Plastomes on the edge: the evolutionary breakdown of mycoheterotroph plastidgenomes

Sean W. Graham, Vivienne K. Y. Lam and Vincent S. F. T. Merckx

Review

New Phytologist

Tansley insight

Plastomes on the edge: the evolutionary breakdown of mycoheterotroph plastid genomes

Author for correspondence:
Sean W. Graham
Tel: +1 604 822 4816
Email: swgraham@interchange.ubc.ca

Received: *17 July 2016*
Accepted: *14 November 2016*

Sean W. Graham[1], Vivienne K. Y. Lam[1] and Vincent S. F. T. Merckx[2]

[1]Department of Botany, University of British Columbia, 6270 University Boulevard, Vancouver, BC V6T 1Z4, Canada;

[2]Understanding Evolution Group, Naturalis Biodiversity Center, Vondellaan 55, 2332 AA Leiden, the Netherlands

Contents

New Phytologist (2017) **214:** 48–55
doi: 10.1111/nph.14398

Key words: comparative genomics, convergent evolution, mixotrophy, mycoheterotrophic plants, photosynthesis genes, plastid function, plastid translation apparatus.

Summary

We examine recent evidence for ratchet-like genome degradation in mycoheterotrophs, plants that obtain nutrition from fungi. Initial loss of the NADH dehydrogenase-like (NDH) complex may often set off an irreversible evolutionary cascade of photosynthetic gene losses. Genes for plastid-encoded subunits of RNA polymerase and photosynthetic enzymes with secondary functions (Rubisco and ATP synthase) can persist initially, with nonsynchronous and quite broad windows in the relative timing of their loss. Delayed losses of five core nonbioenergetic genes (especially *trn*E and *acc*D, which respectively code for glutamyl tRNA and a subunit of acetyl-CoA carboxylase) probably explain long-term persistence of heterotrophic plastomes. The observed range of changes of mycoheterotroph plastomes is similar to that of holoparasites, although greater diversity of both probably remains to be discovered. These patterns of gene loss/retention can inform research programs on plastome function.

I. Introduction

Mycoheterotrophs are a remarkable guild of plants that survive by skimming nutrients from mycorrhizal or saprophytic fungal hyphae housed in modified root systems (Fig. 1), relaxing the need for carbon fixation from photosynthesis. They are often assumed to be parasitic, although their fungal partners may benefit in as yet uncharacterized ways (Selosse & Roy, 2009). Flowering individuals can be cryptic and hidden in the leaf litter of dark forest understoreys, some never emerging from the soil (e.g. *Rhizanthella*; Delannoy *et al.*, 2011). Several plant lineages are partial mycoheterotrophs (mixotrophs), using nutrient recovery from mycorrhizal partners to complement limited photosynthesis (Fig. 2a; trophic status confirmed in only a few cases, e.g. Hynson *et al.*, 2013). A few lineages, including orchids, have only an initial phase of mycoheterotrophy. However, *c.* 50 lineages of mycoheterotrophs in different land plant clades have completely lost photosynthetic ability (Fig. 2a; Merckx & Freudenstein, 2010). These full mycoheterotrophs represent replicated evolutionary experiments in the loss of photosynthetic function (analogous to holoparasitic plants, e.g. Westwood *et al.*, 2010; Wicke *et al.*, 2016), and its effect on genome evolution.

Full mycoheterotrophs display numerous modifications to their physiology, morphology and reproductive biology (Leake, 1994;

