Mitochondrial genomic data are effective at placing mycoheterotrophic lineages in plant phylogeny

Qianshi Lin, Thomas W. A. Braukmann, Marybel Soto Gomez, Juliana Lischka Sampaio Mayer, Fabio Pinheiro, Vincent S. F. T. Merckx , Sasa Stefanovic and Sean W. Graham

DOI:
https://doi.org/10.1111/nph.18335

New Phytologist

Mitochondrial genomic data are effective at placing mycoheterotrophic lineages in plant phylogeny

Qianshi Lin[1,2,3], Thomas W. A. Braukmann[3,4], Marybel Soto Gomez[1,5], Juliana Lischka Sampaio Mayer[6], Fábio Pinheiro[6], Vincent S. F. T. Merckx[7,8], Saša Stefanović[2,3] and Sean W. Graham[1]

[1]Department of Botany, University of British Columbia, 6270 University Boulevard, Vancouver, BC V6T 1Z4, Canada; [2]Department of Biology, University of Toronto Mississauga, Mississauga, ON L5L 1C6, Canada; [3]Ecology and Evolutionary Biology, University of Toronto, Toronto, ON M5S 2Z9, Canada; [4]Department of Pathology, Stanford University, Stanford, CA 94305, USA; [5]Royal Botanic Gardens, Kew, Richmond, Surrey, TW9 3AB, UK; [6]Departamento de Biologia Vegetal, Universidade Estadual de Campinas, 255 Rua Monteiro Lobato, Campinas, São Paulo 13.083-862, Brazil; [7]Naturalis Biodiversity Center, Vondellaan 55, 2332 AA, Leiden, the Netherlands; [8]Department of Evolutionary and Population Biology, Institute for Biodiversity and Ecosystem Dynamics, University of Amsterdam, PO Box 94240,1090 GE, Amsterdam, the Netherlands

Author for correspondence:
Sean W. Graham
Email: swgraham@interchange.ubc.ca

Received: *26 February 2022*
Accepted: *13 June 2022*

New Phytologist (2022) **236:** 1908–1921
doi: 10.1111/nph.18335

Key words: *Afrothismia*, angiosperm phylogenomics, core Thismiaceae, Dioscoreales, Ericales, mitogenomes, phylogenetic congruence, RNA editing.

Summary

- Fully mycoheterotrophic plants can be difficult to place in plant phylogeny due to elevated substitution rates associated with photosynthesis loss. This potentially limits the effectiveness of downstream analyses of mycoheterotrophy that depend on accurate phylogenetic inference. Although mitochondrial genomic data sets are rarely used in plant phylogenetics, theory predicts that they should be resilient to long-branch artefacts, thanks to their generally slow evolution, coupled with limited rate elevation in heterotrophs.
- We examined the utility of mitochondrial genomes for resolving contentious higher-order placements of mycoheterotrophic lineages in two test cases: monocots (focusing on Dioscoreales) and Ericaceae.
- We find Thismiaceae to be distantly related to Burmanniaceae in the monocot order Dioscoreales, conflicting with current classification schemes based on few gene data sets. We confirm that the unusual *Afrothismia* is related to Taccaceae–Thismiaceae, with a corresponding independent loss of photosynthesis. In Ericaceae we recovered the first well supported relationships among its five major lineages: mycoheterotrophic Ericaceae are not monophyletic, as pyroloids are inferred to be sister to core Ericaceae, and monotropoids to arbutoids.
- Genes recovered from mitochondrial genomes collectively resolved previously ambiguous mycoheterotroph higher-order relationships. We propose that mitochondrial genomic data should be considered in standardised gene panels for inferring overall plant phylogeny.

Introduction

Mycoheterotrophic plants (Fig. 1) obtain some or all of their nutrients by effectively cheating on soil fungal partners (Leake, 2004, 2005; Merckx *et al.*, 2009b), reducing or replacing the need to use sunlight to fix carbon (Merckx, 2013) without apparent reward to the fungi (but please refer to Selosse & Roy, 2009; Perez-Lamarque *et al.*, 2020). At least 50 mycoheterotrophic lineages have completely lost photosynthetic function to become fully mycoheterotrophic (FM) (Merckx & Freudenstein, 2010), all marked evolutionary convergences (Merckx *et al.*, 2013; Perez-Lamarque *et al.*, 2020). The majority of these losses occurred in monocots (91% species), with six or seven instances in eudicots, including in Ericaceae (Merckx & Freudenstein, 2010). In monocots, two families (Burmanniaceae, Dioscoreales; Orchidaceae, Asparagales) together represent *c.* 65% of all FM species, and *c.* 75% of all origins (Imhof, 2010; Merckx *et al.*, 2013; Barrett *et al.*, 2019). Our understanding of the phylogenetic placement of FM lineages in angiosperm phylogeny, both broadly and locally, has often been contentious (Dahlgren & Clifford, 1982; Dahlgren & Bremer, 1985; Merckx *et al.*, 2009b; Lam *et al.*, 2016, 2018). For example, earlier classification schemes (Cronquist, 1968; Dahlgren & Bremer, 1985) mistakenly grouped together distantly related mycoheterotroph lineages, due to an overemphasis on their convergent and unusual morphological modifications (i.e. reduced plant stature, loss/reduction of leaves, highly modified underground organs that recover carbon and other nutrients from soil fungi, and sometimes highly modified floral forms; Imhof, 2010; Merckx *et al.*, 2013); molecular data can also be misled by extremely elevated substitution rates in some lineages (Lam *et al.*, 2018; Susko & Roger, 2021). A firmer understanding of the phylogenetic placement of these marked plants would provide refined estimates of their highly convergent evolution. For example, reliable

Fig. 1 Examples of mycoheterotrophic lineages sampled in this study: (a) *Petrosavia stellaris* (Petrosaviaceae, Petrosaviales); (b) *Sciaphila* sp. (Triuridaceae, Pandanales); (c) *Corsia* sp. (Corsiaceae, Liliales); (d) *Gastrodia sesamoides* (Orchidaceae, Asparagales); (e) *Geosiris aphylla* (Iridaceae, Asparagales); (f) *Apteria aphylla* (Burmanniaceae, Dioscoreales); (g) *Afrothismia hydra* (Thismiaceae, Dioscoreales), flower; (h) *Afrothismia foertheriana*, root; (i) *Thismia rodwayi* (Thismiaceae, Dioscoreales), flower; (j) *Thismia rodwayi*, root; (k) *Hypopitys monotropa* (Ericaceae, Ericales); (l) *Pyrola asarifolia* (Ericaceae, Ericales). Photographs are by Vincent S. F. T. Merckx, except (c) by Stephanie Lyon (University of Wisconsin, Stevens Point), (e) by Ehoarn Bidault (Missouri Botanical Garden) and (l) by Qianshi Lin.

phylogenies are needed to better understand the convergence vs divergence of organismal and physiological traits during and after transitions to mycoheterotrophy (Merckx & Freudenstein, 2010; Graham *et al.*, 2017; Jąkalski *et al.*, 2021), the drivers and constraints that may predispose (or prevent) the origin of mycoheterotrophy (Selosse & Roy, 2009; Imhof, 2010), and the degree to which mycoheterotroph evolution and ecology is affected by the ancestry of their fungal partners and vice versa (Perez-Lamarque *et al.*, 2020).

Despite the rapid advance of nuclear phylogenomic methods, plastid (chloroplast) markers still underpin much of plant classification (Angiosperm Phylogeny Group (APG), 2016). Multiple studies have still used them effectively for mycoheterotrophs, including several studies that focused on a handful of plastid genes (Fay *et al.*, 2000; Fuse & Tamura, 2000; Cameron *et al.*, 2003; Lam *et al.*, 2016) to whole plastid genomes (Barrett *et al.*, 2014; Logacheva *et al.*, 2014; Lam *et al.*, 2015, 2018; Mennes *et al.*, 2015; Givnish *et al.*, 2016; Lim *et al.*, 2016; Braukmann *et al.*, 2017; Lallemand *et al.*, 2019). Nonetheless, strong rate elevation in retained plastid genes of heterotrophs (Lam *et al.*, 2016, 2018) – and the resulting very long branches – can strongly interfere with phylogenetic inference, as predicted by Felsenstein (1978) and Hendy & Penny (1989). This is true even with model-based approaches that correct for multiple hits, as demonstrated for several very rapidly evolving heterotrophic lineages (Naumann *et al.*, 2016; Lam *et al.*, 2018). Nuclear genes appear to be less impacted by rate elevation in heterotrophic lineages compared with plastid genes (Wolfe *et al.*, 1987; Lemaire *et al.*, 2011; Bromham *et al.*, 2013), and have recently been used to place several heterotrophic lineages in large-scale studies using target-sequence capture or transcriptome-based approaches (One Thousand Plant Transcriptomes Initiative, 2019; Baker *et al.*, 2021). Nuclear studies can take advantage of multispecies coalescent (MSC) approaches (e.g. ASTRAL-III; Zhang *et al.*, 2018) that are not valid for plastid or mitochondrial genomes (as each organellar genome behaves effectively as a single, highly informative 'coalescent gene'; Doyle, 2022). However, organellar data sets can still play a useful role in plant phylogenetics as a data-rich counterpoint to MSC approaches, and the comparison between nuclear MSC and organellar results enrich our understanding of both (Baker *et al.*, 2021; Doyle, 2022).

