Plant Mol Biol (2012) 79:5–20
DOI 10.1007/s11103-012-9884-3

Plastid genome evolution in mycoheterotrophic Ericaceae

Thomas Braukmann · Saša Stefanović

Received: 14 August 2011 / Accepted: 12 January 2012 / Published online: 23 March 2012

Abstract Unlike parasitic plants, which are linked to their hosts directly through haustoria, mycoheterotrophic (MHT) plants derive all or part of their water and nutrients from autothrophs via fungal mycorrhizal intermediaries. Ericaceae, the heather family, are a large and diverse group of plants known to form elaborate symbiotic relationships with mycorrhizal fungi. Using *PHYA* sequence data, we first investigated relationships among mycoheterotrophic Ericaceae and their close autotrophic relatives. Phylogenetic results suggest a minimum of two independent origins of MHT within this family. Additionally, a comparative investigation of plastid genomes (plastomes) grounded within this phylogenetic framework was conducted using a slot-blot Southern hybridization approach. This survey encompassed numerous lineages of Ericaceae with different life histories and trophic levels, including multiple representatives from mixotrophic Pyroleae and fully heterotrophic Monotropeae and Pterosporeae. Fifty-four probes derived from all categories of protein coding genes typically found within the plastomes of flowering plants were used. Our results indicate that the holo-mycoheterotrophic Ericaceae exhibit extensive loss of genes relating to photosynthetic function and expression of the plastome but retain genes with possible functions outside photosynthesis. Mixotrophic taxa tend to retain most genes relating to photosynthetic functions but are varied regarding the plastid *ndh* gene content. This investigation extends previous inferences that the loss of the NDH complex occurs prior to becoming holo-heterotrophic and it shows that the pattern of gene losses among mycoheterotrophic Ericaceae is similar to that of haustorial parasites. Additionally, we identify the most desirable candidate species for entire plastome sequencing.

Keywords Mycoheteotrophs · Ericaceae · Plastid genome · Southern hybridization · Phylogeny · *PHYA*

Introduction

Heterotrophic plants are usually divided into two morphologically distinct groups: parasitic and mycoheterotrophic (MHT) plants. While parasites establish direct haustorial connection with the host plant tissue, mycoheterotrophs use a third-party intermediary, mycorrhizal fungi and their symbiotic network, to link to their ultimate host. Some of these plants rely only partially on hosts and retain relatively unaffected ability to photosynthesize, following the so-called mixotrophic nutritional strategy. On the other end of the spectrum, holo-heterotrophs acquire all of their water, fixed carbon, and other nutrients from autotrophs. Consequences of these more dramatic trophic shifts can be seen from both morphological and molecular points of view. Namely, full heterotrophy is associated with extreme reduction and/or modification of vegetative structures as well as rampant morphological convergence, thus rendering an assessment of homology with their photosynthetic relatives quite difficult (Kuijt 1969). At the molecular level, estimating a species tree is difficult due to spurious long branch attraction caused by the highly divergent DNA sequences of heterotrophs (Nickrent et al. 1998; Barkman

Electronic supplementary material The online version of this article (doi:10.1007/s11103-012-9884-3) contains supplementary material, which is available to authorized users.

T. Braukmann (✉) · S. Stefanović
Department of Biology, University of Toronto Mississauga, 3359 Mississauga Rd. N, Mississauga, ON L5L 1C6, Canada
e-mail: thomas.braukmann@utoronto.ca

et al. 2007; Lemaire et al. 2010). For these reasons, precise phylogenetic relationships of heterotrophic plants to their respective autotrophic relatives have been notoriously difficult to ascertain (reviewed in Stefanović and Olmstead 2004). Nevertheless, broad-scale molecular investigations within flowering plants have shown that haustorial parasitism has evolved at least 12 times independently (Nickrent 2002; Nickrent et al. 2004; Barkman et al. 2007; Davis et al. 2007) and there are at least 10 independent origins of MHT (Bidartondo 2005; Merckx and Freudenstein 2010). Each of those lineages of heterotrophs could be seen as an independent natural genetic experiment whose plastid genes have evolved under relaxed functional constraints and therefore, each represents a unique opportunity to dissect plastid genome function and evolution.

The typical plastid genome (plastome) of flowering plants is highly conserved in size (~130–165 kbp), structure, gene content (~113 protein coding genes), and synteny (Palmer 1990; Palmer and Delwiche 1998; Ravi et al. 2008). It is generally composed of four functional classes of genes (Ravi et al. 2008), including: (1) genes coding for the photosynthetic apparatus (e.g., *psa, psb, atp, pet*, *rbcL, ndh*), (2) the housekeeping genes (e.g., *rpo, rps, rpl*), (3) genes with other functions (e.g., *accD, clpP, matK*), and 4) open reading frames (ORFs) with unknown function (i.e., *ycf* genes). Owing to relaxation of strong functional constraints normally associated with such a vital function as photosynthesis, the plastid genomes of heterotrophs exhibit a wide range of evolutionary degradation. The majority of currently available data comes from haustorial parasites. Some of their plastomes, primarily among hemiparasites, are impacted relatively little. Two species of *Cuscuta* subg. *Monogyna* (Convolvulaceae), *Cuscuta reflexa* and *Cuscuta exaltata*, have retained much of their plastid genomes (~121–125 kbp), and losses are restricted primarily to the chlororespiratory (*ndh*) genes and non-coding regions, such as intergenic spacers and introns (Funk et al. 2007; McNeal et al. 2007a). Others, especially among holoparasites, experienced further reductions. For example, *Cuscuta obtusiflora* and *Cuscuta campestris*, two closely related species from "clade B" of *Cuscuta* subg. *Grammica* (Stefanović et al. 2007), have substantially reduced plastomes (~85–87 kbp) but still maintain many genes required for photosynthesis (Funk et al. 2007; McNeal et al. 2007a). Plastomes of *Epifagus virginiana* (Orobanchaceae) are even smaller (~70 kbp) and many genes once involved in photosynthesis are either pseudogenes or are entirely lost from the plastome. Finally, there are putative cases of haustorial parasites for which the very existence of plastomes is questionable. Plastid DNA (ptDNA) could not be detected by Southern hybridization in some non-asterid holoparasites, such as *Corynaea* (Balanophoraceae) and *Hydnora* (Hydnoraceae; Nickrent et al. 1997) as well as in a large clade of predominantly South American species of *Cuscuta* subgenus *Grammica* ("clade O"; Stefanović et al. 2007). To date, there are only two MHT plant plastomes entirely sequenced: *Aneura mirabilis*, a liverwort (Wickett et al. 2008) and *Rhizanthella gardneri*, a geophytic orchid (Delannoy et al. 2011). In contrast to the highly reduced genome of fully MHT *Rhizanthella*, the plastid genome of *Aneura* has retained many genes related to photosynthesis, attributed to a presumably recent switch of this species to mycoheterotrophy (Wickett et al. 2008). Aside from these two examples, the plastomes of MHT species remain poorly characterized, especially among angiosperms.

Ericaceae, the heather family, are a large and diverse group of flowering plants, nearly cosmopolitan in distribution. This family is known to have elaborated symbiotic relationships with fungi. As currently circumscribed, based on broad-scale molecular and morphological analyses, Ericaceae s.l. is composed of eight subfamilies (summarized by Kron et al. 2002). Seven of those contain exclusively autotrophic species. All MHT taxa of various trophic levels, previously treated as segregate families (Monotropaceae and Pyrolaceae), are now classified in subfamily Monotropoideae. Within this group, mixotrophic (i.e., hemi-heterotrophic) species are confined to tribe Pyroleae while the fully MHT species are confined to tribes Pterosporeae and Monotropeae (Kron et al. 2002). Monophyly of each of these tribes is strongly supported; however, the phylogenetic relationships among them remain uncertain (Kron et al. 2002). First, it is unclear whether there is a single origin of mycoheterotrophy or whether these three tribes are examples of parallel evolution (Copeland 1941; Cullings 1994; Bidartondo and Bruns 2001; Merckx and Freudenstein 2010). Second, the position of Arbutoideae and its relationship to other autotrophic subfamilies as well as various MHT taxa also remains uncertain (Cullings 1994; Kron et al. 2002; Bidartondo and Bruns 2001). In an attempt to build upon previously available phylogenies and provide further resolution to some of these outstanding questions, we report here results of phylogenetic analyses inferred from the nuclear phytochrome A (*PHYA*) gene. Phytochrome genes have been shown to be powerful markers for phylogenetic studies at the higher phylogenetic levels (e.g., Poaceae, Mathews and Sharrock 1996; Brassicaceae, Beilstein et al. 2008) and in particular for heterotrophs (e.g., Orobanchaceae, Bennett and Mathews 2006).

Comparative analyses of the plastomes along the full trophic spectrum, from autotrophs to mixotrophs to full heterotrophs, would allow us to assess the degree to which genomic changes take place prior to complete loss of photosynthesis and to dissect the evolutionary constraints imposed by the presence of non-photosynthetic genes. In

this investigation, we gather data using a comprehensive Southern hybridization survey of plastid protein coding genes for an extensive sampling across Ericaceae. We interpret those data within a phylogenetic framework and in comparison with plastomes of other heterotrophs. Finally, we seek to identify the most interesting species that have highly modified plastid genomes, thus representing the prime candidates for entire plastome sequencing.

Materials and methods

Taxon sampling

Our sampling encompasses six of eight subfamilies in Ericaceae (Kron et al. 2002; Table 1). We focused most extensively on Monotropoideae, the subfamily traditionally grouping all MHT members of Ericaceae. We included 2/3 of its generic diversity (10 out of 15 genera), representing all three major lineages, tribes Pyroleae, Monotropeae, and Pterosporeae. For a number of their species with broad geographic distribution, we included multiple accessions to evaluate potential polymorphisms among populations (Table 1). As representatives of autotrophic lineages, we included species from five Ericaceae subfamilies: Enkianthoideae, Cassiopoideae, Arbutoideae, Ericodeae, and Vaccinoideae (sampling lacking for Stypheloideae and monotypic Harrimanelloideae). Taken together, our sampling strategy provides a broad phylogenetic background in which to compare the MHT members of various trophic levels to their autotrophic relatives. *Cyrilla racemiflora* and *Clethra barbinervis* were included as close outgroups (Kron et al. 2002). Voucher information for taxa included in the study are listed in Table 1.

A representative subset of 15 of these accessions (Table 1; underlined) was used for the molecular phylogenetic analysis based on single-copy nuclear *PHYA* sequence data. This analysis included sequences from three additional species that were not surveyed, *Ledum groenlandicum* Oeder (voucher: 05051, BIOUG), *Moneses uniflora* (voucher: 05060, BIOUG), and *Monotropastrum globosum* Andres ex Hara (GenBank accession number: AY348569). All sequences newly generated in this study are deposited in GenBank (accessions JQ248014-JQ248029).

