. We cloned *rpl2*, *rpl16*, and *clpP* cDNAs (accession numbers GU066223, GU066224, and GU066222,

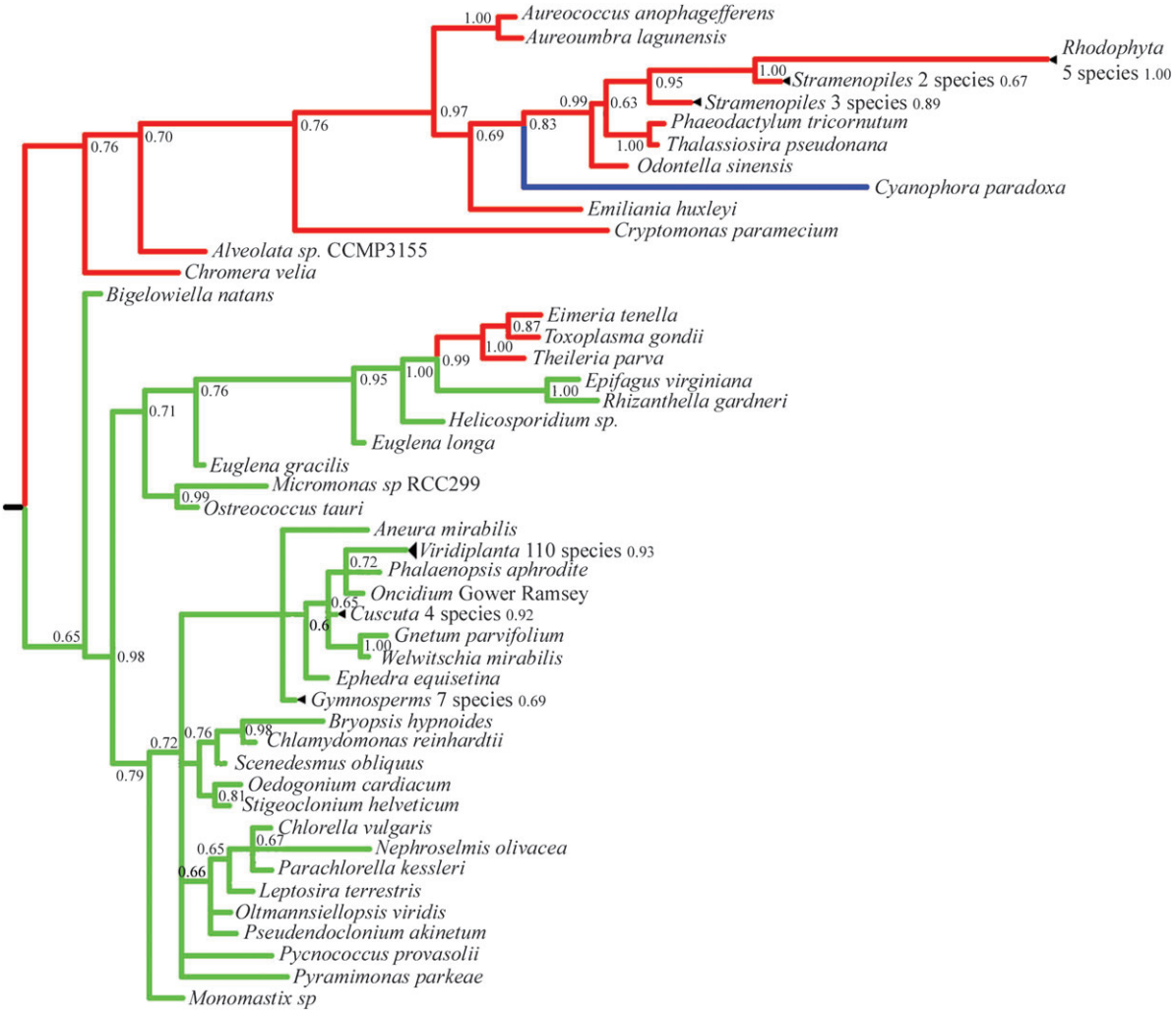

Fig. 3. Convergence of plastid genomes from parasites. One-hundred and seventy-two plastid genomes were clustered by Bayesian inference using the data in supplementary table 1, Supplementary Material online. Black triangles indicate collapsed clusters, with the number and taxonomic group of the species shown. Probabilities above 0.6 are shown at the corresponding nodes. Red, green, and blue branches indicate plastids derived from the red, green, and blue lineage, respectively. Supplementary data 1, Supplementary Material online, is the full tree in NEXUS format.