Mitochondrial genomes have largely been ignored in inferring plant phylogeny, in part due to the very low substitutional rates in most lineages (Knoop, 2004; Yurina & Odintsowa, 2016), which has led to a general perception that they evolve too slowly to be useful (e.g. Fazekas *et al.*, 2009, but please refer to Qiu

et al., 2006). However, theory, simulation and empirical investigations (Kim, 1996; Swofford *et al.*, 1996; Brinkmann *et al.*, 2005; Klopfstein *et al.*, 2017) support the idea that slower (and therefore less saturated) genes should be less prone to the types of long-branch artefacts observed in many heterotrophic lineages for plastid data. While one or a few mitochondrial regions have been deployed effectively to address the placements of several mycoheterotroph lineages (Merckx *et al.*, 2009a; Freudenstein *et al.*, 2016), analysis of mitochondrial genome-scale data sets also show promise for placing individual heterotrophic lineages (Bell *et al.*, 2020; Soto Gomez *et al.*, 2020; Jost *et al.*, 2021). Li *et al.* (2019) recently posited that incongruence between plastid and mitochondrial phylogenomic data sets in Orchidaceae may at least in part be due to elevated rates in the plastid organellar genome for FM orchids (please refer also to Lam *et al.*, 2018). Nonetheless, the broad utility of mitochondrial genomes in estimating plant phylogeny – and in placing problematic heterotrophic lineages in particular – remains to be tested more generally.

Fortunately, it has become relatively straightforward to assemble complete (or nearly complete) mitochondrial gene sets and apply them in plant phylogenomic studies. Here we use the protein-coding complement of mitochondrial genomes to examine their utility in combined phylogenomic analyses. In seed plants this generally comprises genes for subunits of the respiratory chain complexes I (*nad* genes), II (*sdh*), III (*cob*), IV (*cox*) and V (*atp*), subunits of a cytochrome *c* maturation pathway (*ccm*), ribosomal proteins (*rpl* and *rps*), and a subunit of a twin-arginine translocase (*tat*C), representing *c.* 60 genes in total, including three rDNA genes and *c.* 20 tRNA genes, and ignoring nonfunctional transfers of genes from other genomes or even species (Stern & Palmer, 1986; Bergthorsson *et al.*, 2004). In most photosynthetic angiosperms, gene number can vary between different taxa, by up to six genes (Knoop, 2004; Li *et al.*, 2009). However, heterotrophs do not appear to experience substantial gene loss in their mitochondrial genomes, in general, a strong contrast with plastid genomes (Petersen *et al.*, 2015, 2019; Bell *et al.*, 2020; Soto Gomez *et al.*, 2020).

Here we recovered and applied mitochondrial phylogenomic data for two major clades, the monocots and Ericaceae, as test cases to demonstrate how well this allows us to deal with contentious placement of mycoheterotrophs (Merckx & Freudenstein, 2010). For monocots, we considered a selection of taxa comprising all five orders that include FM taxa (Fig. 1a–j), but with a particular focus on the yam order Dioscoreales, which has experienced multiple losses of photosynthesis, and whose plastid genomes include several extremely rapidly evolving lineages (Lam *et al.*, 2016, 2018; Givnish *et al.*, 2018). In this order, Burmanniaceae include both FM and photosynthetic taxa, while Thismiaceae are entirely composed of FM species. Burmanniaceae are estimated to have diverged from other Dioscoreales *c.* 116 million years ago (Ma) (Merckx *et al.*, 2008), but the estimated stem age of Thismiaceae is highly variable, ranging from *c.* 86 (61–106) Ma in Merckx & Smets (2014) to only *c.* 12 Ma in Givnish *et al.* (2018), who also include age estimates for other monocot mycoheterotrophic lineages. Discrepancies in age estimates for Thismiaceae may reflect the general difficulty of including rate-elevated heterotrophic taxa in dating analyses (Iles *et al.*, 2015).

Recent angiosperm classification schemes (APG, 2003, 2009, 2016) combine Burmanniaceae and Thismiaceae, despite phylogenetic evidence from several mitochondrial, nuclear and plastid regions that the two lineages are not closely related in the order (Merckx *et al.*, 2009a; Lam *et al.*, 2016, 2018; summarised in Fig. 2). The decision to combine Burmanniaceae and Thismiaceae in the most recent versions of the APG classification scheme (APG, 2009, 2016) appears to have been based primarily on evidence from Caddick *et al.* (2002; Fig. 2). In addition, although they recognised the need to review the incorporation of Thismiaceae in Burmanniaceae, APG (2016) did not account for the existence of problematic sequences in Caddick *et al.* (2002) noted in Lam *et al.* (2016). By contrast, the distinctness of Thismiaceae from Burmanniaceae was clearly supported by Merckx *et al.* (2009a) based on likelihood analyses of the mitochondrial *atpA* and nuclear 18S rDNA loci, in a similar analysis in Merckx & Smets (2014), and in recent plastid-based phylogenomic studies (Givnish *et al.*, 2018; Lam *et al.*, 2018). Merckx *et al.* (2009a) inferred that FM *Afrothismia*, currently classified in Thismiaceae, may instead be the sister group of a clade comprising both Taccaceae and Thismiaceae based on likelihood analyses (Fig. 2; maximum likelihood (ML) summary), which would be consistent with a further independent loss of photosynthesis (Fig. 1g, h). However, they also found that the placement of *Afrothismia* was sensitive to the inference method and gene(s) examined. *Afrothismia* has also not been included in any phylogenomic analyses to date, which we remedy here.

In eudicots, the heather family Ericaceae provides another example of partly unresolved placements of mycoheterotrophic lineages. In the most current classification based on broad-scale molecular and morphological analyses it is divided into eight subfamilies (please refer to Kron *et al.*, 2002 and references there), seven of which are exclusively autotrophic (i.e. Enkianthoideae, Arbutoideae, Cassiopoideae, Ericoideae, Harrimanelloideae, Styphelioideae and Vaccinioideae). All mycoheterotrophic Ericaceae are currently classified under subfamily Monotropoideae, with all partial/initial heterotrophic species in tribe Pyroleae (which includes only one FM species, *Pyrola aphylla*), and FM species in tribes Pterosporeae and Monotropeae (Fig. 1k,l; Kron *et al.*, 2002; Tedersoo *et al.*, 2007). Several studies based on a few plastid, nuclear and/or mitochondrial genes have suggested that this mycoheterotrophic subfamily is polyphyletic, but with only low to moderate bootstrap support (< 85%) for conflicting clades (Fig. 3; Kron *et al.*, 2002; Braukmann & Stefanovic, 2012; Freudenstein *et al.*, 2016; Lallemand *et al.*, 2016), although photosynthetic arbutoids are well supported as sister to mycoheterotrophic monotropoids in Freudenstein *et al.* (2016) and Lallemand *et al.* (2016). However, overall the relationships among major groups of Ericaceae – that is, the core Ericaceae (also referred to as the 'inverted anther' clade; Kron *et al.*, 2002), pyroloids, monotropoids, arbutoids and *Enkianthus*, and especially the placement of pyroloids – have not been completely resolved.

Our main hypothesis for both test cases is that mitochondrial phylogenomic data can effectively resolve the relationships of

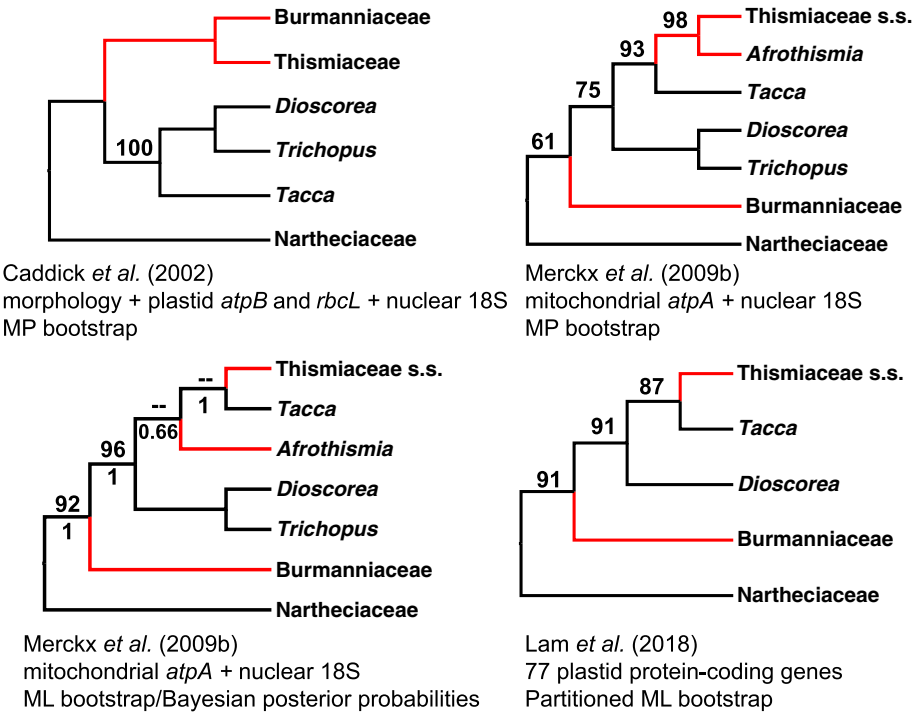

Fig. 2 Schematic summary of previously published phylogenetic hypotheses concerning mycoheterotrophic Dioscoreales. Bootstrap support values and posterior probabilities are noted when they are at least 50% or 0.5, respectively (double dashes or missing values on some branches are below these cut-offs). Red lineages represent mycoheterotrophic lineages in Dioscoreales. (Burmanniaceae include autotrophic, partial and fully mycoheterotrophic taxa.)