DNA extraction, amplification, and sequencing

Total genomic DNA was extracted from fresh, silica dried, or herbarium material using a modified hexa-decyltrimethylammonium bromide (CTAB) technique (Doyle and Doyle 1987). Samples used in phylogenetic analyses were further purified using Wizard mini-columns (Promega). The nuclear genome region containing exon 1 of *PHYA* was amplified and sequenced using five primers (Supplementary Table 1) designed from regions conserved across *Monotropastrum*, *Solanum*, and *Cuscuta* (GeneBank accession numbers: AY348569, DQ208423, and AY348567, respectively). The polymerase chain reaction (PCR) reactions were carried out in 50 μL volumes with an annealing temperature of 60°C for 5 cycles followed by annealing temperature of 50°C for 30 cycles using high fidelity DNA Polymerase (Platinum® Taq; Invitrogen). Amplified products were cleaned by polyethylene glycol/NaCl precipitation and cloned using into the pSTBlue-1 AccepTor™ vector (EMD Biosciences). Multiple clones (2–5 clones) were cleaned and sequenced using the DYEnamic™ ET dye terminator sequencing kit (GE Healthcare) on an Applied Biosystems model 377 automated DNA sequencer (PE Biosystems). There were minimal substitution differences (1–5 bp) between sequenced clones, implying that only a single copy of *PHYA* was present. Sequence chromatograms were proofed, edited, and contigs were assembled using Geneious Pro v5.4.4 (Drummond et al. 2010). Sequences were aligned using the native Geneious alignment algorithm and then checked by eye. For the phylogenetic analyses, gaps were treated as missing data.

Phylogenetic analyses

Phylogenetic analyses were conducted under parsimony and maximum likelihood using PAUP* v4.0b10 (Swofford 2002). Given the moderate number of terminal units, the parsimony searches were conducted with a Branch-and-Bound algorithm, ensuring recovery of all of the most parsimonious (MP) trees. Matrix characters were treated as unordered (Fitch 1971), and all changes were equally weighted. ModelTest v3.7 (Posada and Crandall 1998) was used to determine the model of sequence evolution that best fit the data. According to both the hierarchical likelihood ratio test (hLRT) and Akaike information criterion (AIC), the general time–reversible (GTR) model of DNA substitution (Lanave et al. 1984), with rate variation among nucleotides following a discrete gamma distribution (GTR + G), was selected as the best-fit model. The full heuristic searches for maximum likelihood (ML) trees were performed under the selected model, involving 100 replicates with stepwise random taxon addition, tree bisection–reconnection (TBR) branch swapping, and MULTREES option on.

Under both criteria, the support for clades was inferred by nonparametric bootstrapping (Felsenstein 1985), using 1,000 heuristic bootstrap pseudoreplicates for MP and 100 heuristic bootstrap pseudoreplicates for ML analyses. Both analyses also included TBR branch swapping, and MULTREES option on. Support for a relationship was

Table 1 Ericaceae and close outgroups surveyed for the presence/absence of 47 plastid protein coding genes

FAMILY / subfamily / species	Voucher	23S	ndhA ex1	ndhA ex2	ndhB ex1	ndhB ex2	ndhC	ndhD 5′	ndhD 3′	ndhE	ndhF	ndhG	ndhH 5′	ndhH 3′	ndhI	ndhJ	ndhK	psaA	psaB 5′	psaB 3′	psaC	psbA	psbB	psbC	psbD	psbE	rbcL	petA	petB	petD	atpA	atpB	atpF	atpH	atpI	rpoA	rpoB	rpoC1	rpoC2	rps2	rps4	rps7	rps16	rpl14	rpl20	rpl23	rpl32	accD	cemA	clpP	matK	ycf2 3′	ycf4	16S
	Gene Number:		1		2		3	4		5	6	7	8		9	10	11	12	13		14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	
		NADH dehydrogenase																Photosystem I & II									CO2 fixation	Cytochrome b6/f complex			ATP synthase					RNA synthesis				Large and small ribosomal proteins								Genes with other function							
ERICACEAE																																																					
Vaccinioideae [VACC]																																																					
Gaultheria procumbens L.	*SS-07-97* (TRTE)	++	++	++	++	++	++	++	++	++	++	++	++	++	++	?	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	+	++	?	++	++	++	++	++	++	++
Gaultheria hispidula (L.) Muhl. Ex Bigelow	*SS-07-98* (TRTE)	++	++	++	++	++	++	++	++	++	++	++	++	++	++	?	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	+	++	+	++	++	++	++	++	++	++
Gaultheria procumbens L.	*SS-03-10* (TRTE)	++	++	++	++	++	++	++	++	++	++	++	++	++	+	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	+	++	+	++	++	++	++	++	++	++
Chamaedaphne calyculata Moench	*QIU-94180* (IND)	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	+	++	+	++	++	++	++	++	++	++
Leucothoe axillaris D. Don	*SS-06-147* (TRTE)	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	?	++	++	++	++	++	+	++	+	++
Andromeda glaucophylla Link	*QIU-94179* (IND)	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	+	++	+	++	++	++	+	++	+	++
Vaccinium cespitosum Michx.	*SS-05-148* (TRTE)	++	++	++	++	++	++	++	++	++	++	++	++	++	?	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	?	++	+	+	++	++	++	+	++	++
Vaccinium myrtilloides Michx.	*no voucher*	++	++	++	++	++	++	++	++	++	++	++	++	++	?	+	++	?	?	++	++	++	++	++	++	?	++	++	++	?	++	++	++	++	++	?	++	++	++	++	?	?	++	++	?	?	++	++	?	?	+	++	++
Ericoideae [ERIC]																																																					
***Epigaea repens* L.**	*SS-04-63* (TRTE)	++	++	++	++	++	++	++	++	++	++	++	++	++	++	+	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	+	++	++	?	+	+	++	++	++	++	+	++	++
***Phyllodoce glanduliflora* (Hook.)** Coville	*SS-03-87* (TRTE)	++	++	++	++	++	++	++	++	++	++	++	++	++	++	+	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	+	++	++	?	+	+	++	++	++	++	+	++	++
Phyllodoce empetriformis (Sm.) D. Don	*SS-03-86* (TRTE)	++	++	++	++	++	++	++	++	++	++	++	++	++	++	+	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	+	++	++	?	+	+	++	++	++	++	+	++	++
Phyllodoce caerulea (L.) Bab.	*SS-07-102* (TRTE)	++	++	++	++	++	++	++	++	++	++	++	++	++	++	+	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	+	++	++	?	++	+	++	++	++	++	+	++	++
Kalmia sp. L.	*no voucher*	++	++	++	++	++	++	++	++	++	++	++	++	++	++	+	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	+	++	++	++	++	+	++	++	++	++	+	++	++
Kalmia latifolia L.	*SS-03-08* (TRTE)	++	++	++	++	++	++	++	++	++	++	++	++	++	++	+	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	+	++	++	+	+	+	-	++	++	+	++	++
Menziesia ferruginea Sm.	*SS-03-50* (TRTE)	++	++	++	++	++	++	++	++	++	++	++	++	++	++	+	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	+	++	++	?	+	+	++	++	++	++	+	++	++
Cassiopoideae [CASS]																																																					
Cassiope mertensiana (Bong.) G.Don	*SS-03-85* (TRTE)	++	++	++	++	++	++	++	++	++	?	++	++	++	++	+	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	?	++	++	++	++	++	++	++	++	++	?	+	+	++	++	++	++	?	++	++
Cassiope tetragona (L.) D. Don	*SS-07-101* (TRTE)	++	++	++	++	++	++	++	++	++	++	?	++	++	?	+	++	?	?	++	++	++	++	++	++	?	++	++	++	?	++	++	?	++	++	?	++	++	++	++	++	?	?	++	++	?	+	?	++	++	++	++	?	+	++	++
Arbutoideae [ARBU]																																																					
Arctostaphylos uva-ursi (L.) Spreng.	*SS-03-90* (TRTE)	++	++	++	++	++	++	++	++	++	++	++	++	++	++	+	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	+	++	+	++	+	++	++	++	++	-	++	++
Arctostaphylos uva-ursi (L.) Spreng.	*no voucher*	++	++	++	++	++	++	++	++	++	++	++	++	++	++	+	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	+	++	+	++	++	++	++	++	++	-	++	++
Arctostaphylos nevadensis A. Gray	*SS-03-89* (TRTE)	++	++	++	++	++	++	++	++	++	++	++	++	++	++	+	++	?	?	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	+	++	?	+	?	?	?	++	++	?	?	?	++	++
Arctostaphylos columbiana Piper	*SS-03-88* (TRTE)	++	++	++	++	++	++	++	++	++	++	++	++	++	++	+	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	+	++	+	++	+	++	++	++	++	-	++	++
Monotropoideae [MONO]																																																					
Chimaphila umbellata (L.)W.P.C. Barton [NS]	*no voucher*	++	++	++	++	++	++	++	++	++	++	+	++	++	++	-	+	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	-	-	++	++	++	+	++	++
Chimaphila umbellata (L.)W.P.C. Barton [NS]	*SS-07-158* (TRTE)	++	++	++	++	++	++	++	++	++	+	++	++	++	+	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	-	-	++	++	++	+	++	++
Chimaphila umbellata (L.)W.P.C. Barton [WA]	*SS-03-41* (TRTE)	++	?	++	++	++	++	++	++	++	-	++	++	++	-	+	++	?	?	++	++	++	++	++	?	++	++	++	++	++	++	++	++	++	++	++	++	++	+	++	++	++	?	++	+	-	-	++	?	++	+	-	++
Chimaphila maculata (L.) Pursh	*SS-03-09* (TRTE)	++	++	++	++	++	++	++	++	++	-	++	++	++	++	+	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	-	-	++	++	++	+	++	++
Chimaphila sp. Pursh	*QIU-96084* (IND)	++	++	++	++	++	++	++	++	++	-	++	++	++	-	+	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	-	-	++	++	++	+	++	++
Moneses uniflora A. Gray	*SS-07-155* (TRTE)	++	-	++	++	++	+	++	++	++	-	-	-	-	+	+	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	?	-	-	++	++	++	+	++	++
Pyrola americana Sweet	*SS-07-156* (TRTE)	++	++	++	++	++	++	++	++	++	++	++	+	++	+	+	++	++	++	++	++	++	++	++	++	++	++	++	+	++	++	++	++	++	++	++	++	++	++	++	++	++	++	+	-	-	++	++	++	+	++	++
Pyrola asarifolia Michx.	*SS-03-40* (TRTE)	++	++	++	++	++	++	++	++	++	++	++	++	++	+	+	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	+	++	++	++	++	++	++	++	+	++	+	-	-	++	++	++	+	++	++
Pyrola asarifolia Michx.	*SS-05-144* (TRTE)	++	++	++	++	+	++	++	++	++	++	++	++	++	+	+	++	++	?	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	+	++	+	-	-	++	++	++	+	++	++
Pyrola asarifolia Michx.	*SS-04-188* (TRTE)	++	++	++	++	++	++	++	++	++	++	++	++	++	+	+	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	+	++	+	-	-	++	++	++	+	++	++
Pyrola chlorantha Sw.	*SS-03-42* (TRTE)	++	++	++	++	++	++	++	++	++	++	++	++	++	+	+	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	+	++	+	-	-	++	++	++	+	++	++
Pyrola chlorantha Sw.	*SS-03-47* (TRTE)	++	++	++	++	?	++	++	++	++	++	++	++	++	+	+	++	++	?	++	?	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	+	++	+	-	-	++	?	++	+	++	++
Pyrola picta Sm.	*SS-03-45* (TRTE)	++	++	++	++	++	++	++	++	++	++	++	++	++	+	+	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	?	-	-	++	++	++	+	++	++
Orthilia secunda L. House	*SS-07-157* (TRTE)	++	++	++	++	++	++	++	++	++	++	++	++	++	+	+	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	+	++	+	-	-	++	++	++	+	++	++
Orthilia secunda L. House	*QIU-96081* (IND)	++	++	++	++	++	++	++	++	++	++	++	++	++	+	+	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	+	++	+	-	-	++	++	++	+	++	++
Hypopitys monotropa L. [WA]	*SS-04-144* (TRTE)	++	-	+	-	-	+	+	-	-	-	-	-	-	+	-	-	-	-	-	-	-	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	++	+	-	-	++	+	+	-	-	++	+	-	-	++
Hypopitys monotropa L. [WA]	*SS-03-80* (TRTE)	++	-	+	-	-	+	+	-	-	-	-	-	-	+	-	-	-	-	-	-	-	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	++	+	-	++	-	++	+	+	-	-	+	+	-	-	++
Hypopitys monotropa L. [ON]	*SS-05-196* (TRTE)	++	-	+	-	-	+	+	-	-	-	-	-	-	+	-	-	-	-	-	-	-	+	-	+	-	-	-	-	-	-	-	-	-	-	-	-	-	+	++	+	-	++	-	++	+	+	-	-	+	+	-	-	++
Hypopitys monotropa L. [IN]	*SS-03-48* (TRTE)	++	-	?	+	-	-	+	-	+	-	-	-	-	+	-	-	-	-	-	-	-	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	++	+	-	++	+	++	+	+	-	-	+	+	-	-	++
Allotropa virgata Torr. & A. Gray	*no voucher*	++	-	?	-	-	-	-	-	-	-	-	-	-	?	-	-	-	-	-	-	-	?	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	++	++	++	-	++	+	++	+	-	-	++	++	-	-	++
Monotropa uniflora L. [WA]	*SS-03-82* (TRTE)	++	-	+	-	-	+	+	-	-	-	-	-	-	++	-	-	-	-	-	-	-	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	-	+	+	+	-	-	-	-	++	-	-	-	++
Monotropa uniflora L. [ON]	*SS-05-195* (TRTE)	++	-	+	-	-	+	+	-	-	-	-	-	-	++	-	-	-	-	-	-	-	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	-	+	+	++	+	-	-	-	++	+	-	-	++
Monotropa uniflora L. [ON]	*SS-06-237* (TRTE)	++	+	+	-	-	+	+	-	-	-	-	-	-	++	-	-	-	-	-	-	-	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	-	+	+	+	-	+	-	-	++	-	-	-	++
Monotropa uniflora L. [IN]	*SS-04-155* (TRTE)	++	+	+	-	-	+	+	-	-	-	-	-	-	++	-	-	-	-	-	-	-	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	-	+	-	++	+	-	-	-	++	+	-	-	++
Monotropa uniflora L. [MA]	*SS-06-246* (TRTE)	++	-	+	-	-	+	+	-	-	-	-	-	-	++	-	-	-	-	-	-	-	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	-	+	?	++	-	-	-	-	++	+	-	-	++
Monotropa uniflora L. [MI]	*SS-06-242* (TRTE)	++	-	+	-	-	+	+	-	-	-	-	-	-	++	-	-	-	-	-	-	-	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	-	+	+	++	-	-	-	-	++	+	-	-	++
Monotropa uniflora L. [MI]	*SS-04-180* (TRTE)	++	-	+	-	-	+	+	-	-	-	-	-	-	++	-	-	-	-	-	-	-	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	-	+	+	++	-	-	-	-	++	+	-	-	++
Pleuricospora fimbriolata A. Gray	*no voucher*	++	-	+	-	-	+	-	-	-	-	-	-	-	+	-	-	-	-	+	-	-	-	++	-	-	++	-	-	-	+	-	-	-	-	-	-	-	++	++	++	-	++	++	++	+	++	+	-	++	++	+	-	++
Pterospora andromedea Nutt.	*no voucher*	++	-	+	-	-	+	-	-	-	-	-	-	-	-	-	-	-	-	-	++	-	+	-	-	-	-	-	-	-	+	-	-	-	-	-	-	-	++	++	++	-	++	+	+	+	-	-	++	+	-	-	++
Pterospora andromedea Nutt.	*SS-99-01* (TRTE)	++	-	+	-	-	+	-	-	-	-	-	-	-	-	-	-	-	-	-	++	-	+	-	-	-	-	-	-	-	+	-	-	?	-	-	-	-	+	++	++	-	++	?	-	?	-	++	-	-	+	?	-	-	++
Sarcodes sanguinea Torr.	*Yatskievych et al., 02-186* (IND)	++	-	+	-	-	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	-	-	-	-	-	-	-	+	-	-	-	-	-	-	-	+	++	+	-	++	?	+	-	++	++	-	-	++	++
Enkianthoideae [ENKI]																																																				
Enkianthus campanulatus G. Nicholson	*no voucher*	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	-	++	++	++	++	++	++	++	++	++
Cyrillaceae																																																				
Cyrilla racemiflora L.	*QIU-95109* (IND)	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	+	++	++	++	++	++	++	++	++	++	++	++	++
Clethraceae																																																				
Clethra barbinervis Siebold & Zucc.	*QIU-95103* (IND)	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++	+	++	++	++	++	++	++	++	++	++	++	++