respectively). These three cDNAs were correctly spliced showing that RNA splicing is occurring in *R. gardneri* plastids. Within these mRNAs, we detected two C–U editing events: at the start codon of *rpl2* and in the second exon of *rpl16* (supplementary fig. 3, Supplementary Material online). Hence, *R. gardneri* has normal plastid RNA metabolism with transcription, splicing, and editing occurring.

Discussion

The *R. gardneri* plastid contains a typical quadripartite genome with reduced IRs. This reduction has resulted in the presence of a single copy of the *rrn* genes as opposed to duplicates or even triplicates in all other plastomes containing large repeats (Strauss et al. 1988; Cai et al. 2008). Genes in the IRs generally display lower substitution rates compared with genes present in the single-copy regions (Wolfe et al. 1987; Perry and Wolfe 2002; Raubeson et al. 2007). The relative divergence rates within the *R. gardneri* plastid genome (compared with parasitic *Cuscuta* species) is higher for the *rrn23* gene (duplicated in *Cuscuta*) than for the *rps8* and *rpl36* genes (single copy in all plastid genomes). This acceleration of mutation rates in the *rrn* genes in *R. gardneri* probably results from the loss of the duplicate copies as shown in legumes (Perry and Wolfe 2002).

The reduction of the IRs is the major reason that explains why the *R. gardneri* plastome is smaller than that of the smallest previously known land plant plastid genome and that of the parasitic dicot *E. virginiana*. Despite the very small gene set encoded in this genome, it is not the smallest in size. Plastid genomes from apicomplexan parasites such as *E. tenella* or the parasitic green alga *Helicosporidium* sp. are much smaller in size, although they contain more genes (table 1).

The *R. gardneri* plastome contains one of the smallest gene sets characterized in any plastome analyzed so far. Only peridin-containing dinoflagellates possess plastids with less genes (Barbrook, Santucci, et al. 2006; Howe et al. 2008). Despite rampant gene loss, the *R. gardneri* plastome appears to be the basis of a functioning gene

Table 3. Plant Mitochondria and Plastids Tend to Lose and Retain the Same tRNA Genes.

		Mitochondria		
		Always Conserved	**Sometimes Replaced**	**Sometimes Lost**
Plastids	**Always conserved**	**E-uuc; Q-uug; I-cau; fM-cau; Y-gua**	**D-guc**	**F-gaa; W-cca**
	Sometimes lost		**eM-cau; N-guu; C-gca**	**S-gga; A-ugc; R-acg; G-gcc; R-ucu; G-ucc; T-ggu; I-gau; T-ugu; K-uuu; V-gac; L-uaa; V-uac; H-gug; L-uag; P-ugg; S-uga; S-gcu; L-caa**

NOTE.—Modified from (Lohan and Wolfe 1998). tRNA genes are indicated by single letter amino acid codes and their anticodon sequence. "Sometimes replaced" indicates mitochondrial tRNA genes that have been replaced by orthologous plastid genes in horizontal gene transfer events.

expression system, with transcription, splicing, and RNA editing all detected and translation likely. Although editing requires only nuclear-encoded factors (Schmitz-Linneweber and Small 2008), transcription in land plant plastids usually relies on two distinct transcription machineries: One encoded by the nucleus (NEP) and the other encoded in part by the plastid. The *rpo* genes coding for the PEP are missing in *R. gardneri*. In photosynthetic plastids, each polymerase transcribes a distinct but overlapping set of genes (Hajdukiewicz et al. 1997), with PEP preferentially transcribing photosynthesis-related genes. As no gene typically relying exclusively on PEP for its transcription remains in the *Rhizanthella* plastid genome, the loss of *rpo* genes in *R. gardneri* can be understood. The loss of PEP probably explains the divergence noted in the promoter regions of several of the remaining genes (supplementary fig. 2, Supplementary Material online).