hard-to-place mycoheterotrophic lineages. In monocots, our major focus is on the yam order Dioscoreales, to address in particular the uncertain relationships and boundaries of Burmanniaceae and Thismiaceae, and the placement of *Afrothismia*. In Ericaceae, we surveyed all five major groups, focusing on mycoheterotrophic Ericaceae. In both cases we examine the degree to which these mitochondrial data sets permit inference of well supported placements of mycoheterotrophic lineages. We also report on whether the resulting placements align with inferences made from other recent analyses. Finally, we make recommendations for current large-scale efforts to infer plant relationships using phylogenomic approaches.

Materials and Methods

Taxon sampling

For monocots, we generated partial mitochondrial genome sequences for 47 taxa, including 18 FM taxa and two partial mycoheterotrophs (also known as mixotrophs; Selosse & Roy, 2009). We sampled representatives from all five orders and seven monocot families that include mycoheterotrophic taxa. We also included representatives of two of four orders of commelinid monocots recognised in APG (2016), Arecales and Poales, excluded the monogeneric order Acorales (which we lacked a full mitochondrial gene set for), and retrieved sequences of *Arabidopsis thaliana* (L.) Heynh. and another six monocot taxa from GenBank (Supporting Information Table S1). The final 54-taxon mitochondrial matrix represents eight monocot orders and a eudicot outgroup. For Ericaceae, we generated partial mitochondrial genome sequences for eight taxa, including two FM species (*Allotropa virgata* Torr. & A. Gray and *Pterospora andromedea* Nutt.) and three partial/initial mycoheterotrophic species (*Chimaphila umbellata* (L.) Nutt., *Moneses uniflora* A. Gray and *Pyrola minor* L.). In addition, we retrieved sequences of *Hypopitys monotropa* Crantz (an FM species) and five other taxa from GenBank (*Actinidia chinensis* Planch., *Arbutus unedo* L., *Camellia sinensis* (L.) Kuntzes, *Rhododendron simsii* Planch and *Vaccinium macrocarpon* Aiton; Table S1). We included *A. chinensis* (Actinidiaceae) and *C. sinensis* (Theaceae), which are in the same order (Ericales) with Ericaceae, as outgroups. The resulting 14-taxon mitochondrial matrix represents all five major clades of Ericaceae (core Ericaceae, pyroloids, monotropoids, arbutoids, and *Enkianthus*) and multiple outgroups (Table S1). Our sampling mainly focuses on mycoheterotrophic Ericaceae, because three major green clades – the core Ericaceae, arbutoids, and *Enkianthus* – have repeatedly been shown to be monophyletic with strong support in recent studies (Braukmann & Stefanovic, 2012; Freudenstein *et al.*, 2016; Lallemand *et al.*, 2016). We sampled all three tribes of mycoheterotrophic Ericaceae (Monotropeae, Pterosporeae and Pyroleae). Taken together, our sampling strategy covered the problematic groups from previous phylogenetic studies of Ericaceae and should be sufficient to address the higher-order relationships in this family.

Sequencing, contig assembly, sequence alignment and data matrix construction

We used a range of approaches to prepare and sequence libraries for multiple taxa in monocots and Ericaceae (Methods S1), and freshly extracted mitochondrial genes from these and previously sequenced libraries (Table S1). We did not attempt to assemble

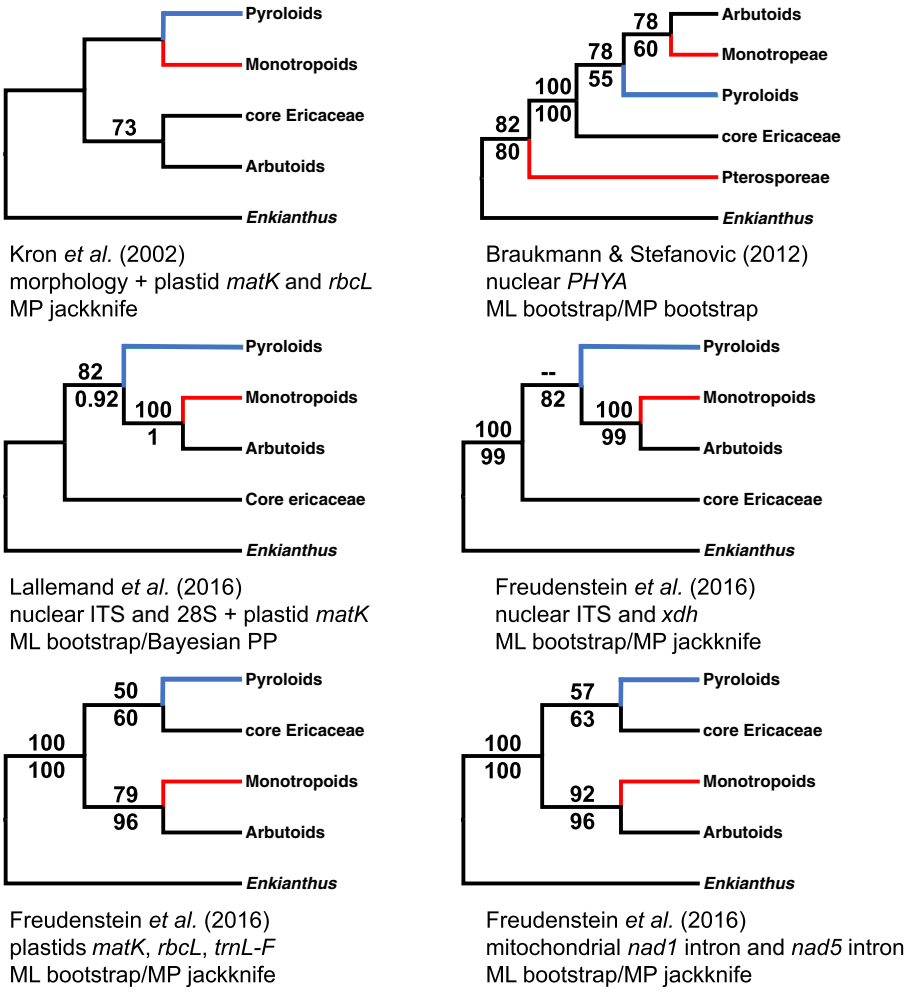

Fig. 3 Schematic summary of previously published phylogenetic hypotheses concerning Ericaceae. The 'core Ericaceae' (*sensu* Kron *et al*., 2002) represents the clade of five subfamilies (Cassiopoideae, Ericoideae, Harrimanelloideae, Styphelioideae and Vaccinioideae) with early anther inversion and typical ericoid mycorrhizas. Bootstrap support values and posterior probabilities are noted when they are at least 50% or 0.5, respectively (double dashes or missing values on some branches are below these cut-offs). Red lineages represent fully mycoheterotrophic monotropoids; blue lineages represent partially/initially mycoheterotrophic pyroloids. ITS, internal transcribed spacer.

complete mitochondrial genomes, as gene sequences evolve very slowly in most plant mitochondrial genomes, whereas intragenomic recombination frequently makes relative gene order quite variable (Wolfe *et al*., 1987; Gualberto *et al*., 2014). For monocots, we assembled mitochondrial contigs following methods used for plastid data in Lam *et al*. (2015), except that we included contigs with an average of > 10× coverage. A partial genome assembly of *Pogoniopsis schenckii* Cogn. was obtained using SPAdes (Bankevich *et al*., 2012), with the standard parameters suggested in the manual. For Ericaceae, we assembled mitochondrial contigs using GetOrganelle v.1.6.2.e (Jin *et al*., 2020), a pipeline using Bowtie2 v.2.3.5.1 (Langmead & Salzberg, 2012) for recruitment of initial target-associated reads, and SPAdes for *de novo* assembly of mitochondrial contigs. For most Ericaceae species we did this using a reference dataset of plant mitochondrial genomes (embplantmt) with 50 rounds of contig extensions and k-mer sizes of 21, 45, 65, 85 and 99. Most species produced assemblies with most expected genes recovered (> 90% of all genes), except *Allotropa*, *Arctostaphylos* and *Pterospora*. We improved the *Arctostaphylos* assembly by using *Pyrola* as an additional reference sequence, and used a reference database of all other assembled Ericaceae species to improve the *Allotropa* and *Pterospora* assemblies.