Large (23S) and small (16S) plastid ribosomal subunits were used as positive controls. Linear arrangement of species follows their presumptive phylogenetic relationships (see Kron et al. 2002 for details). Taxa used in our *PHYA* phylogenetic analysis (Fig. 1) are underlined. Based on slot-blot results, a "++" symbol indicates presence of full hybridization, "+" diminished hybridization, "–" absence of hybridization signal in comparison to positive controls, and "?" unable to score. Selected hybridization results for representatives from all major groups (species indicated in boldface) are shown in Fig. 2. Geographic locations for several broadly distributed species are indicated in brackets [state or province]. Four-letter abbreviations for subfamilies of Ericaceae are provided. Abbreviations of herbaria follow index herbarium

considered weak if bootstrap value was <65%, moderate if between 65 and 85%, and strong if >85%.

Three alternative topologies were constructed to further investigate relationships within Ericaceae. To statistically test and compare these alternatives with the optimal trees we conducted one-tailed Shimodaira-Hasegawa (SH) tests (Shimodaira and Hasegawa, 1999; Goldman et al. 2000) in PAUP* using 1,000 replicates and full parameter optimization of the model. We also carried out the approximately unbiased tests (AU tests; Shimodaira, 2002). The p values for the AU were calculated in CONSEL version v0.20 (Shimodaira and Hasegawa 2001), using 10 repetitions of multiscale bootstrapping, each consisting of 10 sets of 10,000 bootstrap replicates.

Hybridization

Due to limited quantity and poor quality of a number of samples derived from silica-gel or herbarium material, slot-blot hybridization was used. Detailed descriptions and rationale for this approach are provided in Doyle et al. (1995) and Braukmann et al. (2009). In brief, a slot-blot apparatus (Bio-Rad) was used to make seven sets of pseudoreplicate filter-blots, following the manufacturer's protocol. Approximately 500–800 ng of total DNA (per sample and per set) was bound to Immobilon-Ny+ nylon membrane (Millipore). Membranes were prehybridized, hybridized, and washed at 60°C. Probes were labeled with ^{32}P using random oligonucleotide primers (Invitrogen). Autoradiography was carried out using intensifying screens at −80°C for 18–48 h. DNA from tobacco (*Nicotiana tabacum* L.) was included on the blots as a positive control for the plastid probes. Prior to subsequent rounds of hybridization, the absence of carry-over signal was determined by an overexposure of decayed blots on a phosphor imaging screen for 6–8 h (Personal Molecular Imager™; Bio-Rad).

Hybridization probes for 47 plastid protein coding genes (Table 1) as well as controls (23S and 16S rDNA) were derived from tobacco via PCR. Two probes were used to survey genes interrupted by an intron, with each probe covering an exon. Also, longer genes were surveyed using two probes situated at the 5′ and 3′ ends, respectively. A total of 52 probes were used, sampling every major functional category of protein coding genes typically observed in green plant plastomes (refer to Wicke et al. 2011 for a detailed review). Primer names and sequences used to construct the probes are provided in the Supplementary Table 1. For each probe, their length, GC content, and the structural location within the plastome of tobacco are provided in Supplementary Table 2. To estimate the non-specific background hybridization levels, an initial negative hybridization control was performed under the same stringency conditions but without probe added.

Results

Phylogenetic analysis

Except for *Monotropastrum globosum* whose mRNA-derived sequence was downloaded from GenBank (AY348569), multiple clones were sequenced for all other species included in our phylogenetic analysis. Aside from minor (presumably allelic) differences, in all those cases only a single copy of *PHYA* has been recovered. Despite using only the protein coding sequence data, the *PHYA* exon 1 provided substantial amount of variability (1,450 aligned positions; 518 variable sites; 285 parsimony informative characters across 17 ingroup taxa) as well as overall good resolution and support for phylogenetic relationships within Ericaceae. ML-derived phylogram is shown in Fig. 1. Phylogeny obtained through MP analysis recovered a nearly identical topology (two equally parsimonious trees of 907 steps; trees not shown). Similar to other broad-scale phylogenetic analyses of Ericaceae (see Kron et al. 2002 and references therein), we obtained strong support for the monophyly of the family as well as the position of subfamily Enkianthoideae as sister to the rest of Ericaceae. Representatives of three autotrophic subfamilies characterized by a synapomorphy (early inversion of anthers from extrorse to introrse), Ericoideae, Vaccinioideae, and Cassiopoideae (Hermann and Palser 2000; Kron et al. 2002), are also recovered together, as a strongly supported clade (Ericaceae s.s.; Fig. 1). On the other hand, while the three MHT tribes (Pyloleae, Monotropeae, and Pterosporeae) are each strongly supported as monophyletic, their grouping into Monotopoideae is not (Fig. 1). Tests for alternative topologies rejected monophyly of this subfamily as traditionally defined (SH test $p < 0.001$; AU test $p = 3 \times 10^{-8}$). Contributing most notably to this is the position of Pterosporeae, strongly supported as a lineage distinct from other MHT taxa (100%; Fig. 1). Finally, the position of subfamily Arbutoideae remains uncertain. We recovered it as sister to the tribe Monotropeae on the optimal trees but the support for this relationship is only moderate to weak (78 and 60%, respectively for ML and MP analyses). However, alternative topology tests rejected (SH test $p < 0.001$; AU test $p = 7 \times 10^{-7}$) the consensus view where Arbutoideae are sister to other autotrophic Ericaceae (as per Kron et al. 2002). Also, we enforced Arbutoideae as sister to the clade containing both Pyroleae and Monotropeae, a topology suggesting a common origin of mycoheterotrophy for these two tribes. Both tests of alternative topology rejected this relationship of Arbutoideae with Pyroleae and Monotropeae (SH test $p = 0.042$; AU test $p = 0.006$), implying an independent origin of MHT for each of these two groups.