Splicing of introns in chloroplasts requires several nuclearly encoded factors that are specific to one or more introns (Asakura and Barkan 2006; de Longevialle et al. 2008). The chloroplast genome also encodes a maturase-like protein, MatK, which is involved in splicing of the *trnK* intron in which it is embedded (Vogel et al. 1997) but is probably also required for the splicing of the group IIa subset of plastid introns (Jenkins et al. 1997; Vogel et al. 1999; Duffy et al. 2009; McNeal et al. 2009). The loss of *matK* has been observed in *Cuscuta* species from the subgenus *Grammica*, which have also lost the introns thought to require MatK activity. The loss of *matK* in *R. gardneri* is more surprising because the *R. gardneri* plastid genome has retained three group IIa introns, two of which we show to be correctly spliced and which, in tobacco, are bound by MatK (Zoschke et al. 2010). A large proportion of *matK* genes in orchids are pseudogenes because of unequal insertions/deletions (Kores et al. 2000, 2001) suggesting that the role of *matK* for splicing in this family of plants is not as essential as in other families. This may explain why *matK* has been lost in *R. gardneri* without impairing splicing of the remaining group IIa introns.

Plastid translation is generally required for cell viability from embryogenesis onward in land plants (Berg et al. 2005), and we presume that this is the case in *R. gardneri*, despite the loss of many genes normally essential for translation to occur. The missing proteins are most probably imported along with other elements of the translation machinery already imported into plastids in other plants (Stengel et al. 2007). Most striking of all, the *R. gardneri* plastome is characterized by the loss of a large proportion of tRNA genes: Only 10 genes coding for 9 different tRNAs are present (table 2). This is by far the smallest tRNA set in an unfragmented plastid genome (table 1, supplementary table 2, Supplementary Material online) and manifestly insufficient for translation. Despite this reduced tRNA set, there is no significant change in codon usage compared with *P. aphrodite* (supplementary table 3, Supplementary Material online).

Import of cytosolic tRNAs into plastids has never been directly demonstrated and several lines of evidence suggest that it does not occur in photosynthetic land plants (Lung et al. 2006; Rogalski et al. 2008), but we presume that the missing tRNAs are indeed imported, as also suggested for *E. virginiana* (Wolfe et al. 1992). Cytosolic tRNAs are imported into mitochondria of land plants and there are intriguing parallels between the tRNA sets encoded in both organelles (Lohan and Wolfe 1998): 25 of 30 tRNAs are either sometimes imported into mitochondria and not conserved in parasite plastomes or never imported into mitochondria and conserved in plastomes (table 3). This implies that the same constraints apply for the import of tRNAs into both mitochondria and plastids. These constraints can be guessed at for all five tRNAs retained in all plant organelle genomes.

Aminoacylated $tRNA^{Glu}$(UUC) is a precursor not only for protein synthesis but also for tetrapyrrole synthesis (Tanaka R and Tanaka A 2007). In photosynthetic plants, it is required for chlorophyll synthesis, and in all plants for the synthesis of heme for mitochondrial respiratory complexes and other essential proteins. Imported cytosolic $tRNA^{Glu}$ would therefore need to be recognized by not only glutamyl-tRNA synthetase for aminocylation but also glutamyl-tRNA reductase for tetrapyrrole synthesis. The *trnE* gene may be the primary justification for the presence of a genome in all nonphotosynthetic plastids (Barbrook, Howe, et al. 2006). Similarly, $tRNA^{fMet}$ is required for initiating translation in prokaryotic systems, including virtually all organelles; imported $tRNA^{Met}$ would need to be recognized by the methionyl-tRNA transformylase or translation initiation factors in addition to being aminoacylated by methionine-tRNA synthetase. For both these tRNAs, the requirement for an imported replacement to be recognized by several distinct enzymes for which they are not normally substrates makes it extremely unlikely for functional replacement to occur.

The other three ubiquitous plant organellar tRNAs share unusual features about their aminoacylation that make them different from all other tRNAs. Organellar $tRNA^{Ile}$ (CAU) has a typical $tRNA^{Met}$ anticodon, but this is modified by addition of the amino acid lysidine to C_{34}; the organellar isoleucyl-tRNA synthetase has evolved to recognize this unusual modified base which is not found in cytosolic tRNAs. Organellar $tRNA^{Tyr}$ differs extensively from its cytosolic counterpart in that it has a very long variable loop. For both these tRNAs, the imported equivalent would fail to be recognized by the corresponding organellar aminoacyl-tRNA synthetase. Plant organellar $tRNA^{Gln}$-Gln is formed by amidation of $tRNA^{Gln}$-Glu; there is no glutaminyl-tRNA synthetase.