We retrieved protein-coding mitochondrial genes using the Blastn program from Blast+ NCBI (version NCBI-BLAST-2.2.30+, Camacho *et al*., 2009), using *A. thaliana* (NC_037304) and *Oryza sativa* (JF_281153) as queries for monocots, and *H. monotropa* (MK990822 and MK990823) as a query for Ericaceae. We recovered data for the 37 protein-coding genes common to monocots, and the 38 protein-coding genes common to Ericales. We set up individual gene files for these, each with 54 taxa for monocots, or 14 taxa for Ericaceae. For both monocots and Ericaceae, we separately constructed concatenated data matrices following Lam *et al*. (2015), with the following modifications. We conducted initial alignment of individual mitochondrial genes using the Muscle (Edgar, 2004a,b) online portal using default settings and concatenated gene files manually. We then used Mesquite v.3.15 (Maddison & Maddison, 2018) to perform manual adjustments of the matrix, following Graham *et al*. (2000). The final alignment of monocots is a 36 807-bp matrix for the concatenated protein-coding genes (for reference, derived from 21 143 bp of unaligned data in *Acanthochlamys bracteata* P.C. Kao). The final alignment of Ericaceae is a 33 801-bp matrix for the concatenated protein-coding genes (for reference, derived from 31 657 bp of unaligned data in *Enkianthus campanulatus* (Miq.) G. Nicholson). We also translated both

alignments into corresponding amino-acid matrices. We deposited new sequence data (Table S1) to GenBank (MN907133–MN907168, MT023139–MT055781 for monocots; MZ593998–MZ594455 for Ericaceae), and alignments on Figshare (https://doi.org/10.6084/m9.figshare.14879796.v2).

Likelihood and parsimony-based analyses

We performed parsimony analysis on DNA data using PAUP* v.4.0a build 164 (Swofford, 2002). We ran a heuristic search for the shortest trees in monocots using tree bisection–reconnection (TBR) branch swapping with 10 random stepwise addition replicates, holding one tree at each step, and otherwise using default settings. Given the moderate number of terminal units in the Ericaceae data set, we conducted a branch-and-bound search for it, ensuring recovery of all the most parsimonious trees. State transitions were treated as unordered, and all sites were equally weighted. We estimated branch support using 500 bootstrap replicates (Felsenstein, 1985), each with 100 random stepwise addition replicates, and otherwise using default settings.

We also performed ML analyses on both data sets using RAxML v.7.4.2 (Stamatakis, 2006) with a graphical interface (Silvestro & Michalak, 2012), considering separate unpartitioned and partitioned likelihood analysis of DNA data, and unpartitioned likelihood analysis of AA data. In both cases, we determined the partitioning schemes as follows. For DNA-based analyses of the concatenated mitochondrial DNA matrix, we partitioned sites using a gene-by-codon ('G × C') partitioning scheme, starting with initial individual partitions derived from first, second, and third codon positions of each gene (111 initial partitions for the monocot DNA matrix, and 114 initial partitions for the Ericaceae DNA matrix); the corresponding AA analysis started with one partition for each gene (Tables S2, S3). In both cases, we combined partitions that did not have significantly different substitution models using PARTITIONFINDER2 (Lanfear *et al.*, 2016) with the relaxed hierarchical clustering algorithm (r-clustering) and the corrected Akaike Information Criterion (AIC_c), limiting the DNA or AA substitution models under consideration to those implemented in RAxML v.7.4.2 (please refer to Table S4 for final partitions and models). The final partitioning schemes (Table S4) include 63 partitions for the DNA matrix and 24 amino-acid partitions for monocots, as well as 47 partitions for the DNA matrix and 18 amino-acid partitions for Ericaceae. The GTR+G or GTR+I+G DNA substitution models were inferred to be the optimal fit for partitions in the DNA version of both matrices; we used the GTR+G model for all partitions because the 'I' parameter for invariant sites may be well accommodated by the gamma-distribution shape parameter 'alpha' (Yang, 2006). Several optimal substitution models were inferred as the best fit for final AA partitions (with JTT+G+F as the most commonly inferred model for monocots, and STMTREV+F for Ericaceae; Table S4), which we used in analysis. For all analyses, we ran 20 independent searches for the best tree and estimated branch support using 500 bootstrap replicates (Felsenstein, 1985). We considered highly supported branches to have ≥ 95% bootstrap support and poorly supported branches to have < 70% support, following Soltis & Soltis (2003).

Other analyses

We assessed the possible effect of RNA edit sites on phylogenetic inference using approaches to identify and remove these sites across taxa (Methods S2), before re-analysing the DNA data using unpartitioned ML analysis. We compared GC content and codon usage fractions of heterotrophic vs green lineages, and characterised relative rate differences in them using BEAST v.2.6.7 (Bouckaert *et al.*, 2019) (computationally tractable for all genes in Ericaceae, and for a common subset of genes in monocots; Methods S3). Finally, we performed topological constraint tests of the monophyly of clades of interest using the approximately unbiased (AU) (Shimodaira, 2002) and Shimodaira-Hasegawa (SH) tests in CONSEL (Shimodaira & Hasegawa, 2001) (Methods S4).

Results

Gene recovery

Our mitochondrial gene sets include all protein-coding genes found in *A. thaliana* and *O. sativa* for monocots, and *H. monotropa* for Ericaceae. We recovered most mitochondrial genes from most taxa (please refer to Tables S2, S3 for data recovery), without evidence of large-scale gene loss compared with a Southern hybridisation survey of mitochondrial gene content in monocots and Ericales (Adams *et al.*, 2002). Large-scale gene loss was found only in Alismatales in our results, an order known to have extensive loss of mitochondrial genes encoding ribosomal proteins (Adams *et al.*, 2002; Petersen *et al.*, 2017). The unrecovered genes were coded as missing data. These scattered instances of 'gene absence' (Tables S2, S3) may reflect either occasional gene loss, or failure of gene recovery due to the relatively low coverage of mitochondrial data for some taxa, discussed in more detail below.

Mitochondrial phylogenomic inference of overall monocot relationships

We found no significant difference between the GC content of mycoheterotrophic and green plants (Table S5). There were significant differences in codon usage values for a subset of codons (15 of 64; Table S5), but the differences were all very small (of the order of 2% or less, Table S5). Relationships inferred in the four likelihood analyses of the concatenated mitochondrial data are consistent with each other, and are also well supported by bootstrap analysis, with no strong conflicts (Figs 4, S1). Our DNA-based analyses have only four branches with < 95% bootstrap support: (i) the local placement of the FM family Triuridaceae within Pandanales, which is moderately well supported as the sister group of Cyclanthaceae–Pandanaceae; (ii) relationships at the base of Burmanniaceae (*Apteria*–*Gymnosiphon* recovered as