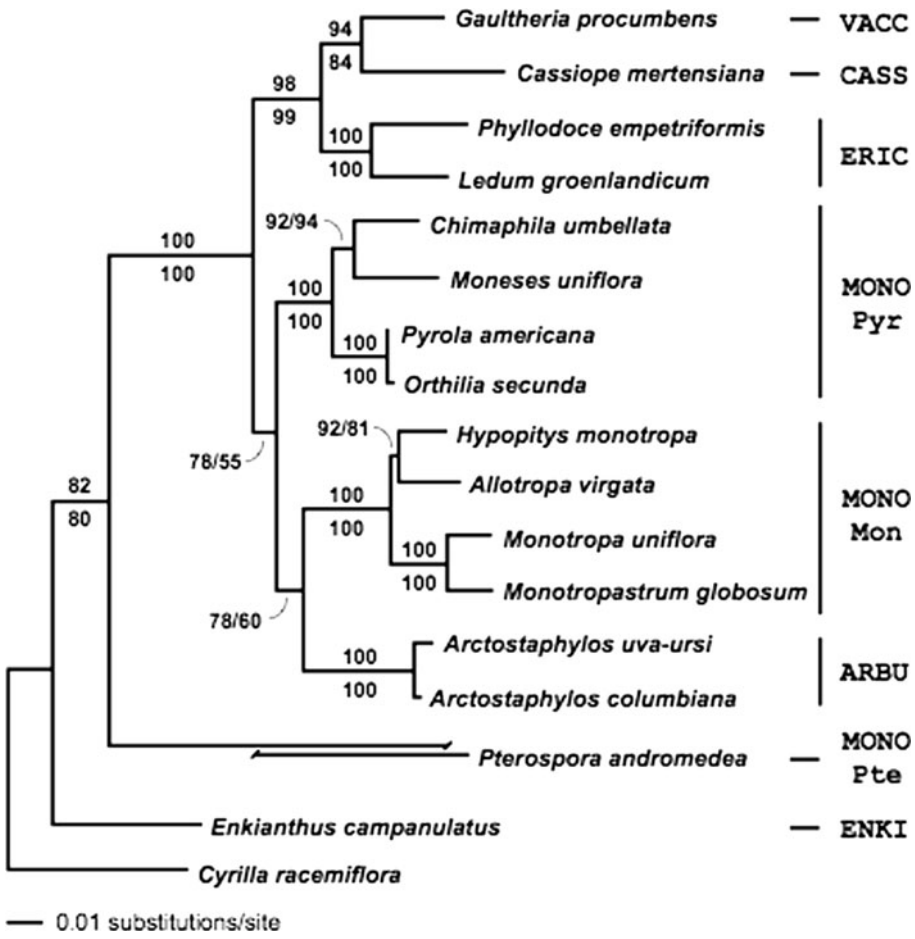

Fig. 1 Ericaceae phylogeny depicted as a phylogram obtained from maximum likelihood analysis of *PHYA* sequence data under the GTA + G model of DNA evolution. Four-letter abbreviations for subfamilies follow those from Table 1, three-letter abbreviations for three Monotropoideae tribes are as follows. *Mon* Monotropeae, *Pte* Pteroideae, *Pyr* Pyroleae. Numbers above and below branches indicate likelihood and parsimony bootstrap values, respectively

Interpretation of slot-blots

The presence or absence of plastid protein coding genes was determined by eye, by comparison of hybridization signal to the corresponding large and small ribosomal subunits probes. Given the conserved nature of 23S and 16S genes and their near ubiquitous presence among plant (Bendich 1987; Wicke et al. 2011), these two probes were used as controls to establish the presence of significant amounts of ptDNA on the blots as well as the baseline measure against which the presence or absence of other plastid genes was estimated. For each blot set and probe, the strength of signal was estimated by comparison to our positive control, tobacco, a species known to contain these genes based on previously available entire ptDNA sequence data (Shinozaki et al. 1986). Additionally, *Cyrilla racemiflora* and *Clethra barbinervis* were included to compare Ericaceae to more closely related autotrophic taxa.

A representative example of slot-blot data, arranged phylogenetically, is depicted in Fig. 2 and results for all of the surveyed species and probes are listed in Table 1. For all probes, the relative absence or presence of signal was scored for each taxon as indicating either full (++), diminished (+), absent (–), or unknown (?) in comparison with 23S and 16S positive controls. The full signal is assumed to indicate that the surveyed gene is present and putatively functional. For genes assayed with two probes (two exons or 5′ and 3′ end), full hybridization signal to both probes is necessary to indicate that a functional copy of the gene is present. Diminished or absent signals can be interpreted in several different ways. Diminished hybridization signal suggests either that the gene is present and functional but divergent with respect to tobacco or alternatively, that the homologous region is present as a pseudogene (i.e., rendered non-functional). Absence was scored if no detectable hybridization to a probe was observed. Given our experimental conditions, a gene transferred to the nucleus would not produce a hybridization signal when compared to a gene copy retained in the plastid genome. Transferred genes are significantly reduced in copy number and have accelerated substitution rates relative to the plastid (Wolfe et al. 1987). Given the typically low substitution rates for functional genes in ptDNA, a lack of signal suggests either loss of the gene or its transfer to the nucleus, rather than a highly divergent yet functional gene.

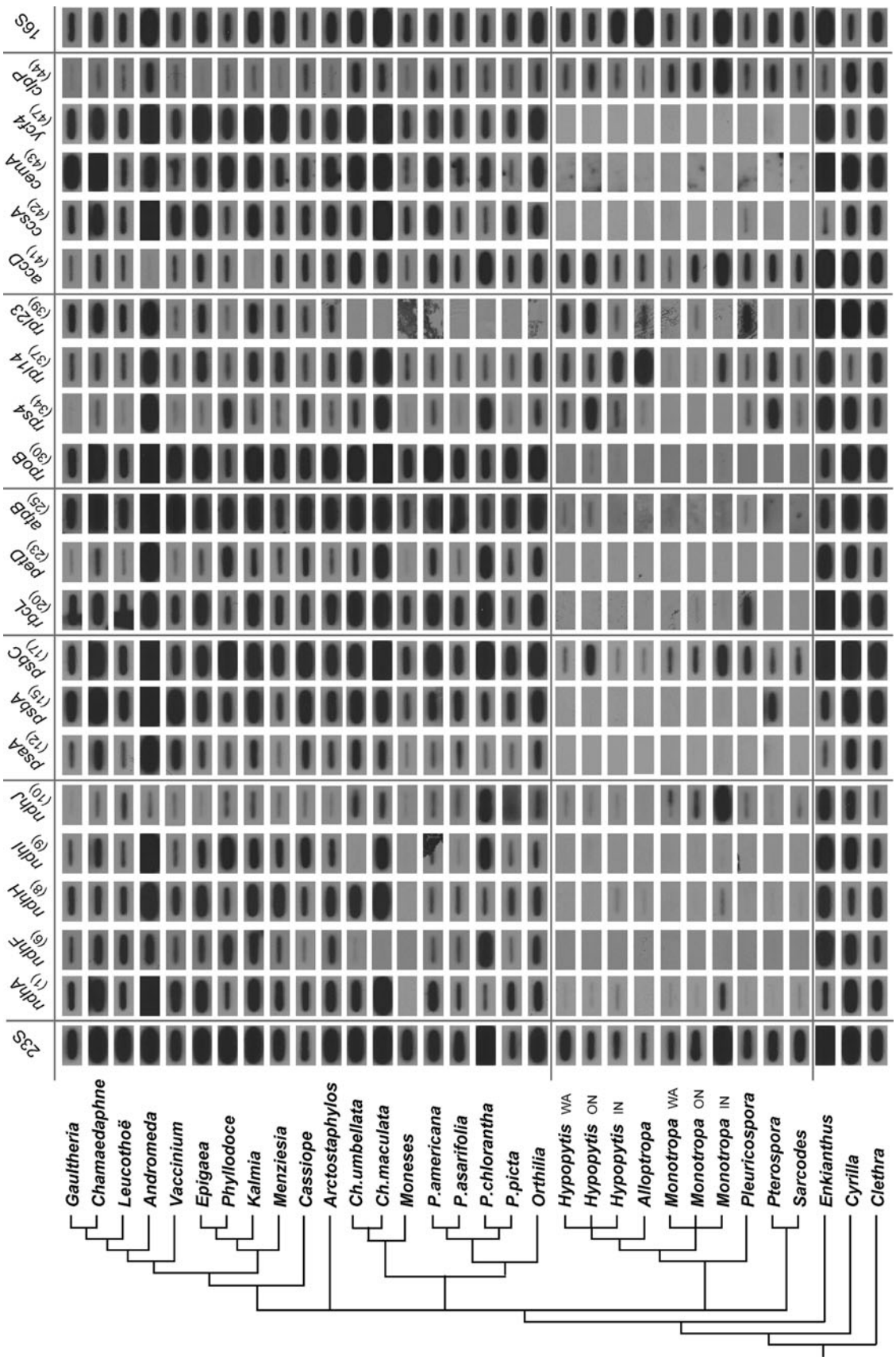

Fig. 2 Autoradiographs representing a subset of slot-blot hybridization results for the presence/absence of 47 plastid protein coding genes in Ericaceae and its close outgroups presented in a phylogenetic context. For full details, compare with Table 1. The large (23S) and small (16S) ribosomal subunits were used as positive controls (shown here is one representative out of seven sets). The topology shown is a composite tree depicting current understanding of relationships within Ericaceae derived from Kron et al. (2002) as well as our phylogenetic analysis based on *PHYA* sequences (Fig. 1). Note that mixotrophic and especially fully heterotrophic Ericaceae show the greatest number of absences or near absences of hybridization signal for the plastid genes