Thus for $tRNA^{Ile}$(CAU), $tRNA^{Tyr}$, and $tRNA^{Gln}$, import of the cytosolic tRNA alone would be unlikely to lead to functional aminoacyl-tRNA for translation. Only coupled import of a suitable aminoacyl-tRNA synthetase could establish functional replacement of the organellar genes. Incidentally, the convergences noted here between retention of plastid and mitochondrial tRNA genes argues that unlike some early suggestions, coupled transport of tRNAs and aminoacyl-tRNA synthetases is probably not the mechanism by which organellar tRNA import occurs.

Apart from genes involved in the translation machinery (rRNAs, ribosomal proteins, and tRNAs), *R. gardneri* has retained a very restricted set of other protein-coding genes, namely *ycf1*, *ycf2*, *accD*, and *clpP* (table 2). *accD*, *ycf1*, and *ycf2* are conserved in almost all land plants (supplementary table 2, Supplementary Material online) but have been lost in grasses (Katayama and Ogihara 1996) and *accD* is missing in a few other lineages (Knox and Palmer 1999; Chumley et al. 2006). The *accD* gene encodes the carboxyltransferase subunit of a multimeric acetyl-CoA carboxylase (ACCase), which provides malonyl-CoA for the biosynthesis of fatty acids. These fatty acids are then used for the synthesis of every cellular membrane (Benning et al. 2006). This fundamental function may explain why mutations in *accD* (Kode et al. 2005) or in the plastid translation machinery (Berg et al. 2005; Rogalski et al. 2008) are lethal. In the few plants where this is not the case, an imported monomeric ACCase replaces the requirement for the *accD* gene product (Konishi et al. 1996). The essential functions of *ycf1* and *ycf2* are unknown (Drescher et al. 2000) but might possibly be linked to expression, assembly, or function of the *accD* gene product, given that grasses have lost both genes in addition to *accD*. Other plants that have lost *accD* have divergent *ycf1* and *ycf2* sequences.

Finally, *clpP* is the only protein-coding gene present in all land plant and green algal plastomes except for that of the parasitic alga *Helicosporidium* sp. (supplementary table 2, Supplementary Material online). It codes for a catalytic subunit of a multimeric protease. In land plants, *clpP* is essential (Kuroda and Maliga 2003) but which of its varied roles are required remains unclear. Recent studies suggest that a function in the regulation of an array of processes including isoprenoid and tetrapyrrole biosynthesis, lipid body stability, and photosynthesis (Kim et al. 2009; Stanne et al. 2009; Zybailov et al. 2009).

The *R. gardneri* plastome is also informative on the selective pressures acting to retain certain genes within organelle genomes. The small number of retained genes and relative lack of pseudogenes (5 vs. 18 in *E. virginiana*) suggests that the *R. gardneri* plastome has progressed further toward shedding or transferring nonessential genes than any other land plant organelle examined so far. The loss of photosynthetic capacity is associated with a strong reduction of the plastid coding capacity, in line with the evolutionary trend of genome reduction in plastids (Martin et al. 1998). The remaining gene set is clearly not random. The analysis of the plastid gene content in 172 organisms (fig. 3, supplementary table 2, Supplementary Material online) showed the remarkable similarity of the apicomplexans and parasitic plants. We believe this similarity primarily reflects convergent rather than shared evolution given the supposed evolutionary distance between these plastid genomes. We note, however, that although several lines of evidence strongly support a red algal origin for apicoplasts (Wilson et al. 1996; Yoon et al. 2002; Janouskovec et al. 2010), it is still a matter of controversy (Kohler et al. 1997; Lau et al. 2009). It has been noted previously (Stiller et al. 2003) that, when considering ribosomal proteins and tRNAs, similarities in plastid gene content reflect convergent evolution rather than shared descent. The similarity of the plastid gene contents of nonphotosynthetic organisms suggests that they tend to converge toward a distinct shared gene set given enough time. The nature of this set gives clues about the evolutionary constraints on plastid gene losses.