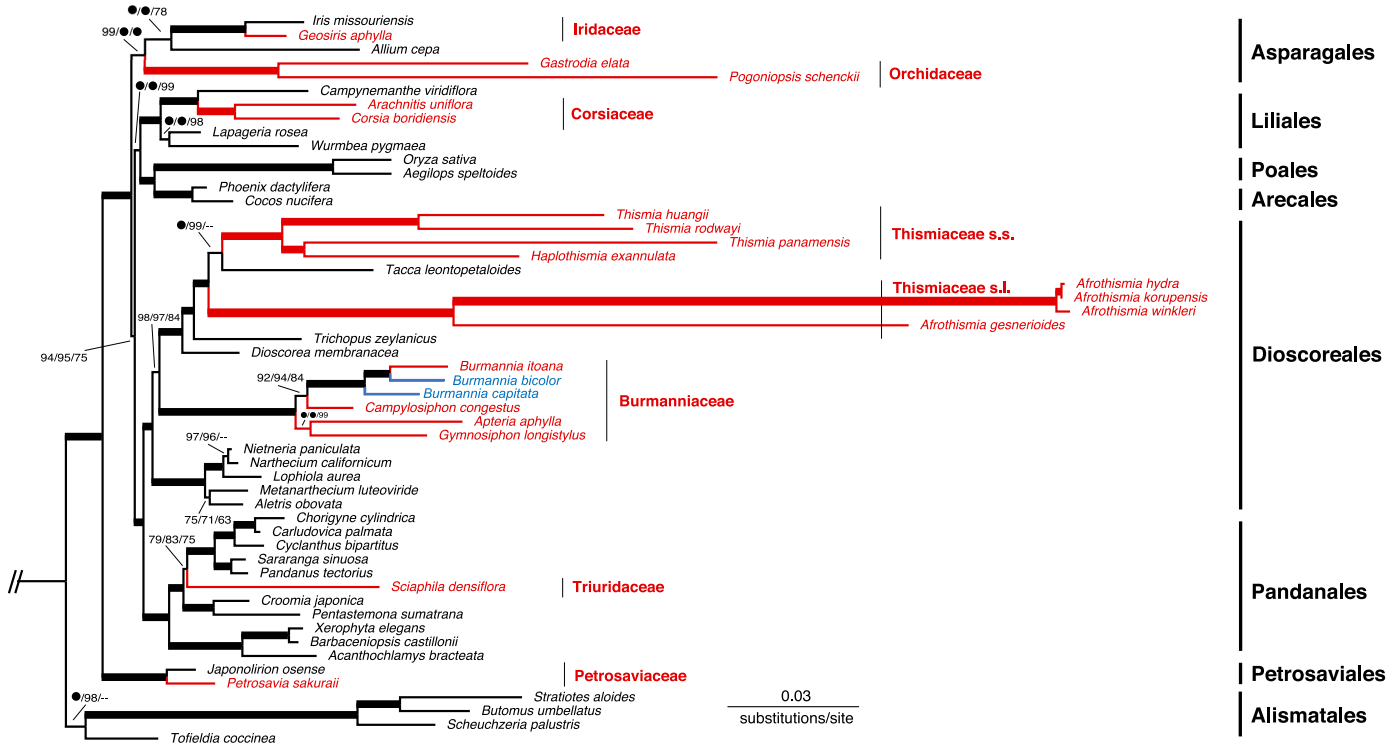

Fig. 4 Phylogenetic placement of mycoheterotrophic monocots inferred from the partitioned maximum likelihood analysis of 37 mitochondrial genes using a 'G × C' partitioning scheme of DNA sequence data (please refer to the Materials and Methods section). Thick lines indicate 100% bootstrap support from both partitioned and unpartitioned likelihood analyses, and parsimony analysis; bootstrap values < 100% are indicated beside branches (partitioned likelihood/unpartitioned likelihood/parsimony, respectively; filled circle = 100%, double dashes < 50%). Red lineages represent fully mycoheterotrophic taxa; terminals with blue labels are suspected or known partial mycoheterotrophs. Families with mycoheterotrophic species are also noted. *Arabidopsis thaliana* (an outgroup) was included but is not shown. Bar indicates estimated substitutions per site.

the sister of the remainder of the family, with moderate support); and (iii) relationships at the base of Nartheciaceae, with *Aletri–Metanarthecium* recovered as the sister of the remainder of the family, again with moderate support. An identical monocot tree topology was inferred when predicted RNA edit sites (Table S6) were removed in the unpartitioned ML analysis of DNA sequence data (only six branches have changes in bootstrap support values of > 10%; Fig. S2). Relationships inferred in the DNA parsimony analysis and two AA likelihood analyses are also generally well supported, but have lower support for several branches (Fig. S1). All eight sampled monocot orders are well supported as monophyletic in the main DNA-based likelihood analyses (Figs 4, S1); the AA analyses recover slightly less support for the monophyly of each of two orders (Alismatales and Dioscoreales), and poor support for Asparagales monophyly (Fig. S1). Asparagales is placed as the sister group to all other monocots in addition to Petrosaviales and Alismatales, and Liliales as the sister group of commelinid monocots. Both arrangements find strong support in the DNA likelihood analyses (Figs 4, S1), and weaker support in the AA and parsimony analyses (Fig. S1). A sister-group relationship between Dioscoreales and Pandanales is also strongly supported across all four likelihood analyses (Figs 4, S1). Petrosaviales are inferred to be the sister group of all other sampled monocots here except Alismatales, with strong support for this relationship in the DNA analyses (Figs 4, S1).

Placement of mycoheterotroph lineages in monocot phylogeny

All seven monocot families containing FM taxa are well supported as monophyletic here, apart from Thismiaceae. This family is divided into two clades with strong support: the FM genus *Afrothismia* (classified in Thismiaceae, Schlechter, 1921; Jonker, 1938; Maas *et al.*, 1986; Maas-van de Kamer, 1998) is inferred to be the sister group of photosynthetic Taccaceae, and a clade comprising other sampled taxa of Thismiaceae (marked as Thismiaceae s.s. in figures), with strong support for this relationship in the DNA ML analyses (Fig. 4). *Trichopus* (the sole genus of Trichopodaceae) is inferred to be the sister group of *Afrothismia*, Thismiaceae s.s. and Taccaceae. *Dioscorea* (Dioscoreaceae) is the next successive sister group of these taxa. These two relationships have strong support across all analyses (Figs 4, S1). Within Thismiaceae, *Haplothismia* is inferred to be the sister group of *T. panamensis* with strong support, also rendering *Thismia* nonmonophyletic (Figs 4, S1). The other lineage with FM taxa in Dioscoreales, Burmanniaceae, includes autotrophic taxa (e.g. two sampled species in *Burmannia* marked in blue in Fig. 4) and multiple lineages of FM taxa (including one sampled FM species of *Burmannia*, *B. itoana*). Burmanniaceae are recovered as the sister group of all taxa in Dioscoreales except Nartheciaceae, a relationship recovered with strong support in DNA ML analyses (Fig. 4). All relationships in Burmanniaceae, including the relative

placements of two autotrophic taxa (*B. bicolor* and *B. capitata*), are also well supported (Figs 4, S1), except for the arrangement around the deepest split in the family, as noted above. We were able to reject the monophyly of (i) a clade comprising Burmanniaceae and Thismiaceae (either including/excluding *Afrothismia* from the analysis) and (ii) a clade comprising Thismiaceae s.s. and *Afrothismia*, based on AU and SH tests ($P < 0.001$ and $P < 0.015$, respectively).

Outside Dioscoreales, FM Triuridaceae, represented here by *Sciaphila*, are inferred to be the sister group of Cyclanthaceae–Pandanaceae, a large photosynthetic clade in Pandanales. The two families of Asparagales that include FM taxa, Orchidaceae and Iridaceae, are grouped together with strong support. The FM genus *Geosiris* is recovered as sister to the other taxon sampled in Iridaceae (*Iris*) with strong support (Figs 4, S1). Corsiaceae, the only FM family in Liliales, is strongly supported as the sister group of Campynemataceae among sampled taxa, an arrangement that is also strongly supported across all analyses here (Fig. S1). Within Petrosaviales, *Petrosavia* (an FM genus) is inferred to be the sister group of autotrophic *Japonolirion*, the only other genus in its family/order; there is strong support for this arrangement in all analyses (Figs 4, S1).

Mitochondrial phylogenomic inference of Ericaceae

The Ericaceae relationships inferred by all five likelihood and parsimony analyses of the concatenated mitochondrial data are topologically identical with each other, and are also moderately to well supported (> 85%) by bootstrap analyses (Figs 5, S3). The inferred topologies are identical (presented for the partitioned ML analysis in Fig. 5). All branches in our DNA-based analyses (likelihood and parsimony) have 100% bootstrap support. Removal of predicted RNA edit sites (Table S7) in the unpartitioned likelihood analysis of DNA sequence data did not substantially affect inference of Ericaceae relationships (identical topologies; only small differences in bootstrap support using different methods for predicting or identifying RNA edit sites; please refer to Figs S4, S5; Methods S2). Relationships inferred from two AA analyses are also generally well supported and topologically identical to the DNA-based analyses, but have slightly reduced support (still > 85%) for five branches (Fig. S3). In all analyses, *Enkianthus* is found as the sister to the rest of Ericaceae. We rejected the monophyly of a clade comprising the mycoheterotrophic Ericaceae (i.e. pyroloids and monotropoids) based on the AU and SH tests ($P < 0.001$). The monotropoids are inferred to be the sister group of arbutoids, and the pyroloids the sister group of the core Ericaceae, with the former clade sister to the latter, all with strong support. Within pyroloids, *Pyrola* is inferred to be the sister group of *Chimaphila* and *Moneses* with strong support. Within monotropoids, *Pterospora* is inferred to be the sister group of *Allotropa* and *Hypopitys* with strong support (Fig. 5).

Rate elevation in mycoheterotrophic lineages

Modest rate elevation is evident in a subset of mycoheterotrophic lineages (Figs S6, S7). In monocots, mycoheterotrophic *Pogoniopsis*, Thismiaceae s.s., green *Chorigyne* and two Alismatales have intermediate rates (Fig. S6); only *Afrothismia* had faster rates, *c.* four-fold higher then green relatives (Fig. S6). In Ericaceae, all mycoheterotrophs have low rates, with the highest rates observed in *Moneses* (Fig. S7) (identified as a partial mycoheterotroph in Hynson *et al.*, 2015, but please refer to Lallemand *et al.*, 2016).

Discussion

Mitochondrial genomes of mycoheterotrophic plants

The gene content we recovered from monocots and Ericaceae differs only for *sdh* and ribosomal protein genes (Tables S2, S3), gene classes that frequently vary between different plant lineages (Adams *et al.*, 2002). We did not find clear evidence of gene loss in Ericaceae (Table S3). Within monocots, unrecovered genes are consistent with previously documented extensive gene loss in Alismatales (Petersen *et al.*, 2017), or are likely to relate to relatively low coverage of mitochondrial data for some taxa (for example *Lophiola*, which has an average coverage between *c.* 10–50×). Several ribosomal proteins genes, for example *rpl*10 and *rps*13, are frequently absent from different taxa (please refer to also Adams *et al.*, 2002). Four *Afrothismia* species have some common missing genes (for example *atp*8, *rpl*5) that may represent actual gene loss, although this should be confirmed by completing whole mitochondrial genomes for these taxa. However, in general for mycoheterotrophic plants, we did not find clear evidence of any extensive mitochondrial gene losses comparable to Alismatales and parasitic Viscaceae (Petersen *et al.*, 2015, 2017; Skippington *et al.*, 2015, 2017). We also found no significant differences in GC content or codon usage between mycoheterotrophic and green plants (Table S5). Therefore, unlike their plastid counterparts, mitochondrial genes appear to be largely unaffected by mycoheterotrophy.