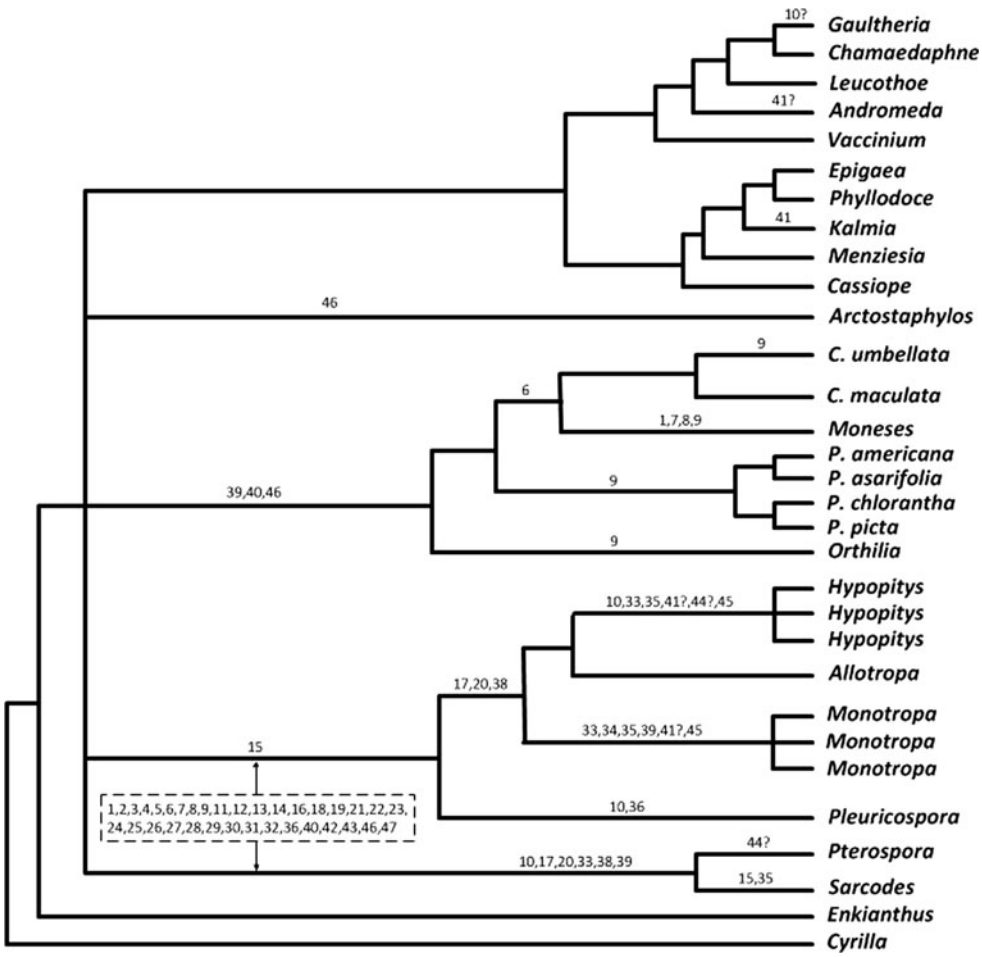

Fig. 3 Summary of functional losses of 47 protein coding genes within Ericaceae and outgroups inferred from hybridization survey. Numbers refer to the genes as enumerated in Table 1. Gene losses within a *box* indicate common losses to Monotropeae and Pterosporeae. Genes that are followed by a "?" indicate potentially divergent copies of a plastid genes, rather than a functional loss. The composite tree depicted is based on the relationships derived from Kron et al. (2002) as well as our phylogenetic analysis based on *PHYA* sequences (Fig. 1)

In certain cases, some taxa were scored as unknown ("?"; see Table 1). These ambiguities are a consequence of insufficient amounts or poor quality DNA for a given pseudoreplicate.

Given our assumptions, caveats of using southern hybridization are potential false positives or false negatives. For example, diminished signals that are interpreted as pseudogenes could be divergent but functional copies of the gene, while genes assumed to be present and functional could be recent pseudogenes. Despite these potential difficulties, southern hybridization allows for the evaluation of the gene content of a broad and diverse set of taxa in an efficient and cost-effective manner.

Distribution of gene losses

According to our investigations, autotrophic members of Ericaceae and the outgroups (*Clethra barbinervis* and *Cyrilla racemiflora*) typically exhibit full signal for all 47 plastid probes used in the survey (see Table 1; Figs. 2, 3 for details). These hybridization results were expected given the presence of these genes within the plastome of most flowering plants (Jansen et al. 2007) and were similar in strength to positive controls on the blots. The relative strength of the tobacco-based probes to most of our ingroup and outgroup taxa indicates the conserved nature of plastid genes across large phylogenetic distances, including the divergence of asterids (~107–117 Mya; Wikström et al. 2001).

Plastid-encoded NADH dehydrogenase (ndh) genes

The autotrophic members of Ericaceae generally showed undiminished signal for the *ndh* genes. An exception to this trend is *ndhJ,* which indicated diminished hybridization signal for all green Ericaceae, i.e., both autotrophic and mixotrophic members of this family. Given the strength of hybridizations of other *ndh* genes, we interpret this as a divergent copy of *ndhJ* with respect to tobacco. On the other end of the continuum, the overall lack of hybridization to *ndh* probes in fully MHT Ericaceae implies the functional loss of the NDH complex (Table 1). Sporadic presence of diminished hybridization signals for *ndh* genes among achlorophyllous Ericaceae most probably represent pseudogenes, and the variable strength of hybridizations among fully MHT taxa is likely a consequence of

Table 2 Comparison of the 47 plastid protein coding genes surveyed across Ericaceae with selected sequenced plastid genomes of heterotrophs and their autotrophic outgroups. Gene losses and pseudogenes are indicated for each protein coding gene category. Taxa with fully sequenced plastid genomes are indicated by an asterisk and achlorophyllous species are indicated in bold

Family *Species*	NADH dehydrogenase	Photosystem I and II	Cytochrom b6/f complex	ATP synthase
Solanaceae				
*Nicotiana tabacum**				
Orobanchaceae				
Epifagus viginiana*	*ndhA, ψndhB, ndhC-K*	*psaA-C, ψpsbA, ψpsbB, psbC, psbD, psbE*	*petA, petB, petD*	*ψatpA, ψatpB, atpF, atpH, atpI*
Convulvulaceae				
*Ipomoea purpurea**				
*Cuscuta gronovii**	*ndhA-K*	*psaI*		
*Cuscuta exaltata**	*ndhA, ψndhB, ndhC, ψndhD, ndhE-K*			
*Cuscuta obtusiflora**	*ndhA-K*	*psaI*		
Ericaceae				
Enkianthus campanulatus				
Chimaphila umbellata	*ψndhF, ψndhI*			
Moneses uniflora	*ndhA, ndhF, ndhG, ndhH, ψndhI, ψndhJ*			
Pyrola Americana	*ψndhI, ψndhJ*			
Orthilia secunda	*ψndhI, ψndhJ*			
Hypopitys monotropa	*ndhA, ndhB, ψndhC, ndhE-I, ψndhJ, ndhK*	*psaA-C, psbA, psbB, ψpsbC, psbD, psbE*	*petA, petB, petD*	*atpA, atpB, atpF, atpH, atpI*
Allotropa virgata	*ndhA-I, ndhJ?, ndhK*	*psaA-C, psbA, psbB, psbC?, psbD, pdbE*	*petA, petB, petD*	*atpA, atpB, atpF, atpH, atpI*
Monotropa uniflora	*ndhA, ndhB, ψndhC, ndhE-I, ndhK*	*psaA-C, psbA, psbB, ψpsbC, psbD, psbE*	*petA, petB, petD*	*atpA, atpB, atpF, atpH, atpI*
Pleuricospora fimbriolata	*ndhA, ndhB, ψndhC, ndhD-I, ψndhJ, ndhK*	*psaA-C, psbA, psbB, ψpsbC, psbD, psbE*	*petA, petB, petD*	*ψatpA, atpB, atpF, atpH, atpI*
Pterospora andromedea	*ndhA, ndhB, ψndhC, ndhD-K*	*psaA-C, psbB, ψpsbC, psbD, psbE*	*petA, petB, petD*	*ψatpA, atpB, atpF, atpH, atpI*
Sarcodes sanguinea	*ndhA, ndhB, ψndhC, ndhD-K*	*psaA-C, psbA, psbB, ψpsbC, psbD, psbE*	*petA, petB, petD*	*atpA, atpB, atpF, atpH, atpI*
Orchidaceae				
*Phalaenopsis aphrodite**	*ndhA, ψndhB, ψndhC, ψndhD, ψndhE, ndhF, ψndhG, ndhH, ψndhI, ψndhJ, ψndhK*			
Rhizanthella gardneri*	*ndhA-J, ψndhK*	*psaA, ψpsaB, psaC, psbA-E*	*petA, petB, petD*	*atpA, atpB, atpF, atpH, atpI*
Aneuraceae				
Aneura mirabilis	*ndhA, ψndhB-F, ndhG, ndhH, ndhI, ψndhJ, ndhK*	*ψpsaA, ψpsaB, ψpsbB-E*	*ψpetA, ψpetB*	

Table 2 continued

Family *Species*	CO_2 fixation	RNA synthesis	Large and small ribosomal proteins	Genes with other function
Solanaceae				
*Nicotiana tabacum**				
Orobanchaceae				
Epifagus viginiana*	*ψrbcL*	*ψrpoA, rpoB-C2*	*rps16, ψrpl14, ψrpl23, rpl32*	*cemA, ccsA, ycf4*
Convulvulaceae				
*Ipomoea purpurea**			*ψrpl23*	
*Cuscuta gronovii**		*rpoA-C2*	*rps16, rpl23, rpl32*	*matK, ψycf2*
*Cuscuta exaltata**			*rps16, ψrpl23*	
*Cuscuta obtusiflora**		*ψrpoA, rpoB-C2*	*rps16, rpl23, rpl32*	*matK*
Ericaceae				
Enkianthus campanulatus			*rps16*	
Chimaphila umbellata			*rpl23, rpl32*	*ψycf2*
Moneses uniflora			*ψrps16, rpl20, rpl23, rpl32*	*ycf2*
Pyrola Americana			*ψrpl20, rpl23, rpl32*	*ycf2*
Orthilia secunda			*ψrpl20, rpl23, rpl32*	*ycf2*
Hypopitys monotropa	*rbcL*	*rpoA, rpoB, ψrpoC1, ψrpoC2*	*ψrps2, ψrps7, rps16, rpl20, ψrpl32*	*ψaccD, ccsA, cemA, ψclpP, ψmatK, ycf2, ycf4*
Allotropa virgata	*rbcL*	*rpoA-C2*	*rps16, ψrpl20, ψrpl32*	*ccsA, cemA, clpP?, ycf2, ycf4*
Monotropa uniflora	*rbcL*	*ψrpoA, rpoB, ψrpoC1, ψrpoC2*	*ψrps2, rps4, ψrps7, rps7, ψrps16, rpl20, rpl23, rpl32*	*ψaccD, ccsA, cemA, ψmatK, ycf2, ycf4*
Pleuricospora fimbriolata		*rpoA-C2*	*rps16, ψrpl32*	*ψccsA, cemA, ψycf2, ycf4*
Pterospora andromedea	*rbcL*	*rpoA-C2*	*rps16, ψrpl20, ψrpl23, rpl32*	*ccsA, cemA, ψclpP, ycf2, ycf4*
Sarcodes sanguinea	*rbcL*	*ψrpoA, rpoB, rpoC1, rpoC2*	*ψrps2, ψrps7, rps16, ψrpl23, rpl32,*	*ccsA, cemA, ycf4*
Orchidaceae				
*Phalaenopsis aphrodite**				
Rhizanthella gardneri*	*rbcL*	*rpoA-C2*	*rps16, rpl32*	*ccsA, cemA, matK, ycf4*
Aneuraceae				
Aneura mirabilis				*ψccsA*

stochastic decay of these remnants (Table 1; Fig. 2). Mixotrophic species generally show presence for most *ndh* genes but are variable at a number of loci. For example, *Moneses uniflora* shows the most variation, with no hybridization signal for exon 1 of *ndhA*, *ndhF, ndhG, and ndhH* as well as diminished signals for *ndhC* and *ndhI*. Similar lack of uniform presence or absence of *ndh* genes was also observed within *Chimaphila*. Interspecific differences of hybridization signal for *ndhI* were evident with *C. maculata* exhibiting full hybridization signal and *C. umbellata* exhibiting either weak or absence of signal. Some intraspecific differences of signal were also observed within *C. umbellata* for *ndhF, ndhI*, and *ndhJ* (Table 1; Fig. 2). The variegated nature of presence and absence of the *ndh* genes in mixotrophic Ericaceae provides an opportunity to investigate more thoroughly the extent, pattern, and tempo of loss of the NDH complex (Table 2). In particular, the taxa with the greatest number of losses, such as *Monotropa uniflora* and *Chimaphila umbellata*, appear to be candidates most suitable for in-depth explorations via entire plastome sequencing.