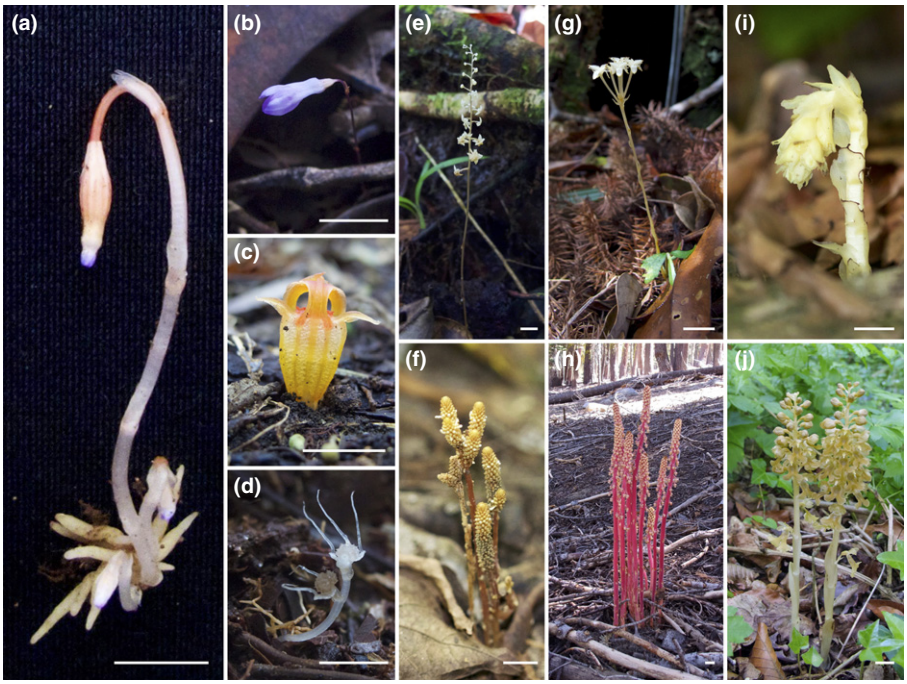

Fig. 1 Examples of fully mycoheterotrophic plants: (a) *Voyria tenella* (Gentianaceae), Brazil; (b) *Apteria aphylla* (Burmanniaceae), French Guiana; (c) *Thismia rodwayi* (Thismiaceae), Australia; (d) *Triuris hyalina* (Triuridaceae), French Guiana; (e) *Soridium spruceanum* (Triuridaceae), French Guiana; (f) *Epirixanthes cylindrica* (Polygalaceae), Malaysia; (g) *Petrosavia stellaris* (Petrosaviaceae), Malaysia; (h) *Pterospora andromedea* (Ericaceae), USA; (i) *Hypopitys monotropa* (Ericaceae), the Netherlands; (j) *Neottia nidus-avis* (Orchidaceae), the Netherlands. Bars, 1 cm.

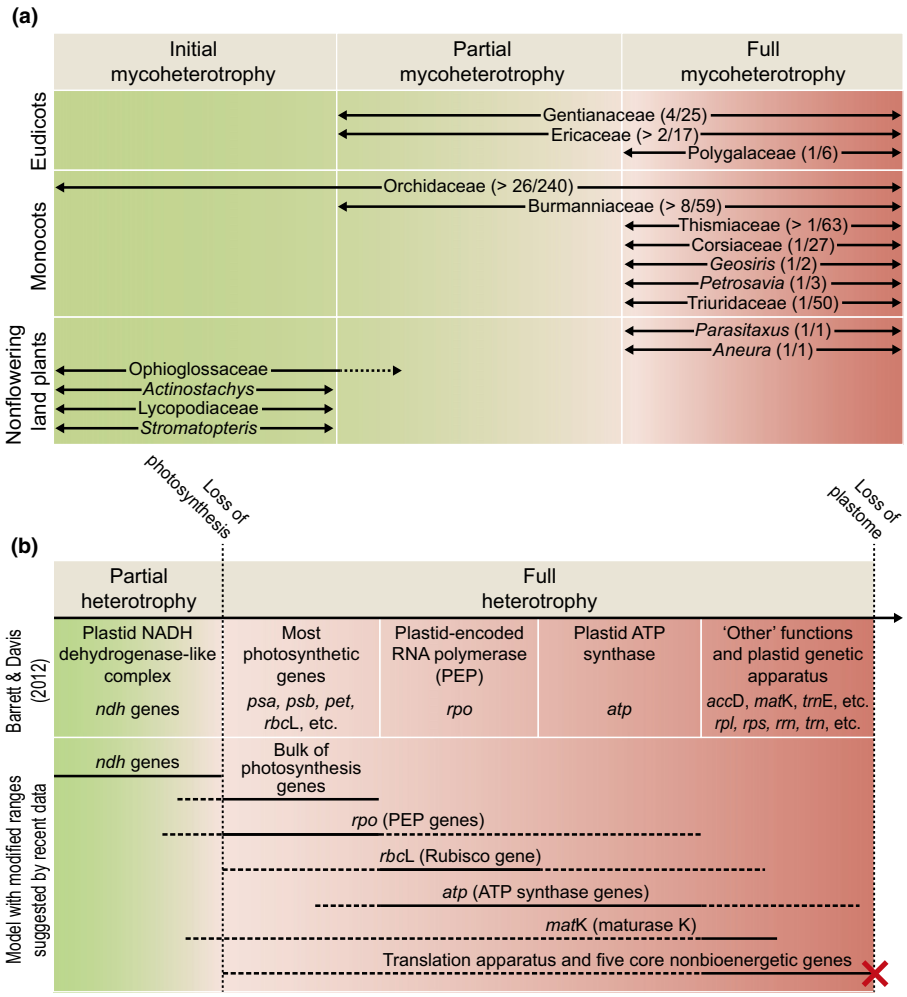

Fig. 2 (a) Diversity of mycoheterotrophs in land plants, with fungal nutrition obtained either in an initial life phase versus for the full life cycle, the latter with either partial or complete reliance on fungal partners (in parentheses: the numbers of origins of full mycoheterotrophy, and the number of fully mycoheterotrophic species in each group, after Merckx, 2013). (b) An irreversible model of gene loss in heterotrophic plants (Barrett & Davis, 2012), with modified ranges suggested by recent data (Table 1; Supporting Information Table S2; thick lines represent most likely points of loss; dashed lines are possible ranges; X, loss of final retained nonbioenergetic gene; pseudogenes may persist beyond boundaries shown in each case). For gene definitions, see Wicke *et al.* (2011).