Performance of whole mitochondrial gene data in phylogenetic inference

When extreme, fast substitution rates in plastid genomes of heterotrophic lineages (Naumann *et al.*, 2016; Lam *et al.*, 2018; Susko & Roger, 2021) can lead to strong long-branch attraction (Felsenstein, 1978; Hendy & Penny, 1989; Bergsten, 2005). Mitochondrial genes of heterotrophic taxa can experience limited rate elevation (Bromham *et al.*, 2013; Petersen *et al.*, 2015), apparent to a degree here (e.g. *Pogoniopsis* and *Afrothismia* in Fig. S6). However, the plastids of the fastest evolving mycoheterotrophic lineages evolve hundreds of times faster than those of their photosynthetic relatives (e.g. for *Thismia* vs *Tacca* in Givnish *et al.*, 2018), and *Afrothismia* plastid genes appear to evolve even more rapidly (Soto Gomez, 2020). By contrast, the overall rates of evolution observed in mitochondrial genes here (Figs 4, 5; note scale) are within optimal ranges recommended for phylogenetic analysis (i.e. within *c.* 0.1–0.5 substitutions per site from root to tip, Klopfstein *et al.*, 2017), even for *Afrothismia* and *Thismia*, which are among the most rapidly evolving lineages

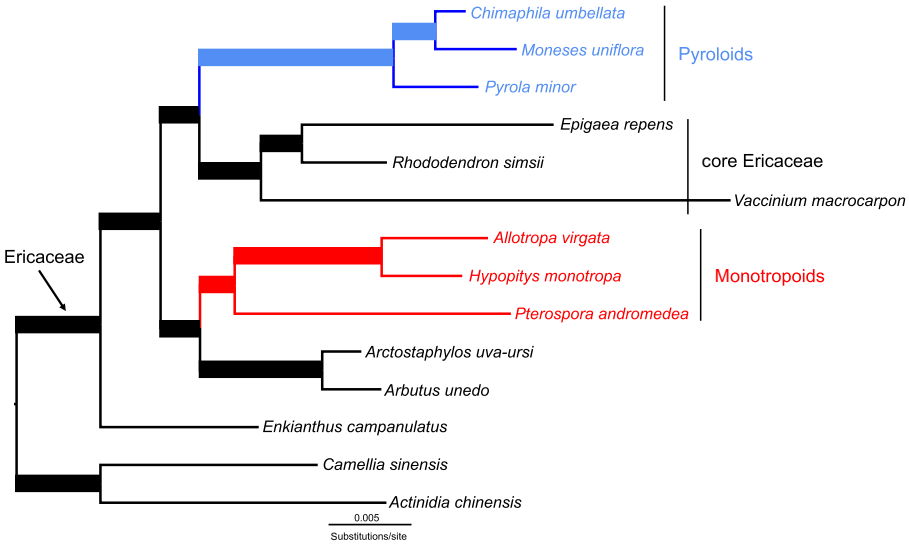

Fig. 5 Phylogenetic placement of mycoheterotrophic Ericaceae inferred from the partitioned maximum likelihood analysis of 38 mitochondrial genes using a 'G × C' partitioning scheme of DNA sequence data (please refer to the Materials and Methods section). Thick lines indicate 100% bootstrap support from both partitioned and unpartitioned likelihood analyses, and parsimony analysis. Red lineages represent fully mycoheterotrophic taxa; blue lineages represent partial/initial mycoheterotrophs. Five major clades within Ericaceae are also noted. Bar indicates estimated substitutions per site.

sampled here (the mitochondrial genomes of these two lineages evolve only *c.* two to four times faster than those of related photosynthetic lineages; Fig. S6).

Our DNA- and AA-based analyses of mitochondrial data agreed on overall tree topology within the monocot and Ericaceae data sets, and are also generally well supported (Figs 4, 5). The sometimes weaker support in AA analyses (Figs S1, S3) may reflect the fewer informative sites in amino-acid data compared with the corresponding untranslated DNA data, in addition to the generally very low substitution rates of individual mitochondrial genes (Knoop, 2004; Yurina & Odintsowa, 2016). Different likelihood models (partitioned, unpartitioned) for DNA and AA data also appear to have little impact on inferences of tree topology or branch support (Figs 4, 5, S1, S3), and we also observed little difference in parsimony vs likelihood inferences (Figs 4, 5); when observed, this type of conflict can be a hallmark of long-branch attraction (Lam *et al.*, 2018).

All but three sampled branches across monocots have strong support (at least 95%) in our DNA-based likelihood analyses (Fig. 4), and all branches within Ericaceae also have strong or maximal support across DNA analyses (Fig. 5). The order-level relationships inferred from our monocot mitochondrial genomic data are also consistent with studies based on plastid genome data (Lam *et al.*, 2015, 2018; Givnish *et al.*, 2016, 2018), except for the placement of Asparagales. This order is strongly supported to be the sister group of commelinids based on plastid phylogenomic data (Givnish *et al.*, 2016, 2018; Lam *et al.*, 2018). This major discrepancy may be caused by insufficient taxon sampling here in Asparagales and commelinids, two of the largest clades of monocots; future studies should focus on more thorough taxon sampling in these two clades, especially for the largest mycoheterotrophic lineage, Orchidaceae.

Placing mycoheterotrophic lineages in Dioscoreales with confidence

The order Dioscoreales includes two families with FM taxa, Burmanniaceae (Fig. 1f) and Thismiaceae (Fig. 1g–j). Our analyses strongly support that these two families are distantly related lineages in Dioscoreales, confirming more weakly supported results in recent plastid genome-based analyses of Givnish *et al.* (2018) and Lam *et al.* (2018). The clear distinction of Thismiaceae from Burmanniaceae was also moderately to strongly supported in the individual or few gene mitochondrial/nuclear analyses of Merckx *et al.* (2009a), summarised for the combined case in Fig. 2 here. Consistent with this, our AU and SH tests strongly reject the monophyly of a clade comprising Burmanniaceae and Thismiaceae, and underline the need to update the APG (2016) classification (please refer also to Lam *et al.*, 2018). Our analyses of mitochondrial data also recover the FM genus *Afrothismia* (Fig. 1g,h) as the sister group of a clade comprising FM Thismiaceae (s.s.) and photosynthetic Taccaceae (Fig. 4). This arrangement is consistent with a subset of analyses of between 1–6 mitochondria and nuclear loci by Merckx *et al.* (2009a) (for example, please refer to likelihood summary here in Fig. 2), Merckx & Smets (2014) and Merckx *et al.* (2017, their appendix S1.4). It is strongly supported here in likelihood analysis for the first time. The branch comprising Taccaceae and Thismiaceae is poorly supported in parsimony analysis but has strong support from likelihood analyses (Fig. 4), and AU and SH tests reject a placement of *Afrothismia* with Thismiaceae s.s. to the exclusion of Taccaceae. Our result is consistent with an independent loss of photosynthesis in Dioscoreales for *Afrothismia*, and with possible recognition of the genus as a distinct family, depending on how broadly family boundaries are defined in future Dioscoreales classification schemes. Characteristics distinguishing *Afrothismia* from other Thismiaceae include the underground organ of the former, which consists of clusters of bulbils, whereas other Thismiaceae have coralloid, vermiform or tuberous roots (Fig. 1g–j; Cheek, 2003; Merckx *et al.*, 2009a).

Within the Thismiaceae s.s. clade, the South American species *T. panamensis* is sister to the Indian monotypic genus *Haplothismia* with strong support in all analyses (Figs 4, S2), disrupting the monophyly of *Thismia*, consistent with Shepeleva *et al.* (2020). Two unsampled genera in Thismiaceae here (*Oxygyne* and *Tiputinia*) belong in the main Thismiaceae clade along

with *Thismia* and *Haplothismia*, based on analyses of one to three genes (mitochondrial, nuclear) (Yokoyama *et al.*, 2008; Merckx *et al.*, 2009a; Shepeleva *et al.*, 2020), which should also be addressed in the future with mitochondrial phylogenomic data. Further studies of Thismiaceae, including its broader evolution and biogeography, await more broadly sampled phylogenies of the family (Shepeleva *et al.*, 2020).