Genes encoding the photosynthetic pathways

Various functional gene classes directly involved in photosynthesis (*psa, psb, atp, pet*, *rbcL*) were present in all green Ericaceae. An exception is weak hybridization for *petB* observed in mixotrophic *Pyrola americana*. Not surprisingly, among the full MHT taxa, there was generally an absence of signal for genes of the photosynthetic apparatus with four notable exceptions. First, *Pterospora andromedea* had full hybridization signal to *psbA*. Second, albeit weak, most full heterotrophs showed some signal for *psbC* and *Pleuricospora fimbriolata* had full hybridization to this probe. Third, the same species, *P. fimbriolata*, was scored as present for *rbcL* (Fig. 2; Table 1). Lastly, diminished signal for *atpA* is observed for *P. andromedea* and *P. fimbriolata*. The loss of photosynthetic genes in achlorophyllous Ericaceae is similar to other full heterotrophs which have abandoned photosynthesis and rely solely on their hosts for survival. However, due to their continued reliance on photosynthesis, mixotrophic taxa retain genes essential to photosynthetic function (Table 2).

Housekeeping genes

Similar to the genes involved in the photosynthetic pathways, green Ericaceae showed no decrease in hybridization signal for the RNA polymerase (*rpo*) genes. Amongst the fully MHT taxa, there was typically a weak to completely absent hybridization signal for *rpo* genes, indicative of the loss of plastid-encoded *rpo* (Table 2; Figs. 2, 3). For example, *Monotropa uniflora* showed some weak signals for *rpoA, rpoC1,* and *rpoC2*. In contrast, there was a complete absence of hybridization signal for all four *rpo* genes in *Pleuricospora fimbriolata* and *Pterospora andromedea*. This loss of the plastid-encoded polymerase (*rpo*) genes in fully MHT Ericaceae is similar to that seen in other holo-heterotrophs (Table 2). Green Ericaceae exhibited full to weak hybridization signal for the small and large subunit ribosomal protein probes (*rps* and *rpl* genes). Full hybridization signal was observed for *rps2, rps4,* and *rpl14*. For the remaining *rps* and *rpl* probes, hybridization ranged from weak to absent (Table 1). The absence of hybridization signal in Pyroleae for *rpl23* and *rpl32* is unique within Ericaceae and distinguishes this tribe from the rest of the family (Table 1; Figs. 2, 3). Fully MHT taxa were also highly variable in their hybridization signal to ribosomal proteins. Notably, *Moneses uniflora* exhibited the greatest number of absences, with no hybridization signal for five genes (*rps4, rpl20, rpl23*, and *rpl32*), while *P. fimbriolata* had the fewest absences (only *rps16* absent; see Table 1). The extensive loss of ribosomal proteins observed in *Moneses uniflora* is more pronounced in comparison to any other known heterotroph (Table 2), rendering this fully MHT species a prime candidate for the entire plastome sequencing.

Genes of other or unknown functions

Intron maturase (*matK*) is present in most taxa but has a weak to absent signal in *Monotropa uniflora*. Hybridizations for β-carboxyl transferase subunit of acetyl-CoA carboxylase (*accD*) typically exhibited presence within Ericaceae but there was diminished signal for *Hypopitys monotropa, Andromeda glaucophylla* and a complete absence of signal in *Kalmia latifolia* (see Table 1). To date, *accD* has been observed in all sequenced plastomes of heterotrophic plants, but is known to have been functionally transferred to the nucleus among autotrophic flowering plants at least 6 times independently (Jansen et al. 2007). Hybridizations for a membrane envelope protein (*cemA*) and heme attachment to cytochrome c biogenesis protein (*ccsA*) were present in green Ericaceae, but generally absent in achlorophyllous taxa. Finally, the hybridization signal for ATP-dependent protease (*clpP*) was diminished in most of Ericaceae except for the full signal in the MHT taxa *Pleuricospora fimbriolata, Sarcodes sanguinea,* and *Monotropa uniflora* (Table 1; Fig. 2). Generally weak hybridization for *clpP* across Ericaceae is potentially a result of a divergent plastid copy of *clpP* common to the family rather than a loss. Alternatively, the diminished signal could be due to a pseudogene of *clpP* present in the plastomes of many Ericaceae.

Overall, there was highly variable signal from the hybridization probes for the hypothetical chloroplast open

reading frames (*ycf*). Only *ycf4*, a gene putatively involved in photosystem I assembly (Boudreau et al. 1997), exhibits a pattern typical to most other plastid loci, with full hybridization to green taxa, and no hybridization in fully MHT taxa. The *ycf2 3′* probe was weak for most green Ericaceae with the absences restricted to *Artcostaphylos* and Pyroleae. MHT taxa did not produce hybridization signal for the *ycf2 3′* probe. The known plastomes of other heterotrophs have retained *ycf2*, and among heterotrophs the loss of *ycf2* appears to be restricted only to MHT Ericaceae.

Discussion

Phylogenetic relationships within Ericaceae

Similar to other studies using *PHYA* sequences for phylogenetic purposes (e.g., Mathews and Sharrock 1996; Bennett and Mathews 2006; Beilstein et al. 2008), assessment of primary homology among sequences was straightforward. Despite extensive cloning, only a single, presumably orthologous, copy of the gene was recovered in all species. Also, the protein-coding nature of this sequence allows for an easy and unambiguous alignment not only across diverse ingroup taxa but also between ingroups and outgroups. Given its relatively short length (~1,450 bp), the data matrix obtained from *PHYA* exon 1 was quite variable and phylogentically informative (518 variable sites across 17 operational units), resulting in a well-resolved topology (Fig. 1).

Most aspects of *PHYA*-derived results (Fig. 1) are in accordance with previous phylogenetic inferences for Ericaceae based on multiple plastid and/or nuclear ribosomal DNA sequences (Kron et al. 2002). An example of this includes the monophyly of the family in a broad sense, i.e., including members previously treated as segregate families, such as Pyrolaceae, Monotropaceae, etc. Also, the position of *Enkianthus* as sister to the rest of the family, the monophyly of Ericaceae s.s., a clade characterised by the early anther inversion character, as well as the monophyly of MHT tribes are all points of congruence with published phylogenies. In contrast to previous studies (Cullings 1994; Bidartondo and Bruns 2001; Kron et al. 2002), our gene tree suggests that Pterosporeae diverged early from the rest of Ericaceae, implying an additional origin of holo-heterotrophy in the family, independent from those in Monotropeae. However, caution is warranted when interpreting the relationships of Pterosporeae with the remaining members of the family. The position of this tribe could be an artifact stemming from long-branch attraction (LBA), which is known to result in strongly supported yet spurious results (Felsenstein 1978). Namely, as can be observed from the phylogram (Fig. 1), *Pterospora andromedea* and *Enkianthus campanulatus* are among the most divergent taxa for the *PHYA* sequences, and the recovered topology is potentially a result of the LBA phenomenon. Nevertheless, this result was recovered not only under the MP criterion employing equal evolutionary rates (Felsenstein 1978; Hendy and Penny 1989), but also by ML, a method using model of DNA evolution that explicitly accounts for rate heterogeneity (Felsenstein 1981; Lockhart et al. 1996; Stefanovic and Olmstead 2004). Also, both SH and AU tests strongly rejected alternative placement of *Pterospora*, as sister to Monotropeae.

Regardless of the exact position of Pterosporeae, mycoheterotrophy appears to have also evolved independently in Pyroleae and Monotropeae, given that autotrophic Arbutoideae are phylogenetically interjected between these two clades, being resolved as sister to Monotropeae. This sister relationship has received only weak to moderate support in our analyses (Fig. 1) but the SH and AU tests of alternative topologies rejected the position of Arbutoideae with Ericaceae s.s., a traditional placement for this subfamily (Kron et al. 2002), or as sister to a clade containing both Pyroleae and Monotropeae. In addition to our results, sister-group relationship between Arbutoideae and Monotropeae (but not Pyroleae) has been previously reported with both weak (Cullings 1994) and moderate support (Feldenkris et al. 2011). Further support for the affinity of Arbutoideae with MHT taxa comes from structural ptDNA feature, the shared loss of *ycf2* (Table 1; Fig. 3).

Relationships within Pyroleae suggested by our *PHYA* data (Fig. 1) are consistent with topologies recovered by nuclear ribosomal ITS and large subunit (26S) sequences (Freudenstein 1999; Liu et al. 2011), as well as some morphology-based studies (Krisa 1971), but not all (see Kron et al. 2002 for alternative views). Within Monotropeae, our topology indicates that *Monotropa uniflora* and *Hypopitys monotropa* (also known as *Monotropa hypopitys*) are not sisters to one another but rather *H. monotropa* is recovered as sister to *Allotropa,* and *Moneses uniflora* as sister to *Monotropastrum*. These relationships are strongly supported and are consistent with the results of other studies based on nuclear ribosomal ITS and 26S sequences (Bidartondo and Bruns 2001; Neyland and Hennigan 2004; Bidartondo 2005; Feldenkris et al. 2011), thereby supporting *Hypopitys* as a genus distinct from *Monotropa*. Additional *PHYA* sequence data, in particular those from introns found in this gene, combined with more extensive sampling would be useful in further elucidating relationships within both of these MHT tribes, and beyond.

Patterns of gene loss within Ericaceae

While there is a general trend in heterotrophic taxa for plastid gene loss, the extent by which these losses occur

depends largely on selective pressure to maintain any photosynthetic function (McNeal et al. 2007b; Krause 2008). The extent of gene loss in fully MHT Ericaceae with respect to the photosynthetic genes is similar to what has been observed for *Epifagus virginiana* and *Rhizanthella gardneri* (Wolfe et al. 1992; Delannoy et al. 2011).