What are these constraints? The most popular previously suggested explanations (Daley and Whelan 2005; Barbrook, Howe, et al. 2006) revolve around constraints on protein import due to high hydrophobicity (de Grey 2005) or essential regulation of gene expression by reactive oxygen signaling cascades (Allen et al. 2005; Howe et al. 2008). However, they cannot easily account for the retention of *accD* and *clpP*, which encode typical soluble globular proteins that are not involved in respiration or photosynthesis. An alternative hypothesis suggests that certain genes encoding key products required for controlling the assembly of multiprotein complexes cannot be easily lost or transferred (Zerges 2002). This could be explained by "control by epistasy of synthesis" (Wostrikoff et al. 2004) which is an elegant theory on the regulation of assembly of protein complexes that requires at least one subunit to be organelle encoded. The *accD* and *clpP* gene products are both essential components of large protein complexes that must be assembled within the plastid. We postulate that herein lies the explanation for the very small retained gene set of the *R. gardneri* plastome.

Supplementary Material

Supplementary tables 1–5, figures 1–3, and supplementary data 1 are available at *Molecular Biology and Evolution* online (http://www.mbe.oxfordjournals.org/).

Acknowledgments

This work was supported by the Australian Research Council Centre of Excellence in Plant Energy Biology (grant CE0561495) and a grant from the Faculty of Natural Science and Agriculture of the University of Western Australia. M.B. is supported by Lotterywest funding for the Wheatbelt Orchid Rescue Project. I.S. was supported by the West Australian Premier's Fellowship scheme. The authors would like to thank the West Australian Native Orchid Study and Conservation Group without whose help this project would not have been possible.

References

Ahlert D, Ruf S, Bock R. 2003. Plastid protein synthesis is required for plant development in tobacco. *Proc Natl Acad Sci U S A*. 100:15730–15735.

Allen JF, Puthiyaveetil S, Strom J, Allen CA. 2005. Energy transduction anchors genes in organelles. *Bioessays* 27:426–435.

Archibald JM. 2009. The puzzle of plastid evolution. *Curr Biol*. 19:R81–R88.

Asakura Y, Barkan A. 2006. Arabidopsis orthologs of maize chloroplast splicing factors promote splicing of orthologous and species-specific group II introns. *Plant Physiol*. 142:1656–1663.

Barbrook AC, Howe CJ, Purton S. 2006. Why are plastid genomes retained in non-photosynthetic organisms? *Trends Plant Sci*. 11:101–108.

Barbrook AC, Santucci N, Plenderleith LJ, Hiller RG, Howe CJ. 2006. Comparative analysis of dinoflagellate chloroplast genomes reveals rRNA and tRNA genes. *BMC Genomics*. 7:297.

Benning C, Xu C, Awai K. 2006. Non-vesicular and vesicular lipid trafficking involving plastids. *Curr Opin Plant Biol*. 9:241–247.

Berg M, Rogers R, Muralla R, Meinke D. 2005. Requirement of aminoacyl-tRNA synthetases for gametogenesis and embryo development in Arabidopsis. *Plant J*. 44:866–878.

Bougoure J, Brundrett M, Grierson PG. 2010. Carbon and nitrogen supply to the underground orchid. *New Phytol*. 186:947–956.

Bougoure J, Ludwig M, Brundrett M, Grierson P. 2009. Identity and specificity of the fungi forming mycorrhizas with the rare mycoheterotrophic orchid *Rhizanthella gardneri*. *Mycol Res*. 113:1097–1106.

Brundrett M. 2009. Mycorrhizal associations and other means of nutrition of vascular plants: understanding the global diversity of host plants by resolving conflicting information and developing reliable means of diagnosis. *Plant Soil*. 320:37–77.

Byrne M, Yeates DK, Joseph L, et al. (14 co-authors). 2008. Birth of a biome: insights into the assembly and maintenance of the Australian arid zone biota. *Mol Ecol*. 17:4398–4417.

Cai Z, Guisinger M, Kim HG, Ruck E, Blazier JC, McMurtry V, Kuehl JV, Boore J, Jansen RK. 2008. Extensive reorganization of the plastid genome of *Trifolium subterraneum* (*Fabaceae*) is associated with numerous repeated sequences and novel DNA insertions. *J Mol Evol*. 67:696–704.