Bidartondo, 2005; Imhof, 2010; Merckx, 2013). Photosynthesis loss also affects all three plant genomes (e.g. Bromham *et al.*, 2013; Petersen *et al.*, 2015; Wicke *et al.*, 2016), but the plastid genome (plastome) can experience especially severe disruption, including changes in structure and GC content, elevation of synonymous and nonsynonymous substitution rates, and gene losses leading to genome compaction (e.g. Wicke *et al.*, 2011). About 100–120 protein-coding, transfer RNA (tRNA) and ribosomal DNA (rDNA) genes are retained in the plastomes of photosynthetic land plants, a pattern of long-term stability that contrasts with inferred massive functional transfers to the nucleus in the early photosynthetic eukaryotes (e.g. Wicke *et al.*, 2011). Retained genes are mostly related to photosynthesis (light and dark reactions) and plastid genetic apparatus (transcription, transcript processing and translation; Wicke *et al.*, 2011). Their retention in green-plant plastids may be a consequence of difficulties in importing hydrophobic proteins produced by potential nuclear-transferred copies of plastid genes, or reflect a need to maintain plastid-gene-based regulation of redox balance or of other aspects of plastid physiology (Barbrook *et al.*, 2006). After a transition to full heterotrophy, loss of purifying selection in photosynthesis genes should lead to rapid gene extinction, although a subset of photosynthetic genes appears to be retained initially according to plastome studies from a variety of mycoheterotrophs (e.g. Barrett *et al.*, 2014; Lam *et al.*, 2015; Lim *et al.*, 2016) and parasitic plants (e.g. Naumann *et al.*, 2016; Bellot & Renner, 2016; Wicke *et al.*, 2016; see Supporting Information Table S1 for additional published examples, and Table S2 for a summary of gene loss/pseudogenization across taxa). These studies allow us to take stock of current hypotheses regarding how mycoheterotrophic genomes degrade following loss of photosynthesis, and to ask whether mycoheterotrophs and parasites differ substantially in trajectories of gene loss and retention.

II. Modelling plastid genome degradation in mycoheterotrophs

Closely related models with staged and irreversible losses of plastid-encoded genes in heterotrophic plants were proposed by Barrett & Davis (2012) and Barrett *et al.* (2014) to explain differences in gene content (summarized in the upper half of Fig. 2b). Here we review how well recent studies fit this stepwise model of genome degradation, focusing on the major proposed stages. It is assumed that individual genes, once lost, cannot be regained in the plastome (e.g. Selosse & Roy, 2009).

Loss of NADH dehydrogenase-like (*ndh*) genes

The NADH dehydrogenase-like (NDH-1) complex may often be the first system involving plastid-encoded genes to be lost, probably before the loss of photosynthesis (Barrett & Davis, 2012). This complex is now recognized to be a ferredoxin-plastoquinone reductase, rather than a genuine NAD(P)H dehydrogenase (Peltier *et al.*, 2016), and is thought to regulate photosynthetic electron flow to moderate the effects of photo-oxidative stresses, and to fine-tune photosynthesis performance under dynamic conditions (Yamori & Shikanai, 2016). *ndh* genes have been lost in several photosynthetic groups that are neither mycoheterotrophic nor parasitic (summarized in Ross *et al.*, 2016), and may be stably and functionally replaced by an alternative nuclear-encoded system in these taxa (Wicke *et al.*, 2011; Peltier *et al.*, 2016). By contrast, partial mycoheterotrophs often live in the forest understorey where photo-oxidative stress is reduced. Increased adaptation to understorey conditions may lessen the need to deal with this stress, and so the loss of NDH function in partial mycoheterotrophs (e.g. Table S2) may be accommodated with little immediate consequence for plant fitness. Once lost, an inability to deal with photo-oxidative stress could curtail them from reinvading high-light conditions, if an alternative plastid-targeted nuclear-encoded NDH system is lost for the same reason. NDH loss may prompt further investment in heterotrophy, initiating a series of additional irreversible gene losses in the evolution of mycoheterotrophy. Although not framed in terms of terms of NDH function, this is the scenario proposed by Selosse & Roy (2009) for partial heterotrophs before full loss of photosynthesis.

Loss of the bulk of photosynthesis genes

The next step in the evolution of full mycoheterotrophy is complete loss of photosynthetic function and the removal of selective pressure to retain photosynthetic plastid genes (Barrett & Davis, 2012; Barrett *et al.*, 2014). This probably requires pre-existing systems for nutrient recovery from fungal partners with little or no carbon repayment on investment, before the full loss of photosynthesis. Initial knockout mutations could affect any plastid- or nuclear-encoded gene for light and dark reactions (Fig. 2b), and should lead to pseudogenization and eventual complete loss of the other (now nonfunctional) photosynthetic genes by rapid accumulation of small deletions (e.g. Leebens-Mack & dePamphilis, 2002; Wicke *et al.*, 2016), accompanied by selection for reduced plastome size (Selosse *et al.*, 2001). Reduced selection on individual photosynthesis genes may also precede any loss of photosynthesis, as demonstrated by Wicke *et al.* (2016) in parasitic Orobanchaceae; it would be useful to confirm this in additional groups with mixed trophic status (Fig. 2a). There have also been reported losses of minor subunits of photosystem I and II (*psa* and *psb* genes) in photosynthetic *Corallorhiza* (partial mycoheterotrophs) and *Cuscuta* (hemiparasites) (Table 1). Similar reductions may occur in other lineages leading to reduced photosynthetic efficiency, but may be rarely captured in sequencing projects if the evolution of full heterotrophy usually follows rapidly. Estimates of rates of loss of these and other genes will be possible when more lineages are included in phylogenomic studies.