A recurring pattern among heterotrophs is the loss of the plastid-encoded *ndh* genes, which are presumed to be the first genes lost in the transition to heterotrophy (McNeal et al. 2007a; Martin and Sabater 2010). Outside heterotrophic lineages, this complex is very rarely lost from plastomes (Braukmann et al. 2009), and among entirely or extensively sequenced plastomes of autotrophic angiosperms (see Jansen et al. 2007 for the most recent summary), its loss has been documented only in few members of Orchidaceae, Lentibulariaceae, and Geraniaceae (Wu et al. 2010; Blazier et al. 2011; Wicke et al. 2011). The absence of the entire suite of *ndh* genes in Monotropeae and Pterosporeae parallels the losses observed in *Epifagus, Cuscuta,* and *Rhizanthella* (Wolfe et al. 1992; Funk et al. 2007; McNeal et al. 2007a; Delannoy et al. 2011). The loss of the NDH complex is correlated with a decreased reliance on photosynthesis and it is thought to be dispensable in mild environments when maintaining photosynthetic rigour is no longer essential for survival (Martin and Sabater 2010). This is particularly true for mixotrophic plants living under forest canopies in which the ability to exploit neighbouring hosts improves the ability to survive low light conditions (Selosse and Roy 2009). If the NDH complex is dispensable under these conditions, then we can predict parallel loss of this complex in other lineages of mixotrophic plants with an otherwise preserved photosynthetic apparatus analogous to that in hemiparasitic *Cuscuta* species. Within mixotrophic Pyroleae, many of the *ndh* genes are still observed, and this provides an opportunity to investigate the loss of these genes from the plastome. *Moneses uniflora* and *Chimaphila umbellata* are the most affected species, with eight and three functional losses, respectively, and therefore they represent prime candidates for the whole plastid genome sequencing.

In contrast to the *ndh* genes, none of the genes involved directly in the photosynthetic pathway (i.e., *psa*, *psb*, *pet*, *atp*, and *rbcL*) are lost in any of mixotrophic Ericaceae. Comparable to autotrophic Ericacaeae, hybridization signal for the genes are strong, indicating strong selection for maintaining genes directly involved in photosynthesis in mixotrophic taxa. This is similar to hemiparasitic *Cuscuta* in which genes in the photosynthetic pathways have generally been retained, except for the loss of *psaI* (Funk et al. 2007; McNeal et al. 2007a). On the other hand, fully MHT Ericaceae has lost most of the genes in the photosynthetic pathway, including the *ndh* genes and have primarily retained pseudogene remnants, similarly to *Epifagus* and *Rhizanthella* (Wolfe et al. 1992; Delannoy et al. 2011). Notably, the large subunit of RuBisCO (*rbcL*) appears to have been retained in *Pleuricospora fimbriolata*. It has been previously hypothesized that *rbcL* potentially has function outside photosynthesis (Bungard 2004). Specifically, it can be involved in fatty acid synthesis in the cell or in transcriptional suppression during oxidative stress (Schwender et al. 2004; Moset et al. 2004; Krause 2008). The retention of an *rbcL* open reading frame has been observed in other fully heterotrophic taxa and requires further investigation to elucidate its role outside photosynthesis (Wolfe and dePamphilis 1997, 1998; Lusson et al. 1998; Delavault and Thalouarn 2002; Wickett et al. 2008; Krause 2008; Barrett and Freudenstein 2008).

The absence of plastid-encoded *rpo* genes from the plastome of fully MHT suggests a shift from plastid-encoded polymerase (PEP) to nuclear-encoded polymerase (NEP) for their remaining transcriptional units (Krause 2008). Similar transitions to NEP have been observed in *Epifagus* and *Cuscuta* and these transitions are presumed to precede loss of photosynthesis (Wolfe et al. 1992; McNeal et al. 2007a; Krause 2008; Delannoy et al. 2011). MHT Ericaceae also exhibit a number of losses of large and small ribosomal protein (*rpl* and *rps*) genes, more extensive than those observed in *Rhizanthella, Cuscuta,* and *Epifagus* (Wolfe et al. 1992; Funk et al. 2007; McNeal et al. 2007a; Delannoy et al. 2011). These losses indicate a greater reliance on nuclear encoded polymerases and ribosomal proteins to translate the remaining plastid genes. In several angiosperms, *rps16* is encoded in the nucleus and targeted to both the chloroplast and mitochondria (Ueda et al. 2008), as are many other proteins and tRNAs (Carrie et al. 2009). The loss of large and small ribosomal protein genes from plastid do not necessarily represent loss of these genes from the cell but perhaps point out toward an increased reliance on nuclear encoded products for plastid expression.

Group IIA intron maturase (*matK*) appears to be present, albeit divergent, across Ericaceae. A couple of populations of *Monotropa uniflora* lack hybridization signal for *matK* (see Table 1), but overall this gene appears to be present across MHT species as well. Hence, given the currently available data, the loss of this maturase seems to be restricted to some members of *Cuscuta* and *Rhizanthella* (McNeal et al. 2007a, b; Krause 2008). The retention of *matK* in fully MHT Ericaceae implies that there is still a demand to splice group IIA intron(s). Another common pattern shared with other heterotrophs is that both *clpP* and *accD* are retained in MHT Ericaceae. This strongly suggests that these genes have function outside photosynthesis and are expected to be retained by heterotrophic plants, as previously hypothesized by Bungard (2004) and Barbrook et al. (2006). Nevertheless, it is known that *clpP* and *accD*

can be functionally transferred to the nucleus and these loci have been lost multiple times from the plastids of flowering plants (3 and 6 times, respectively; Jansen et al. 2007). Interestingly, within our data set, an autotrophic species, *Kalmia latifolia*, appears to have lost its plastid *accD*, which potentially represents yet another functional transfer to the nucleus. Unexpectedly, *clpP* is divergent in most of Ericaceae and a divergent *clpP* differentiates Ericaceae from its close outgroups, Cyrillaceae and Clethraceae. Also, unique to Ericaceae is the weak hybridization signal for *ycf2*. Similar to *clpP*, this may represent a divergent gene, but could also be a pseudogene copy of *ycf2*.

Summary and prospects

This study provides the first comprehensive investigation of gene content in plastomes of MHT Ericaceae. There is a strong contrast in plastid gene content amongst Ericaceae of different trophic levels. Autotrophic Ericaceae generally retain all plastid genes investigated and within mixotrophic Pyroleae gene losses are restricted to the *ndh* genes (particularly *Moneses uniflora;* Table 1; Fig. 3), *ycf2*, and a few proteins of the large ribosomal subunit (*rpl23* and *rpl32*). However, a distinctive characteristic of some ericoid mixotrophs compared to all other published cases of sequenced plastomes is their variability regarding the presence and absence of plastid-encoded *ndh* genes. Plastid gene losses are concentrated primarily among fully MHT Ericaceae. These gene losses are associated with the loss of photosynthetic function, and for the most part, only genes with function outside photosynthesis seem to be retained. This trend is similar to other full heterotrophs sequenced to date, primarily among haustorial parasites, which also exhibit loss of most genes pertaining to photosynthesis (see comparison in Table 2).

This work, grounded in a phylogenetic framework, lays the foundation for further investigations of MHT species by whole plastome sequencing. Given the potential difficulties with obtaining the whole plastome sequences of fully heterotrophic plants (McNeal et al. 2007a, b; Delannoy et al. 2011), it is advantageous to have a priori information on heterotrophic plants to direct future sequencing efforts on the most interesting and information rich taxa. Our Southern hybridization revealed that the most promising cases for plastome sequencing among mixotrophic Ericaceae are *Moneses uniflora* and *Chimaphila maculata*. An in-depth investigation in these species will allow us, for example, to further explore the extent and tempo of losses of *ndh* genes. Among fully MHT species, *Monotropa uniflora* appears to be an ideal candidate for entire plastome sequencing. On one hand, this species exhibits more extensive losses of genes compared to other closely related holo-heterotrophs (e.g., a number of housekeeping genes). On the other hand, *Moneses uniflora* shows unexpected presence of hybridization signal for some *ndh* and *rpo* genes.

Acknowledgments For providing generous access to their live plant collections, the authors are grateful to directors/managers of the following greenhouses: Indiana University (Bloomington, IN, USA), the University of Toronto (Toronto, ON, Canada), and the University of Washington (Seattle, WA, USA). We would also like to thank Masha Kuzmina for collecting and providing plant material. Special thanks are due to Dan Nickrent and two anonymous reviewers for their valuable suggestions that considerably improved earlier versions of the manuscript. Financial support from the Natural Sciences and Engineering Research Council of Canada (grant no. 326439), the Canada Foundation for Innovation (grant no. 12810), and the Ontario Research Funds to S. Stefanović is gratefully acknowledged. We also thank the Natural Sciences and Engineering Research Council of Canada for the scholarship award provided to T. Braukmann.

References

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

Barkman TJ, McNeal JR, Lim S-H, Coat G, Croom HB, Young ND, dePamphilis CW (2007) Mitochondrial DNA suggests at least 11 origins of parasitism in angiosperms and reveals genomic chimerism in parasitic plants. BMC Evol Biol 7:248

Barrett CF, Freudenstein JV (2008) Molecular evolution of *rbcL* in mycoheterotrophic coralroot orchids (*Corallorhiza* Gagnebin, Orchidaceae). Mol Phylogenet Evol 47(2):665–679

Beilstein MA, Al-Shehbaz IA, Mathews S, Kellogg EA (2008) Brassicaceae phylogeny inferred from phytochrome A and *ndhF* sequence data: tribes and trichomes revisited. Am J Bot 95(10):1307–1327

Bendich AJ (1987) Why do chloroplasts and mitochondria contain so many copies of their genome? Bioessays 6:279–282

Bennett JR, Mathews S (2006) Phylogeny of the parasitic plant family Orobanchaceae inferred from phytochrome A. Am J Bot 93(7):1039–1051

Bidartondo MI (2005) The evolutionary ecology of myco-heterotrophy. New Phytol 167:335–352

Bidartondo MI, Bruns TD (2001) Extreme specificity in epiparasitic Monotropoideae (Ericaceae): widespread phylogenetic and geographical structure. Mol Ecol 10:2285–2295

Blazier JC, Gusinger MM, Jansen RK (2011) Recent loss of plastid encoded *ndh* genes within *Erodium* (Gerianaceae). Plant Mol Biol 76(3–5):263–272

Boudreau E, Takahashi Y, Lemieux C, Turmel M, Rochaix JD (1997) The chloroplast *ycf3* and *ycf4* open reading frames of *Chlamydomonas reinhardtii* are required for the accumulation of the photosystem I complex. EMBO J 16:6095–6104

Braukmann TWA, Kuzmina M, Stefanović S (2009) Loss of all plastid *ndh* genes in gnetales and conifers: extent and evolutionary significance for the seed plant phylogeny. Curr Genet 55(3):323–337

Bungard RA (2004) Photosynthetic evolution in parasitic plants: insight from the chloroplast genome. Bioessays 26:235–247

Carrie C, Giraud E, Whelan J (2009) Protein transport in organelles: dual targeting of proteins to mitochondria and chloroplasts. FEBS J 276:1187–1195