Chang CC, Lin HC, Lin IP, et al. (11 co-authors). 2006. The chloroplast genome of *Phalaenopsis aphrodite* (*Orchidaceae*): comparative analysis of evolutionary rate with that of grasses and its phylogenetic implications. *Mol Biol Evol*. 23:279–291.

Chumley TW, Palmer JD, Mower JP, Fourcade HM, Calie PJ, Boore JL, Jansen RK. 2006. The complete chloroplast genome sequence of *Pelargonium* x *hortorum*: organization and evolution of the largest and most highly rearranged chloroplast genome of land plants. *Mol Biol Evol*. 23:2175–2190.

Daley DO, Whelan J. 2005. Why genes persist in organelle genomes. *Genome Biol*. 6:110.

de Grey AD. 2005. Forces maintaining organellar genomes: is any as strong as genetic code disparity or hydrophobicity? *Bioessays* 27:436–446.

de Koning AP, Keeling PJ. 2006. The complete plastid genome sequence of the parasitic green alga Helicosporidium sp. is highly reduced and structured. *BMC Biol*. 4:12.

de Longevialle AF, Hendrickson L, Taylor NL, Delannoy E, Lurin C, Badger M, Millar AH, Small I. 2008. The pentatricopeptide repeat gene OTP51 with two LAGLIDADG motifs is required for the cis-splicing of plastid ycf3 intron 2 in *Arabidopsis thaliana*. *Plant J*. 56:157–168.

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

Donaher N, Tanifuji G, Onodera NT, Malfatti SA, Chain PS, Hara Y, Archibald JM. 2009. The complete plastid genome sequence of the secondarily nonphotosynthetic alga *Cryptomonas paramecium*: reduction, compaction, and accelerated evolutionary rate. *Genome Biol Evol*. 1:439–448.

Drescher A, Ruf S, Calsa T Jr., Carrer H, Bock R. 2000. The two largest chloroplast genome-encoded open reading frames of higher plants are essential genes. *Plant J*. 22:97–104.

Duffy AM, Kelchner SA, Wolf PG. 2009. Conservation of selection on *matK* following an ancient loss of its flanking intron. *Gene* 438:17–25.

Edgar RC. 2004. MUSCLE: a multiple sequence alignment method with reduced time and space complexity. *BMC Bioinformatics*. 32:1792–1797.

Funk HT, Berg S, Krupinska K, Maier UG, Krause K. 2007. Complete DNA sequences of the plastid genomes of two parasitic flowering plant species, *Cuscuta reflexa* and *Cuscuta gronovii*. *BMC Plant Biol*. 7:45.

Hajdukiewicz PT, Allison LA, Maliga P. 1997. The two RNA polymerases encoded by the nuclear and the plastid compartments transcribe distinct groups of genes in tobacco plastids. *EMBO J*. 16:4041–4048.

Hibberd JM, Bungard RA, Press MC, Jeschke WD, Scholes JD, Quick WP. 1998. Localization of photosynthetic metabolism in the parasitic angiosperm. *Planta* 205:506–513.

Howe CJ, Nisbet RE, Barbrook AC. 2008. The remarkable chloroplast genome of dinoflagellates. *J Exp Bot*. 59:1035–1045.

Janouskovec J, Horak A, Obornik M, Lukes J, Keeling PJ. 2010. A common red algal origin of the apicomplexan, dinoflagellate, and heterokont plastids. *Proc Natl Acad Sci U S A*. 107:10949–10954.

Jenkins BD, Kulhanek DJ, Barkan A. 1997. Nuclear mutations that block group II RNA splicing in maize chloroplasts reveal several intron classes with distinct requirements for splicing factors. *Plant Cell*. 9:283–296.

Katayama H, Ogihara Y. 1996. Phylogenetic affinities of the grasses to other monocots as revealed by molecular analysis of chloroplast DNA. *Curr Genet*. 29:572–581.

Kim J, Rudella A, Ramirez Rodriguez V, Zybailov B, Olinares PD, van Wijk KJ. 2009. Subunits of the plastid ClpPR protease complex have differential contributions to embryogenesis, plastid biogenesis, and plant development in Arabidopsis. *Plant Cell*. 21:1669–1692.