Retention and loss of photosynthesis genes with secondary functions

Several plastid-encoded systems involved in photosynthetic expression or function persist at least temporarily after the loss of photosynthesis and most photosynthetic genes (Fig. 2). Rubisco, the enzyme involved in photosynthetic carbon fixation, has been retained in several fully heterotrophic lineages (Table 1). Its

Table 1 Summary of loss (or retention) of plastid-encoded genes based on open-reading frame retention in heterotrophic plants (red text, mycoheterotrophs; blue text, parasitic plants). For gene definitions, see Wicke *et al.* (2011)

Plastid complexes (genes)	Photosynthetic phase	Partial mycoheterotrophy/ obligate hemiparasitism[1]	Full mycoheterotrophy/holoparasitism→ (i) Most photosynthesis genes lost/pseudogenes ('−') (retained reading frames probably nonfunctional) (ii) Functional retention ('+') of photosynthesis genes with other functions	Loss of remaining genes, ending in plastome loss
NADH dehydrogenase-like complex (*ndh* genes)	Lost in some photosynthetic taxa (unrelated to heterotrophy)	Lost in partial mycoheterotrophs and obligate hemiparasites: ERIC: *Pyrola*; ORCH: *Corallorhiza* CONV: *Cuscuta*; OROB: *Schwalbea* SANT: *Osyris*; VISC: *Viscum*		
Main photosynthesis genes (*psa*, *psb*, *pet*, *ccsA*, *cem*A, *lhb*A and *paf*)		Minor losses: ORCH: *psa*I, *psb*M in some *Corallorhiza* CONV: *psa*I in some *Cuscuta*		
Plastid-encoded polymerase PEP (*rpo* genes)		Can precede loss of main photosynthesis genes: CONV: *Cuscuta*	Retention (= delayed loss?) ANEU: *Aneura* (+*rpo*)	
Rubisco (*rbc*L) & plastid ATP synthase genes (*atp* genes)			Retention (=delayed loss?) ANEU: *Aneura* (+*atp*, +*rbcL*) ORCH: *Corallorhiza* (×2 lineages) (+*atp*) PETR: *Petrosavia* (+*atp*, +*rbc*L) OROB: *Lathraea* (+*atp*) OROB: *Myzorrhiza* (+*atp*, +*rbc*L) *Phelipanche*+*Orobanche* spp. (+*atp*)	
Group IIA intron maturase (*mat*K)[2]		*mat*K lost in: CONV: *Cuscuta*	*mat*K lost in: CORS: *Arachnitis*; THIS: *Thismia* ORCH: *Epipogium*; *Rhizanthella* APOD: *Pilostyles*; CYNO: *Cynomorium*; CYTI: *Cytinus*; HYDN: *Hydnora*	
Translation apparatus genes (*inf*A, *rpl*, *rps*, *rrn* and *trn*)	Rare losses of individual genes (mostly nuclear transfers?)		Translation apparatus genes losses accumulate (most/ all represent functional transfers outside plastid?): CORS: *Arachnitis*+*Corsia*; ERIC: *Monotropa*, *Hypopitys* ORCH: Multiple lineages; THIS: *Thismia*; TRIU: *Sciaphila* APOD: *Pilostyles*; CYNO: *Cynomorium*; CYTI: *Cytinus*; HYDN *Hydnora*; OROB: Multiple lineages	
Genes with roles outside photosynthetic and genetic apparatus (*acc*D, *clp*P, *trn*E, *ycf*1 and *ycf*2)	Rare losses (nuclear transfers?) of individual genes (e.g. *acc*D, *ycf*1 and *ycf*2 in Poales)		Gradual loss (or functional transfer outside plastid) CORS: *Arachnitis* (−*ycf*1, −*ycf*2) ERIC: *Hypopitys*, *Monotropa* (−*ycf*1, −*ycf*2) ORCH: *Epipogium* (−*ycf*1, −*ycf*2); TRIU: *Sciaphila* (−*ycf*1, −*ycf*2) THIS: *Thismia* (−*clp*P, −*ycf*1, −*ycf*2) APOD: *Pilostyles* (−*clp*P, −*trn*E, −*ycf*1, −*ycf*2) CYTI: *Cytinus*[3] (−*ycf*1, −*ycf*2); HYDN: *Hydnora* (−*clp*P) OROB: *Cistanche* (−*acc*D, −*ycf*1); *Phelipanche* (−*clp*P)	Plastid genome loss: RAFF: *Rafflesia*?

[1]Family abbreviations: APOD, Apodanthaceae; ANEU, Aneuraceae; CONV, Convolvulaceae; CORS, Corsiaceae; CYNO, Cynomoriaceae; CYTI, Cytinaceae; ERIC, Ericaceae; HYDN, Hydnoraceae; ORCH, Orchidaceae; OROB, Orobanchaceae; SANT, Santalaceae; THIS, Thismiaceae; TRIU, Triuridaceae; VISC, Viscaceae.
[2]*mat*K loss and patterns of group IIA intron loss summarized in Table 2; *mat*K status unclear in some *Neottia* (Orchidaceae) and *Viscum album* (Viscaceae).
[3]*acc*D status in *Cytinus* unclear (5′-end of *acc*D is truncated by several hundred bp; the remainder is an open reading frame).