Copeland HF (1941) Further studies on Monotropoideae. Madroño 6:97–119
Cullings K (1994) Molecular phylogeny of the Monotropoideae (Ericaceae) with a note on the placement of Pyroloideae. J Evol Biol 7(5):501–516
Davis CC, Latvis M, Nickrent DL, Wurdack KJ, Baum DA (2007) Floral gigantism in Rafflesiaceae. Science 315:1812
Delannoy E, Fijii S, Colas des Francs C, Brundett M, Small I (2011) Rampant gene loss in the underground orchid *Rhizanthella gardneri* highlights evolutionary constraints on plastid genomes. Mol Biol Evol 28(56):2077–2086. doi:10.1093/molbev/msr028
Delavault P, Thalouarn P (2002) The obligate root parasite *Orobanche cumara* exhibits several *rbcL* sequences. Gene 297:85–92
Doyle JJ, Doyle JL (1987) A rapid DNA isolation procedure for small quantities of fresh leaf tissue. Phytochem Bull 19:11–15
Doyle JJ, Doyle JL, Palmer JD (1995) Multiple independent losses of two genes and one intron from legume chloroplast genomes. Mol Phylogenet Evol 5:429–438
Drummond AJ, Ashton B, Buxton S, Cheung M, Cooper A, Heled J, Kearse M, Moir M, Stones-Havas S, Sturrock S, Thierer T, Wilson A (2010) Geneious v5.4.4 Biomatters Ltd., Auckland. http://www.geneious.com
Feldenkris E, Broe M, Freudenstein J (2011) A mitochrondrial DNA and combined analysis at the base of Ericaceae. Botany 2011, St. Louis MO, Botanical Society of America, St. Louis Mo, abstract 19010
Felsenstein J (1978) The number of evolutionary trees. Syst Zool 27:27–33
Felsenstein J (1981) Evolutionary trees from DNA sequences: A maximum likelihood approach. J. Mol. Evol. 17:368–376
Felsenstein J (1985) Confidence limits on phylogenies—an approach using bootstrap. Evolution 39:783–791
Fitch WM (1971) Toward defining the course of evolution: minimum change for a specific tree topology. Syst Biol 20(4):406–416
Freudenstein JV (1999) Relationship and character transformation in Pyroloideae (Ericaceae) based on ITS sequences, morphology, and development. Syst Bot 24:398–408
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
Goldman N, Anderson JP, Rodrigo AG (2000) Likelihood based tests of topologies in phylogenetics. Syst Bot 49:652–670
Graham SW, Olmstead RG (2000) Utility of 17 chloroplast genes for inferring the phylogeny of the basal angiosperms. Am J Bot 87:1712–1730
Hendy MD, Penny D (1989) A framework for the quantitative study of evolutionary trees. Syst Zool 38:296–309
Hermann PM, Palser BF (2000) Stamen development in the Ericaceae. I. Anther wall, microsporogenesis, inversion, and appendages. Am J Bot 87(7):934–957
Hoot SB, Culham A, Crane PR (1995) The utility of the *atpB* gene sequences in resolving phylogenetic relationships: comparison with *rbcL* and 18S ribomsomal DNA sequences in the Lardizabalaceae. Ann MO Bot Gard 82(2):194–207
Jansen RK, Cai Z, Raubeson LA, Daniell H, dePamphilis CW, Leebens-Mack J, Muller KF, Guisinger-Bellian M, Haberle RC, Hansen AK, Chumley TW, Lee S, Peery R, McNeal JR, Kuehl JV, Boore JL (2007) Analysis of 81 genes from 64 plastid genomes resolves relationships in angiosperms and identifies genome-scale evolutionary patterns. Proc Natl Acad Sci USA 104:1369–1374
Krause K (2008) From chloroplasts to "cryptic" plastids: evolution of plastid genomes in parasitic plants. Curr Genet 54:111–121
Krisa B (1971) Beitrage zur taxonomie und chorologie der Gattung *Pyrola* L. Bot Jahrb Syst 90:476–508
Kron KA, Judd WS, Stevens PF, Crayn DM, Anderberg AA, Gadek PA, Quinn CJ, Luteyn JL (2002) Phylogenetic classification of Ericaceae: molecular and morphological evidence. Bot Rev 68:35–423
Kuijt J (1969) The biology of parasitic flowering plants. University of California Press, Berkeley
Lanave C, Preparata G, Saccone C, Serio G (1984) A new method for calculating evolutionary substitution rates. J Mol Evol 20:86–93
Lemaire B, Huysmans S, Smets E, Merckx V (2010) Rate accelerations in nuclear 18S rDNA of mycoheterotrophic and parasitic angiosperms. J Plant Res 124(5):561–576. doi:10.1007/s10265-010-0395-5
Liu ZW, Wang Z, Zhou J, Peng H (2011) Phylogeny of Pyroleae (Ericaceae): implications for character evolution. J Plant Res 124:325–337
Lockhart PJ, Larkum AWD, Steel MA, Waddell PJ, Penny D (1996) Evolution of chlorophyll and bacteriochlorophyll: the problem of invariant sites in sequence analysis. Proc Natl Acad Sci USA 93:1930–1934
Lusson NA, Delavault PM, Thalouarn PA (1998) The *rbcL* gene from the non-photosynthetic parastite *Lathraea clandestina* is not transcribed by a plastid-encoded RNA polymerase. Curr Genet 34:212–215
Martin M, Sabater B (2010) Plastid *ndh* genes in plant evolution. Plant Physiol Biochem 48:636–645
Mathews S, Sharrock RA (1996) The phytochrome gene family in grasses (Poaceae): a phylogeny and evidence that grasses have a subset of loci found in dicot angiosperms. Mol Biol Evol 13(8):1141–1150
McNeal JR, Kuehl JV, Boore JL, dePamphilis CW (2007a) Complete plastid genome sequences suggest strong selection for retention of photosynthetic genes in the parasitic plant genus *Cuscuta*. BMC Plant Biol 7:57
McNeal JR, Arumugunathan K, Kuehl JV, Boore JL, dePamphilis CW (2007b) Systematics and plastid genome evolution of the cryptically photosynthetic parasitic plant genus *Cuscuta* (Convolvulaceae). BMC Biol 5:55
Merckx V, Freudenstein JV (2010) Evolution of mycoheterotrophy in plants: a phylogenetic perspective. New Phytol 185:605–609
Neyland R, Hennigan MK (2004) A cladistic analysis of *Monotropa uniflora* (Ericaceae) inferred from large ribosomal subunit (26S) rRNA gene sequences. Castanea 69:265–271
Nickrent DL (2002) Orígenes filosgenéticos de las plantas parásitas. In: López-sáez JA, Catalán P, Sáez L (eds) Plantas parásitas de la Península Ibérica e Islas Baleares. Mundi-Prensa Libros, Madrid, pp 29–56
Nickrent DL, Ouyang Y, Duff RJ, dePamphilis CW (1997) Do nonasterid holoparasitic flowering plants have plastid genomes? Plant Mol Biol 34:731–743
Nickrent DL, Duff RJ, Colwell AE, Wolfe AD, Young ND, Steiner KE, dePamphilis CW (1998) Molecular phylogenetic and evolutionary studies of parasitic plants. In: Soltis D, Soltis P, Doyle J (eds) Molecular systematics of plants II. DNA sequencing. Kluwer, Boston, pp 211–241
Nickrent DL, Blarer A, Qiu Y-L, Vidal-Russel R, Anderson FE (2004) Phylogenetic inference in Rafflesiales: the influence of rate heterogeneity and horizontal gene transfer. BMC Evol Biol 4:40
Olmstead RG, Sweere JA (1994) Combining data in phylogenetic systematics: an empirical approach using three molecular data sets in the Solanaceae. Syst Biol 43:467–481
Palmer JD (1990) Contrasting modes and tempos of genome evolution in land plant organelles. Trends Genet 6:115–120
Palmer JD, Delwiche CF (1998) The origin and evolution of plastids their genomes. In: Soltis DE, Soltis PS, Doyle JJ (eds) Molecular systematics of plants II. Kluwer, Boston, pp 375–409

Posada D, Crandall KA (1998) Modeltest: testing the model of DNA substitution. Bioinformatics 14:817–818

Ravi V, Khurana JP, Tyagi AK, Khurana P (2008) An update on chloroplast genomes. Plant Syst Evol 271:101–122

Selosse MA, Roy M (2009) Green plants that feed on fungus: facts and questions about mixotrophy. Trends Plant Sci 14(2):64–70

Shimodaira H (2002) An approximately unbiased test of phylogenetic tree selection. Syst Bot 51:369–381

Shimodaira H, Hasegawa M (1999) Multiple comparisons of loglikelihoods with applications to phylogenetic inference. Mol Biol Evol 16:1114–1116

Shimodaira H, Hasegawa M (2001) CONSEL: for assessing the confidence of phylogenetic tree selection. Bioinformatics 17:1246–1247

Shinozaki K, Ohme M, Tanaka M, Wakasugi T, Hayashida N, Matsubayashi T, Zaita N, Chunwongse J, Obokata J, Yamaguchi-Shinozaki K, Ohto C, Torazawa K, Meng BY, Sugita M, Deno H, Kamogashira T, Yamada K, Kusuda J, Takaiwa F, Kato A, Tohdoh N, Shimada H, Sugiura M (1986) The complete nucleotide sequence of the tobacco chloroplast genome: its gene organization and expression. EMBO J 5:2043–2049

Stefanović S, Olmstead RG (2004) Testing the phylogenetic position of a parasitic plant (*Cuscuta*, Convolvulaceae, Asteridae): Bayesian inference and the parametric bootstrap on data drawn from three genomes. Syst Biol 53:384–399

Stefanović S, Kuzmina M, Costea M (2007) Delimitation of major lineages within *Cuscuta* subgenus *Grammica* (Convolvulaceae) using plastid and nuclear DNA sequences. Am J Bot 94(4):568–589

Swofford DL (2002) Phylogenetic analysis using parsimony (* and other methods), version 4.0b10. Sinauer, Sunderland

Ueda M, Nishikawa T, Fujimoto M, Takanashi H, Arimura S, Tsutsumi N, Kadowaki K (2008) Substitution of the gene for chloroplast RPS16 was assisted by generation of dual targeting signal. Mol Biol Evol 25(8):1566–1575

Wicke S, Schneeweiss GM, dePamphilis CW, Muller KF, Quandt D (2011) The evolution of the plastid chromosome in land plants: gene content, gene order, gene function. Plant Mol Biol 76:273–297

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

Wikström N, Savolainen V, Chase MW (2001) Evolution of the angiosperms: Calibrating the family tree. Proceedings. Biological Sciences 268: 2211–2220

Wolfe AD, dePamphilis CW (1997) Alternate pathways of evolution for the photosynthetic gene *rbcL* in four non-photosynthetic species of Orobanche. Plant Mol Biol 33:965–977

Wolfe AD, dePamphilis CW (1998) The effect of relaxed functional constraints on the photosynthetic gene *rbcL* in photosynthetic and non-photosynthetic parasitic plants. Mol Biol Evol 15:1243–1258

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

Wolfe KH, Morden CW, Ems SC, Palmer JD (1992) Rapid evolution of the plastid translational apparatus in a nonphotosynthetic plant: loss or accelerated sequence evolution of tRNA and ribosomal protein genes. J Mol Evol 35:304–317

Wu F, Chan M, Liao D, Hsu C, Lee Y, Daniell H, Duvall MR, Lin C (2010) Complete chloroplast genome of *Oncidium* Gower Ramsey and evaluation of molecular markers for identification and breeding in Oncidiinae. BMC Plant Biol 10:68