Knox EB, Palmer JD. 1999. The chloroplast genome arrangement of *Lobelia thuliniana* (*Lobeliaceae*): expansion of the inverted repeat in an ancestor of the *Campanulales*. Pl. *Syst Evol*. 214:49–64.

Kode V, Mudd EA, Iamtham S, Day A. 2005. The tobacco plastid *accD* gene is essential and is required for leaf development. *Plant J*. 44:237–244.

Kohler S, Delwiche CF, Denny PW, Tilney LG, Webster P, Wilson RJ, Palmer JD, Roos DS. 1997. A plastid of probable green algal origin in Apicomplexan parasites. *Science*. 275:1485–1489.

Konishi T, Shinohara K, Yamada K, Sasaki Y. 1996. Acetyl-CoA carboxylase in higher plants: most plants other than gramineae have both the prokaryotic and the eukaryotic forms of this enzyme. *Plant Cell Physiol*. 37:117–122.

Kores PJ, Molvray M, Weston PW, Hopper SD, Brown AP, Cameron KM, Chase MW. 2001. A phylogenetic analysis of *Diurideae* (*Orchidaceae*) based on plastid DNA sequence data. *Am J Bot*. 88:1903–1914.

Kores PJ, Weston PW, Molvray M, Chase MW. 2000. Phylogenetic relationships within *Diurideae*: inferences from plastid *matK* DNA sequences. In: Wilson KL, Morrison DA, editors. Monocots: systematics and evolution. Victoria (Australia): CSIRO Publishing.

Kuroda H, Maliga P. 2003. The plastid clpP1 protease gene is essential for plant development. *Nature* 425:86–89.

Lang BF, Burger G. 2007. Purification of mitochondrial and plastid DNA *Nat Protoc*. 2:652–660.

Laslett D, Canback B. 2004. ARAGORN, a program to detect tRNA genes and tmRNA genes in nucleotide sequences. *Nucleic Acids Res*. 32:11–16.

Lau AO, McElwain TF, Brayton KA, Knowles DP, Roalson EH. 2009. Babesia bovis: a comprehensive phylogenetic analysis of plastid-encoded genes supports green algal origin of apicoplasts. *Exp Parasitol*. 123:236–243.

Lohan AJ, Wolfe KH. 1998. A subset of conserved tRNA genes in plastid DNA of nongreen plants. *Genetics* 150:425–433.

Lung B, Zemann A, Madej MJ, Schuelke M, Techritz S, Ruf S, Bock R, Huttenhofer A. 2006. Identification of small non-coding RNAs from mitochondria and chloroplasts. *Nucleic Acids Res*. 34:3842–3852.

Martin W, Stoebe B, Goremykin V, Hapsmann S, Hasegawa M, Kowallik KV. 1998. Gene transfer to the nucleus and the evolution of chloroplasts. *Nature* 393:162–165.

McNeal JR, Kuehl JV, Boore JL, Leebens-Mack J, dePamphilis CW. 2009. Parallel loss of plastid introns and their maturase in the genus Cuscuta. *PLoS One*. 4:e5982.

Miller JR, Delcher AL, Koren S, Venter E, Walenz BP, Brownley A, Johnson J, Li K, Mobarry C, Sutton G. 2008. Aggressive assembly of pyrosequencing reads with mates. *Bioinformatics* 24:2818–2824.

Perry AS, Wolfe KH. 2002. Nucleotide substitution rates in legume chloroplast DNA depend on the presence of the inverted repeat. *J Mol Evol*. 55:501–508.

Ramirez SR, Gravendeel B, Singer RB, Marshall CR, Pierce NE. 2007. Dating the origin of the *Orchidaceae* from a fossil orchid with its pollinator. *Nature* 448:1042–1045.

Raubeson LA, Peery R, Chumley TW, Dziubek C, Fourcade HM, Boore JL, Jansen RK. 2007. Comparative chloroplast genomics: analyses including new sequences from the angiosperms Nuphar advena and Ranunculus macranthus. *BMC Genomics*. 8:174.

Rogalski M, Karcher D, Bock R. 2008. Superwobbling facilitates translation with reduced tRNA sets. *Nat Struct Mol Biol*. 15:192–198.

Ronquist F, Huelsenbeck JP. 2003. MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* 19:1572–1574.

Schmitz-Linneweber C, Small I. 2008. Pentatricopeptide repeat proteins: a socket set for organelle gene expression. *Trends Plant Sci*. 13:663–670.