retention could reflect its secondary role in lipid synthesis (e.g. Schwender *et al.*, 2004). ATP synthase has also been retained in multiple lineages (Table 1), which may be a common feature in moderately old lineages of full mycoheterotrophs. In photosynthetic plants, this enzyme complex is associated with the thylakoid, and uses a proton gradient generated by light reactions to generate ATP. It has also been implicated in producing proton gradients (via ATP hydrolysis) to enable protein transport across thylakoid membranes of photosynthetic plants (Kohzuma *et al.*, 2012; Kamikawa *et al.*, 2015). Whether it functions in this way in full mycoheterotrophs that retain plastid ATP synthase is unknown; this would also presumably require retention of plastid internal membrane systems, present in some heterotrophs (e.g. Walsh *et al.*, 1980). Finally, the transcription of plastid-encoded genes in green plants is mediated by PEP (plastid-encoded polymerase, an RNA polymerase encoded by plastid *rpo* genes) and a nuclear-encoded RNA polymerase (NEP). Most mycoheterotrophic lineages have lost PEP, although it appears to be retained in *Aneura*, a mycoheterotrophic liverwort (Table 1). This retention, if functional, could be related to a requirement for high transcription of retained Rubisco large subunit (*rbc*L) or ATP synthase genes in this taxon, as PEP is involved in high transcriptional activity in chloroplasts of green plants (Börner *et al.*, 2014). Alternatively, its *rpo* (PEP) genes may be nonfunctional but retained as a consequence of a recent loss of photosynthesis (Wickett *et al.*, 2008). The window for PEP loss may also be quite broad, as it has also been lost in hemiparasitic *Cuscuta* lineages capable of low levels of photosynthesis that retain other photosynthesis genes. The available data point to a tendency to lose *rpo* genes first, and *atp* genes last (Table 1), and to broader windows for the retention of these plastid-encoded complexes than the closely related models of Barrett & Davis (2012) and Barrett *et al.* (2014) predict (the latter combined PEP loss with the bulk of photosynthetic gene loss, and ATP synthase loss with later stages of plastid genome degradation).

Degradation and eventual loss of the plastid genetic apparatus

One or more of five nonbioenergetic genes (*acc*D, *clp*P, *trn*E, *ycf*1 and *ycf*2, which code, respectively, for a subunit of acetyl-CoA carboxylase, a proteolytic subunit of Clp-protease, glutamyl tRNA, and two plastid proteins of uncertain function) are retained in all plastomes sequenced to date, even highly degraded ones, potentially explaining the long-term persistence of the plastid translation apparatus that transcribes them, and the genes that code for this machinery. All heterotroph plastomes sequenced to date maintain at least some translation apparatus genes as open reading frames, even when other genes are lost or highly degraded (Tables 2, S2).

Table 2 Group IIA intron status in plastomes lacking *mat*K, a plastid gene that codes for plastid group IIA[1] intron maturase ('–' = intron missing)

		Status of introns (see Wicke *et al*. (2011) for gene definitions)							
Taxon	Status of *mat*K	*clp*P-intron 2[1]	*trn*A-UGC	*trn*I-GAU	*trn*K-UUU	*trn*V-UAC	*atp*F	*rpl*2	3′-*rps*12
Mycoheterotroph									
Corsiaceae									
Arachnitis uniflora	Lost	–[2]	–[3]	–[3]	–[3]	–[3]	–[3]	Present	Present
Orchidaceae									
Epipogium aphyllum	Lost	Present	–[3]	–[3]	–[3]	–[3]	–[3]	Present	–[3]
Epipogium roseum	Lost	Present	–[3]	–[3]	–[3]	–[3]	–[3]	Present	–[3]
Neottia nidus-avis[4]	Present?[4]	Present	Present	–[3]	Present	–[3]	–[3]	Present	Present
Rhizanthella gardneri	Lost	Present	–[3]	–[3]	–[3]	–[3]	–[3]	Present	–[3]
Thismiaceae									
Thismia tentaculata	Lost	–[3]	–[3]	–[3]	–[3]	–[3]	–[3]	Present	–[2]
Parasitic plants									
Apodanthaceae									
Pilostyles aethiopica	Lost	–[3]	–[3]	–[3]	–[3]	–[3]	–[3]	–[3]	–[3]
Pilostyles hamiltonii	Lost	–[3]	–[3]	–[3]	–[3]	–[3]	–[3]	–[3]	–[2]
Convolvulaceae									
Cuscuta gronovii	Lost	Present	–[3]	–[3]	–[3]	–[3]	–[2]	–[2]	–[2]
Cuscuta obtusiflora	Lost	Present	–[3]	–[3]	–[3]	–[3]	–[2]	–[2]	–[2]
Cynomoriaceae									
Cynomorium coccineum	Lost	Present	–[3]	–[3]	–[3]	–[3]	–[3]	Present	–[2]
Cytinaceae									
Cytinus hypocistis	Lost	–[2]	–[3]	–[3]	–[3]	–[3]	–[3]	–[2]	–[2]
Hydnoraceae									
Hydnora visseri	Lost	–[3]	–[3]	–[3]	–[3]	–[3]	–[3]	–[2]	Present
Viscaceae									
Viscum album	Present?[5]	Present	Present	Present	Present	–[3]	Present	Present	Present

[1]*clp*P second intron does not require *mat*K for splicing (Zoschke *et al.*, 2010).
[2]Intron lost (gene retained in plastome).
[3]Gene and intron lost from plastome (or gene pseudogenized).
[4]*mat*K open reading frame present, but position of start site unclear in some members of tribe Neottieae (taxa in Supporting Information Table S1).
[5]*mat*K status unclear (an indel disrupting the reading frame is potentially consistent with sequencing error).