Higher-order phylogeny of Ericaceae

With the help of molecular markers, much progress has been made to resolve the higher-order phylogeny of Ericaceae recently (Kron *et al.*, 2002; Liu *et al.*, 2011; Braukmann & Stefanovic, 2012; Freudenstein *et al.*, 2016; Lallemand *et al.*, 2016). Three green lineages – the core Ericaceae (i.e. the inverted anther clade), arbutoids, and the genus *Enkianthus* – have been repeatedly shown to be monophyletic with high support, with *Enkianthus* sister to the rest of Ericaceae (Fig. 3; Braukmann & Stefanovic, 2012; Freudenstein *et al.*, 2016; Lallemand *et al.*, 2016). However, the placement of the mycoheterotrophic Ericaceae taxa differ between studies (Fig. 3). Based on different molecular markers, essentially all possible topological permutations have been recovered at one point or the other, but consistently without strong support, particularly the position of the pyroloids (Fig. 3; Braukmann & Stefanovic, 2012; Freudenstein *et al.*, 2016; Lallemand *et al.*, 2016). Our current, mitogenome-based results provide the first phylogeny of Ericaceae with strong support across the whole backbone (Fig. 5). The sister-group relationship of Monotropeae and Pterosporeae supports the monophyly of monotropoids, consistent with recent studies (Freudenstein *et al.*, 2016; Lallemand *et al.*, 2016). Monotropoids (Fig. 1k) are sister to arbutoids, an arrangement that has also been suggested based on morphological evidence and molecular data (Copeland, 1941; Freudenstein *et al.*, 2016), and which is strongly supported in Braukmann & Stefanovic (2012) and Lallemand *et al.* (2016). Our AU and SH results also reject the monophyly of mycoheterotrophic Ericaceae (pyroloids + monotropoids), in line with these studies, and these lines of evidence support two independent origins of mycoheterotrophy in Ericaceae. However, for the first time we infer that the mycoheterotrophic pyroloids (Fig. 1l) are the sister group of the core Ericaceae clade with strong support (Fig. 5). In future classification updates, tribe Pyroleae should therefore be split from the subfamily Monotropoideae, and recognised as its own subfamily, Pyroloideae Kostel. A more extensive sampling of Ericaceae is needed to further clarify relationships within each of its major groups. Even so, we predict that this higher-order phylogeny will remain the same, considering that the monophyly of each group has been strongly supported by multiple studies based on more extensive sampling (Liu *et al.*, 2011; Freudenstein *et al.*, 2016; Lallemand *et al.*, 2016), and the well supported backbone relationships recovered here among the major clades.

Conclusion

Mitochondrial genomes have been largely ignored in phylogenomic analysis of plant relationships, for several possible reasons: (i) their genomes are highly rearranged (Gualberto *et al.*, 2014); (ii) they are prone to RNA editing (Small *et al.*, 2019); (iii) they are recovered at lower coverage in genome skims (in the sense of Steele *et al.*, 2012; Straub *et al.*, 2012) due to stoichiometric differences compared with plastid genomes for gene coverage in genomic assemblies (Gualberto & Newton, 2017); (iv) they are perceived to be too slowly evolving, with relatively little variation present in individual genes (Wolfe *et al.*, 1987). However, gene recovery from genome assemblies does not depend on completing whole genomes, and RNA editing appears not to have substantial noticeable effect in our inferences, in line with the findings of Qiu *et al.* (2010) on inference of overall angiosperm relationships. In addition, the generally slow rate of evolution and specifically the lower rate observation observed in mitochondrial compared with plastid genes appear to be advantageous in inferences that employ mitochondrial phylogenomic data, in which information from multiple slowly evolving genes is combined. The mitochondrial genome effectively operates as a single linkage group, and so may mislead phylogenetic inference of species trees when incomplete lineage sorting or hybridisation/introgression events occur (although the impact of these phenomena may be relatively localised on neighbouring branches, for example Maddison, 1997). Despite this limitation, Doyle (2022) argued that ideal phylogenomic studies should include nuclear and organellar markers to better understand the evolution of both. To date the plastid genome has attracted the bulk of attention in phylogenomic studies that include organellar data, and we suggest that the mitochondrial genome should be included, when feasible, in phylogenomic studies aimed at inferring higher-order plant relationships, particularly in lineages thought to harbour heterotrophic plants. Although mitochondrial phylogenomic data can be relatively easily recovered in genome skims, their recovery as unbaited byproducts in target-sequence capture approaches may be too poor to be generally useful (Baker *et al.*, 2021). However, mitochondrial genes could be used in gene panels for target-sequence capture kits, potentially as a patch to the PAFTOL panel of Johnson *et al.* (2019).

We demonstrated that mitochondrial phylogenomics can successfully solve the uncertain placement of mycoheterotrophic lineages within monocots and Ericaceae, two distinct ancient groups with comparable crown ages (i.e. *c.* 136 million years (Myr) vs *c.* 117 Myr, respectively; Givnish *et al.*, 2018; Schwery *et al.*, 2015), and that they are likely to be generally useful for inferring backbone relationships among sampled photosynthetic lineages. The conservative nature of mitochondrial genes may make them particularly useful for placing lineages that are hard to place based on highly rate-elevated plastid genes (please refer to also Jost *et al.*, 2021). We predict, for example, that mitochondrial phylogenomics will help with other unresolved placements of mycoheterotrophic taxa, as in Orchidaceae as a whole, which has the greatest number of origins of fully heterotrophic plants (Merckx & Freudenstein, 2010).

Acknowledgements

The authors thank Vivienne Lam and Nathaniel Klimpert for helping with laboratory work, data analyses and writing the

manuscript, Andy Fyon for help in locating *Chimaphila*, *Epigaea*, and *Moneses* populations in the field, and multiple workers for permission to use photographs here (Fig. 1). We are grateful to the reviewers and Associate Editor for their helpful comments. This research was supported by the Natural Sciences and Engineering Research Council of Canada (NSERC) Discovery grants to SWG and SS.

Author contributions

QL, SS and SWG designed the research. TWAB, JLSM, FP, VSFTM and SS provided materials. QL and TWAB carried out the data collection. QL performed most of the data analysis, with input from MSG. QL, SS and SWG performed the data interpretation and wrote the manuscript. TWAB, MSG, JLSM, FP and VSFTM participated in manuscript writing and editing.

ORCID

Thomas W. A. Braukmann https://orcid.org/0000-0002-2452-3776
Sean W. Graham https://orcid.org/0000-0001-8209-5231
Qianshi Lin https://orcid.org/0000-0001-5477-8000
Juliana Lischka Sampaio Mayer https://orcid.org/0000-0001-7439-5234
Vincent S. F. T. Merckx https://orcid.org/0000-0002-3959-8623
Fábio Pinheiro https://orcid.org/0000-0003-3243-2652
Marybel Soto Gomez https://orcid.org/0000-0003-1812-7416
Saša Stefanović https://orcid.org/0000-0001-8290-895X

Data availability

All data supporting the findings of this study are available within the paper, its supplementary materials published online, or are openly available in GenBank (MN907133–MN907168, MT023139–MT055781 for monocots; MZ593998–MZ594455 for Ericaceae) and figshare (https://doi.org/10.6084/m9.figshare.14879796.v2).

References

Adams KL, Qiu YL, Stoutemyer M, Palmer JD. 2002. Punctuated evolution of mitochondrial gene content: high and variable rates of mitochondrial gene loss and transfer to the nucleus during angiosperm evolution. *Proceeding of the National Academy of Sciences, USA* **99**: 9905–9912.

APG. 2003. An update of the Angiosperm Phylogeny Group classification for the orders and families of flowering plants: APG II. *Botanical Journal of the Linnean Society* **141**: 399–436.

APG. 2009. An update of the Angiosperm Phylogeny Group classification for the orders and families of flowering plants: APG III. *Botanical Journal of the Linnean Society* **161**: 105–121.

APG. 2016. An update of the Angiosperm Phylogeny Group classification for the orders and families of flowering plants: APG IV. *Botanical Journal of the Linnean Society* **181**: 1–20.

Baker WJ, Dodsworth S, Forest F, Graham SW, Johnson MG, McDonnell A, Pokorny L, Tate JA, Wicke S, Wickett NJ. 2021. Exploring Angiosperms353: an open, community toolkit for collaborative phylogenomic research on flowering plants. *American Journal of Botany* **108**: 1059–1065.

Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Pevzner PA. 2012. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *Journal of Computational Biology* **19**: 455–477.

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

Barrett CF, Sinn BT, Kennedy AH. 2019. Unprecedented parallel photosynthetic losses in a heterotrophic orchid genus. *Molecular Biology and Evolution* **36**: 1884–1901.

Bell D, Lin Q, Gerelle WK, Joya S, Chang Y, Taylor ZN, Rothfels CJ, Larsson A, Villarreal JC, Li FW *et al.* 2020. Organellomic data sets confirm a cryptic consensus on (unrooted) land-plant relationships and provide new insights into bryophyte molecular evolution. *American Journal of Botany* **107**: 91–115.

Bergsten J. 2005. A review of long-branch attraction. *Cladistics* **21**: 163–193.

Bergthorsson U, Richardson AO, Young GJ, Goertzen LR, Palmer JD. 2004. Massive horizontal transfer of mitochondrial genes from diverse land plant donors to the basal angiosperm *Amborella*. *Proceeding of the National Academy of Sciences, USA* **101**: 17747–17752.

Bouckaert R, Vaughan TG, Barido-Sottani J, Duchêne S, Fourment M, Gavryushkina A, Heled J, Jones G, Kühnert D, De Maio N *et al.* 2019. Beast 2.5: an advanced software platform for Bayesian evolutionary analysis. *PLoS Computational Biology* **15**: e1006650.

Braukmann TWA, Broe MB, Stefanovic S, Freudenstein JV. 2017. On the brink: the highly reduced plastomes of nonphotosynthetic Ericaceae. *New Phytologist* **216**: 254–266.