Stanne TM, Sjogren LL, Koussevitzky S, Clarke AK. 2009. Identification of new protein substrates for the chloroplast ATP-dependent Clp protease supports its constitutive role in Arabidopsis. *Biochem J*. 417:257–268.

Stengel A, Soll J, Bolter B. 2007. Protein import into chloroplasts: new aspects of a well-known topic. *Biol Chem*. 388:765–772.

Stiller JW, Reel DC, Johnson JC. 2003. A single origin of plastids revisited: convergent evolution in organellar genome content. *J Phycol*. 39:95–105.

Strauss SH, Palmer JD, Howe GT, Doerksen AH. 1988. Chloroplast genomes of two conifers lack a large inverted repeat and are extensively rearranged. *Proc Natl Acad Sci U S A*. 85: 3898–3902.

Swiatecka-Hagenbruch M, Liere K, Borner T. 2007. High diversity of plastidial promoters in Arabidopsis thaliana. *Mol Genet Genomics*. 277:725–734.

Tanaka R, Tanaka A. 2007. Tetrapyrrole biosynthesis in higher plants. *Annu Rev Plant Biol*. 58:321–346.

Vogel J, Borner T, Hess WR. 1999. Comparative analysis of splicing of the complete set of chloroplast group II introns in three higher plant mutants. *Nucleic Acids Res*. 27:3866–3874.

Vogel J, Hubschmann T, Borner T, Hess WR. 1997. Splicing and intron-internal RNA editing of *trnK-matK* transcripts in barley plastids: support for MatK as an essential splice factor. *J Mol Biol*. 270:179–187.

Westwood JH. Yoder JI, Timko M, de Pamphilis CW. Forthcoming. 2010. The evolution of parasitism in plants. *Trends Plant Sci*. 15:227–235.

Wickett NJ, Zhang Y, Hansen SK, Roper JM, Kuehl JV, Plock SA, Wolf PG, DePamphilis CW, Boore JL, Goffinet B. 2008. Functional gene losses occur with minimal size reduction in the plastid genome of the parasitic liverwort *Aneura mirabilis*. *Mol Biol Evol*. 25:393–401.

Wilson RJ, Denny PW, Preiser PR, et al. (11 co-authors). 1996. Complete gene map of the plastid-like DNA of the malaria parasite *Plasmodium falciparum*. *J Mol Biol*. 261:155–172.

Wolfe KH, Li WH, Sharp PM. 1987. Rates of nucleotide substitution vary greatly among plant mitochondrial, chloroplast, and nuclear DNAs. *Proc Natl Acad Sci U S A*. 84:9054–9058.

Wolfe KH, Morden CW, Palmer JD. 1992. Function and evolution of a minimal plastid genome from a nonphotosynthetic parasitic plant. *Proc Natl Acad Sci U S A*. 89:10648–10652.

Wostrikoff K, Girard-Bascou J, Wollman FA, Choquet Y. 2004. Biogenesis of PSI involves a cascade of translational autoregulation in the chloroplast of *Chlamydomonas*. *EMBO J*. 23: 2696–2705.

Yoon HS, Hackett JD, Bhattacharya D. 2002. A single origin of the peridinin- and fucoxanthin-containing plastids in dinoflagellates through tertiary endosymbiosis. *Proc Natl Acad Sci U S A*. 99: 11724–11729.

Young ND, dePamphilis CW. 2005. Rate variation in parasitic plants: correlated and uncorrelated patterns among plastid genes of different function. *BMC Evol Biol*. 5:16.

Zerges W. 2002. Does complexity constrain organelle evolution? *Trends Plant Sci*. 7:175–182.

Zoschke R, Nakamura M, Liere K, Sugiura M, Borner T, Schmitz-Linneweber C. 2010. An organellar maturase associates with multiple group II introns. *Proc Natl Acad Sci U S A*. 107: 3245–3250.

Zybailov B, Friso G, Kim J, Rudella A, Rodriguez VR, Asakura Y, Sun Q, van Wijk KJ. 2009. Large scale comparative proteomics of a chloroplast Clp protease mutant reveals folding stress, altered protein homeostasis, and feedback regulation of metabolism. *Mol Cell Proteomics*. 8:1789–1810.