This is consistent with the idea that many genes for the translation apparatus were successfully functionally transferred to the nucleus for reimport of protein or tRNAs (predicted by Wolfe *et al.*, 1992). The mechanism for purported tRNA import has been studied in mitochondria, but is still uncharacterized in plastids (e.g. Smith & Lee, 2014). If the translation apparatus of mycoheterotrophic *Thismia* and holoparasitic *Pilostyles* is still functional, the RNA products of rDNA genes may also be imported (Table S2). Patterns of gene loss from multiple independent lineages of heterotrophic plants are also suggestive of nonrandom gene loss (e.g. plastid ribosomal genes *rpl*23, *rps*16, *trn*A-UGC and *trn*G-UCG have been lost repeatedly; Table S2), and so some genes may be more easily transferred to the nucleus or replaced by nuclear counterparts, a hypothesis that could be tested in a phylogenetic framework (e.g. Cusimano & Wicke, 2016).

The plastid group IIA intron maturase *mat*K locus codes for a maturase that splices seven group IIA introns in land plant plastomes; an eighth group IIA intron (the second intron in *clp*P) does not require *mat*K to splice; other introns (group I or IIB) rely on nuclear-encoded splicing factors (Zoschke *et al.*, 2010). Barring functional transfer to the nucleus, *mat*K should only be lost when the seven group IIA introns that require it for splicing (or the genes that contain them) are lost. This appears to be the case in *Cuscuta* (McNeal *et al.*, 2009). Instances of purported *mat*K loss with retention of four to six group IIA introns may represent sequencing errors or alternative start sites (Table 2 footnote), or functional transfers of *mat*K to the nucleus. However, other *mat*K-lacking heterotrophic lineages retain at least some of these introns (*rpl*2 and/or 3′-*rps*12 in multiple mycoheterotroph and parasitic lineages; Table 2). These group IIA introns may therefore be able to self-splice (e.g. Delannoy *et al.*, 2011) or splice by other means in some taxa. Genes with retained introns may also be pseudogenes (Naumann *et al.*, 2016); but if so, it is odd that introns in these two genes are retained repeatedly (Table 2).

Retention and loss of the five core nonbioenergetic genes

Retention of plastid genomes in ancient mycoheterotrophs probably depends on at least one of five core nonbioenergetic genes being retained. The last retained core gene may usually be *trn*E (Barbrook *et al.*, 2006), as its gene product glutamyl tRNA plays a dual role in translation and in activation of haem biosynthesis (e.g. Howe & Smith, 1991; Barbrook *et al.*, 2006). Of the other four 'core' nonphotosynthetic genes in land plants, two have known roles. *acc*D has a critical role in lipid biosynthesis (e.g. Wicke *et al.*, 2011); *clp*P, a caseinolytic protease involved in plastid protein turnover (Wicke *et al.*, 2011), is also widely retained, with some losses documented (Tables 1, S2; Lam *et al.*, 2016). The *ycf*1 and *ycf*2 genes are lost more frequently. The former has a role in the plastid protein import (TOC/TIC translocon) machinery (De Vries *et al.*, 2015) and functions in photosynthetic protein complex biogenesis (Yang *et al.*, 2016); the role of the latter is unclear, but both appear to be essential for plastid function in most plants. Any one of these genes may be the 'last gene out'. However, *acc*D and *trn*E may be especially 'sticky', as the latter is found in all lineages except endoparasitic *Pilostyles* (Tables 2, S2), and the former is retained in all mycoheterotrophic lineages surveyed to date (Tables 2, S2; see Lam *et al.* (2016) for a more extensive survey). The *acc*D locus has been lost in some green plants, however. It can also be hard to identify in heterotrophs because of a high rate of sequence divergence (Logacheva *et al.*, 2016). Retention of any of these five genes in the plastome (or even ATP synthase or Rubisco) would explain plastome retention. We therefore predict that different scenarios of core gene retention in advanced full heterotrophs will be revealed by characterization of additional lineages.

III. Conclusions: a modified model for plastome degradation, and future research directions

The trajectories of gene loss appear to be essentially similar for mycoheterotrophs and parasitic plants (Tables 1, S2), although sequencing of additional heterotroph plastomes may reveal some differences. Recently sequenced plastomes largely uphold the core hypothesis proposed by Barrett & Davis (2012) and Barrett *et al.* (2014) (Fig. 2b). However, windows for retention for PEP, Rubisco and ATP synthase may be broader than supposed, with different tendencies for loss among these three systems; we therefore propose a slightly modified model in the lower half of Fig. 2(b). Small losses of peripheral photosystem I or II gene products may also precede full-scale loss of photosynthesis, before or after loss of *ndh* genes. The retention of one of the five nonbioenergetic genes (*acc*D, *clp*P, *trn*E, *ycf*1 and *ycf*2) long after the loss of other genes may explain plastome retention in ancient heterotrophic lineages (e.g. Lam *et al.*, 2015, 2016; Wicke *et al.*, 2016). Of these, *acc*D and *trn*E typically appear to be retained the longest. Plastome-less mycoheterotrophs may also exist, parallel to reported losses in *Polytomella* (a heterotrophic alga) and *Rafflesia* (a holoparasitic angiosperm); see Smith & Asmail (2014). This prediction is amenable to study by looking for absence of plastid-targeted nuclear-encoded genes in genomic data from these organisms.

Comparative analysis of plastomes has led to multiple functional predictions that could be tested in studies of plant physiology, including import of tRNAs (and rRNAs) into the plastid, and the possible splicing of several group IIA introns (beyond *clp*P) without *mat*K assistance. Future studies could aim to characterize transfers of missing genes from plastomes to the nucleus, for systems still thought to be functional (e.g. missing plastid translation apparatus genes in full mycoheterotrophs). They could also look for changes in selective regime in different stages of plastome degradation for partially and fully mycoheterotrophic taxa (e.g. applying the approaches of Wicke *et al.*, 2016). We also have only a limited idea of how long different phases in the trajectory of gene loss last, which needs to be followed up with molecular dating analyses. Dated phylogenies would be valuable for testing predictions that retained nonphotosynthetic genes experience an elevated rate of loss as a consequence of increasing dependence on external carbon. They would also allow us to test whether there is a general and increasing disruption of evolutionary stasis in heterotrophic plastomes, as a result of relaxation of selection on nuclear-encoded plastid DNA processing and repair systems (Wicke *et al.*, 2016). A better sampling of plastomes of mycoheterotrophic and related lineages is needed to address these unanswered questions.

Acknowledgements

We thank the Natural Sciences and Engineering Research Council of Canada (NSERC) for Discovery grant support of S.W.G., NSERC for a Postgraduate Fellowship and UBC for a Four-Year Fellowship awarded to V.K.Y.L., and anonymous reviewers for their comments.

References

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

Barrett CF, Davis JI. 2012. The plastid genome of the mycoheterotrophic *Corallorhiza striata* (Orchidaceae) is in the relatively early stages of degradation. *American Journal of Botany* **99**: 1513–1523.

Barrett CF, Freudenstein JV, Li J, Mayfield-Jones DR, Perez L, Pires JC, Santos C. 2014. Investigating the path of plastid genome degradation in early-transitional heterotrophic orchids, and implications for heterotrophic angiosperms. *Molecular Biology and Evolution* **31**: 3095–3112.

Bellot S, Renner SS. 2016. The plastomes of two species in the endoparasite genus *Pilostyles* (Apodanthaceae) each retain just five or six possibly functional genes. *Genome Biology and Evolution* **8**: 189–201.

Bidartondo MI. 2005. The evolutionary ecology of myco-heterotrophy. *New Phytologist* **167**: 335–352.

Börner T, Zhelyazkova P, Legen J, Schmitz-Linneweber C. 2014. Chloroplast gene expression–RNA synthesis and processing. In: Theg SM, Wollman F-A, eds. *Plastid biology*. New York, NY, USA: Springer, 3–48.

Bromham L, Cowman PF, Lanfear R. 2013. Parasitic plants have increased rates of molecular evolution across all three genomes. *BMC Evolutionary Biology* **13**: 1.

Cusimano N, Wicke S. 2016. Massive intracellular gene transfer during plastid genome reduction in nongreen Orobanchaceae. *New Phytologist* **210**: 680–693.

De Vries J, Sousa FL, Bölter B, Soll J. 2015. *Ycf*1: a green TIC? *Plant Cell* **27**: 1827–1833.

Delannoy E, Fujii S, Des Francs-Small CC, Brundrett M, Small I. 2011. Rampant gene loss in the underground orchid *Rhizanthella gardneri* highlight evolutionary constraints on plastid genomes. *Molecular Biology and Evolution* **28**: 2077–2086.

Howe CJ, Smith AG. 1991. Plants without chlorophyll. *Nature* **349**: 109.

Hynson NA, Madsen TP, Selosse M-A, Adam IKU, Ogura-Tsujita Y, Roy M, Gebauer G. 2013. The physiological ecology of mycoheterotrophic plants. In: Merckx VSFT, ed. *Mycoheterotrophy: the biology of plants living on fungi*. New York, NY, USA: Springer-Verlag, 297–342.

Imhof S. 2010. Are monocots particularly suited to develop mycoheterotrophy? In: Seberg O, Petersen G, Barfod A, Davis JI, eds. *Diversity, phylogeny, and evolution in the monocotyledons*. Copenhagen, Denmark: Aarhus University Press, 11–23.

Kamikawa R, Tanifuji G, Ishikawa SA, Ishii KI, Matsuno Y, Onodera NT, Ishida KI, Hashimoto T, Miyashita H, Mayama S *et al.* 2015. Proposal of a twin-arginine translocator system-mediated constraint against loss of ATP synthase genes from nonphotosynthetic plastid genomes. *Molecular Biology and Evolution* **32**: 2598–2604.

Kohzuma K, Dal Bosco C, Kanazawa A, Dhingra A, Nitschke W, Meurer J, Kramer DM. 2012. Thioredoxin-insensitive plastid ATP synthase that performs moonlighting functions. *Proceedings of the National Academy of Sciences, USA* **109**: 3293–3298.

Lam VKY, Merckx VSFT, Graham SW. 2016. A few-gene plastid phylogenetic framework for mycoheterotrophic monocots. *American Journal of Botany* **103**: 692–708.

Lam VKY, Soto Gomez M, Graham SW. 2015. The highly reduced plastome of mycoheterotrophic *Sciaphila* (Triuridaceae) is colinear with its green relatives and is under strong purifying selection. *Genome Biology and Evolution* **7**: 2220–2236.

Leake JR. 1994. The biology of myco-heterotrophic ('saprophytic') plants. *New Phytologist* **127**: 171–216.

Leebens-Mack JH, dePamphilis CW. 2002. Power analysis of tests for loss of selective constraint in cave crayfish and nonphotosynthetic plant lineages. *Molecular Biology and Evolution* **19**: 1292–1302.

Lim GS, Barrett CF, Pang C-C, Davis JI. 2016. Drastic reduction of plastome size in the mycoheterotrophic *Thismia tentaculata* relative to that of its autotrophic relative *Tacca chantrieri*. *American Journal of Botany* **103**: 1–9.

Logacheva MD, Schelkunov MI, Shtratnikova VY, Matveeva MV, Penin AA. 2016. Comparative analysis of plastid genomes of non-photosynthetic Ericaceae and their photosynthetic relatives. *Scientific Reports* **6**: 30042.

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**: e5983.

Merckx VSFT. 2013. *Mycoheterotrophy: the biology of plants living on fungi*. New York, NY, USA: Springer-Verlag.

Merckx VSFT, Freudenstein JV. 2010. Evolution of mycoheterotrophy in plants: a phylogenetic perspective. *New Phytologist* **185**: 605–609.

Naumann J, Der JP, Wafula EK, Jones SS, Wagner ST, Honaas LA, Ralph PE, Bolin JF, Maass E, Neinhuis C *et al.* 2016. Detecting and characterizing the highly divergent plastid genome of the nonphotosynthetic parasitic plant *Hydnora visseri* (Hydnoraceae). *Genome Biology and Evolution* **8**: 345–363.

Peltier G, Aro E-M, Shinakai T. 2016. NDH-1 and NDH-2 plastoquinone reductases in oxygenic photosynthesis. *Annual Review of Plant Biology* **67**: 55–80.

Petersen G, Cuenca A, Møller M, Seberg O. 2015. Massive gene loss in mistletoe (*Viscum*, Viscaceae) mitochondria. *Scientific Reports* **5**: 17588.

Ross TG, Barrett CF, Soto Gomez M, Lam VKY, Henriquez CL, Les DH, Davis JI, Cuenca A, Petersen G, Seberg O *et al.* 2016. Plastid phylogenomics and molecular evolution of Alismatales. *Cladistics* **32**: 160–178.

Schwender J, Goffman F, Ohlrogge JB, Shachar-Hill Y. 2004. Rubisco without the Calvin cycle improves the carbon efficiency of developing green seeds. *Nature* **432**: 779–782.

Selosse M-A, Albert B, Godelle B. 2001. Reducing the genome size of organelles favours gene transfer to the nucleus. *Trends in Ecology & Evolution* **16**: 135–141.

Selosse M-A, Roy M. 2009. Green plants that feed on fungi: facts and questions about mixotrophy. *Trends in Plant Science* **14**: 64–69.

Smith DR, Asmail SR. 2014. Next-generation sequencing data suggest that certain nonphotosynthetic green plants have lost their plastid genomes. *New Phytologist* **204**: 7–11.

Smith DR, Lee RW. 2014. A plastid without a genome: evidence from the nonphotosynthetic green algal genus *Polytomella*. *Plant Physiology* **164**: 1812–1819.

Walsh MA, Recehel EA, Popvich TM. 1980. Observations on plastid fine-structure in the holoparasitic angiosperm *Epifagus virginiana*. *American Journal of Botany* **67**: 833–837.

Westwood JH, Yoder JI, Timko MP, dePamphilis CW. 2010. The evolution of parasitism in plants. *Trends in Plant Science* **15**: 227–235.

Wicke S, Müller KF, dePamphilis C, Quandt D, Bellot S, Schneeweiss GM. 2016. Mechanistic model of evolutionary rate variation en route to a nonphotosynthetic lifestyle in plants. *Proceedings of the National Academy of Sciences, USA* **113**: 9045–9050.

Wicke S, Schneeweiss GM, dePamphilis CW, Müller KF, Quandt D. 2011. The evolution of the plastid chromosome in land plants: gene content, gene order, gene function. *Plant Molecular Biology* **76**: 273–297.

Wickett NJ, Zhang Y, Hansen SK, Roper JM, Kuehl JV, Plock SA, Wolfe 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*. *Molecular Biology and Evolution* **25**: 393–401.

Wolfe KH, Morden CW, Palmer JD. 1992. Function and evolution of a minimal plastid genome from a nonphotosynthetic parasitic plant. *Proceedings of the National Academy of Sciences, USA* **89**: 10648–10652.

Yamori W, Shikanai T. 2016. Physiological functions of cyclic electron transport around Photosystem I in sustaining photosynthesis and plant growth. *Annual Review of Plant Biology* **67**: 81–106.

Yang X-F, Wang Y-T, Chen S-T, Li J-K, Shen H-T, Guo F-Q. 2016. PBR1 selectively controls biogenesis of photosynthetic complexes by modulating translation of the large chloroplast gene *Ycf1* in *Arabidopsis*. *Cell Discovery* **2**: 16003.

Zoschke R, Nakamura M, Liere K, Sugiura M, Börner T, Schmitz-Linneweber C. 2010. An organellar maturase associates with multiple group II introns. *Proceedings of the National Academy of Sciences, USA* **107**: 3245–3250.

Supporting Information

Additional Supporting Information may be found online in the Supporting Information tab for this article:

Table S1 Sequenced plastid genomes of mycoheterotrophs, hemiparasites and holoparasites

Table S2 Heatmap summarizing retention, loss or pseudogenization of plastid genes in mycoheterotrophs and parasitic plants

Please note: Wiley Blackwell are not responsible for the content or functionality of any Supporting Information supplied by the authors. Any queries (other than missing material) should be directed to the *New Phytologist* Central Office.