Braukmann TWA, Stefanovic S. 2012. Plastid genome evolution in mycoheterotrophic Ericaceae. *Plant Molecular Biology* **79**: 5–20.

Brinkmann H, van der Giezen M, Zhou Y, Poncelin de Raucourt G, Philippe H. 2005. An empirical assessment of long-branch attraction artefacts in deep eukaryotic phylogenomics. *Systematic Biology* **54**: 743–757.

Bromham L, Cowman PF, Lanfear R. 2013. Parasitic plants have increased rates of molecular evolution across all three genomes. *BMC Evolutionary Biology* **13**: 126.

Caddick LR, Rudall PJ, Wilkin P, Hedderson TAJ, Chase MW. 2002. Phylogenetics of Dioscoreales based on analyses of morphological and molecular data. *Botanical Journal of the Linnean Society* **138**: 123–144.

Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, Madden TL. 2009. Blast+: architecture and applications. *BMC Bioinformatics* **10**: 421.

Cameron KM, Chase MW, Rudall PJ. 2003. Recircumscription of the monocotyledonous family Petrosaviaceae to include *Japonolirion*. *Brittonia* **55**: 214–225.

Cheek M. 2003. A new species of *Afrothismia* (Burmanniaceae) from Kenya. *Kew Bulletin* **58**: 951–955.

Copeland HF. 1941. Further studies on the Monotropoideae. *Madroño* **6**: 97–119.

Cronquist A. 1968. *The evolution and classification of flowering plants*. London, UK: Nelson.

Dahlgren RM, Bremer K. 1985. Major clades of the angiosperms. *Cladistics* **1**: 349–368.

Dahlgren RM, Clifford HT. 1982. *The monocotyledons: a comparative study*. London, UK: Academic Press.

Doyle JJ. 2022. Defining coalescent genes: theory meets practice in organelle phylogenomics. *Systematic Biology* **71**: 476–489.

Edgar RC. 2004a. Muscle: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research* **32**: 1792–1797.

Edgar RC. 2004b. Muscle: a multiple sequence alignment method with reduced time and space complexity. *BMC Bioinformatics* **5**: 113.

Fay MF, Rudall PJ, Sullivan S, Stobart KL, De Brujin AY, Reeves G, Qamaruz-Zaman F, Hong WP, Joseph J, Hahn WJ *et al.* 2000. Phylogenetic studies of Asparagales based on four plastid DNA regions. In: Wilson KL, Morrison DA, eds. *Monocots: systematics and evolution*. Collingwood, Vic., Australia: CSIRO, 360–371.

Fazekas AJ, Kesanakurti PR, Burgess KS, Percy DM, Graham SW, Barret SCH, Newmaster SG, Hajibabaei M, Husband BC. 2009. Are plant species inherently harder to discriminate than animal species using DNA barcoding markers? *Molecular Ecology Resources* **9**: 130–139.

Felsenstein J. 1978. Cases in which parsimony or compatibility methods will be positively misleading. *Systematic Zoology* **27**: 401–410.

Felsenstein J. 1985. Confidence limits on phylogenies: an approach using the bootstrap. *Evolution* **39**: 783–791.

Freudenstein JV, Broe MB, Feldenkris ER. 2016. Phylogenetic relationships at the base of Ericaceae: implications for vegetative and mycorrhizal evolution. *Taxon* **65**: 794–804.

Fuse S, Tamura MN. 2000. A phylogenetic analysis of the plastid *mat*K gene with emphasis on Melanthiaceae sensu lato. *Plant Biology* **2**: 415–427.

Givnish TJ, Zuluaga A, Marques I, Lam VKY, Gomez MS, Iles WJD, Ames M, Spalink D, Moeller JR, Briggs BG *et al.* 2016. Phylogenomics and historical biogeography of the monocot order Liliales: out of Australia and through Antarctica. *Cladistics* **32**: 581–605.

Givnish TJ, Zuluaga A, Spalink D, Soto Gomez M, Lam VKY, Saarela JM, Sass C, Iles WJD, de Sousa DJL, Leebens-Mack J *et al.* 2018. Monocot plastid phylogenomics, timeline, net rates of species diversification, the power of multi-gene analyses, and a functional model for the origin of monocots. *American Journal of Botany* **105**: 1–23.

Graham SW, Lam VKY, Merckx VSFT. 2017. Plastomes on the edge: the evolutionary breakdown of mycoheterotroph plastid genomes. *New Phytologist* **214**: 48–55.

Graham SW, Reeves PA, Burns ACE, Olmstead RG. 2000. Microstructural changes in noncoding chloroplast DNA: interpretation, evolution, and utility of indels and inversions in basal angiosperm phylogenetic inference. *International Journal of Plant Sciences* **161**: S83–S96.

Gualberto JM, Mileshina D, Wallet C, Niazi AK, Weber-Lotfi F, Dietrich A. 2014. The plant mitochondrial genome: dynamics and maintenance. *Biochimie* **100**: 107–120.

Gualberto JM, Newton KJ. 2017. Plant mitochondrial genomes: dynamics and mechanisms of mutation. *Annual Review of Plant Biology* **68**: 225–252.

Hendy MD, Penny D. 1989. A framework for quantitative study of evolutionary trees. *Systematic Zoology* **38**: 297–309.

Hynson NA, Bidartondo MI, Read DJ. 2015. Are there geographic mosaics of mycorrhizal specificity and partial mycoheterotrophy? A case study in *Moneses uniflora* (Ericaceae). *New Phytologist* **208**: 1003–1007.

Iles WJD, Smith SYS, Gandolfo MA, Graham SW. 2015. Monocot fossils suitable for molecular dating analyses. *Botanical Journal of the Linnean Society* **178**: 346–374.

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.

Jąkalski M, Minasiewicz J, Caius J, May M, Selosse M-A, Delannoy E. 2021. The genomic impact of mycoheterotrophy in orchids. *Frontiers in Plant Science* **12**: 632033.

Jin J-J, Yu W-B, Yang J-B, Song Y, dePamphilis CW, Yi T-S, Li D-Z. 2020. GetOrganelle: a fast and versatile toolkit for accurate *de novo* assembly of organelle genomes. *Genome Biology* **21**: 241.

Johnson MG, Pokorny L, Dodsworth S, Botigué LR, Cowan RS, Devault A, Eiserhardt WL, Epitawalage N, Forest F, Kim JT *et al.* 2019. A universal probe set for targeted sequencing of 353 nuclear genes from any flowering plant designed using k-medoids clustering. *Systematic Biology* **68**: 594–606.

Jonker FP. 1938. A monograph of the Burmanniaceae. *Mededeelingen van het Botanisch Museum en Herbarium van de Rijks Universiteit te Utrecht* **51**: 1–279.

Jost M, Samain MS, Marques I, Graham SW, Wanke S. 2021. Discordant phylogenomic placement of Hydnoraceae and Lactoridaceae within Piperales using data from all three genomes. *Frontiers in Plant Science* **12**: 642598.

Kim J. 1996. General inconsistency conditions for maximum parsimony: effects of branch lengths and increasing numbers of taxa. *Systematic Biology* **45**: 363–374.

Klopfstein S, Massingham T, Goldman N. 2017. More on the best evolutionary rate for phylogenetic analysis. *Systematic Biology* **66**: 769–785.

Knoop V. 2004. The mitochondrial DNA of land plant: peculiarities in phylogenetic perspective. *Current Genetics* **46**: 123–139.

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. *Botanical Review* **68**: 35–423.

Lallemand F, Gaudeul M, Lambourdiere J, Matsuda Y, Hashimoto Y, Selosse MA. 2016. The elusive predisposition to mycoheterotrophy in Ericaceae. *New Phytologist* **212**: 314–319.

Lallemand F, Logacheva M, Le Clainche I, Bérard A, Zheleznaia E, May M, Jakalski M, Delannoy É, Le Paslier MC, Selosse MA. 2019. Thirteen new plastid genomes from mixotrophic and autotrophic species provide insights into heterotrophy evolution in Neottieae orchids. *Genome Biology and Evolution* **11**: 2457–2467.

Lam VKY, Darby H, Merckx VSFT, Lim G, Yukawa T, Neubig KM, Abbott JR, Beatty GE, Provan J, Soto Gomez M *et al.* 2018. Phylogenomic inference *in extremis*: a case study with mycoheterotroph plastomes. *American Journal of Botany* **105**: 480–494.

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.

Lanfear R, Frandsen PB, Wright AM, Senfeld T, Calcott B. 2016. PartitionFinder 2: new methods for selecting partitioned models of evolution for molecular and morphological phylogenetic analyses. *Molecular Biology and Evolution* **34**: 772–773.

Langmead B, Salzberg SL. 2012. Fast gapped-read alignment with Bowtie 2. *Nature Methods* **9**: 357–359.

Leake JR. 2004. Myco-heterotrophic/epiparasitic plant interactions with ectomycorrhizal and arbuscular mycorrhizal fungi. *Current Opinion in Plant Biology* **7**: 422–428.

Leake JR. 2005. Plants parasitic on fungi: unearthing the fungi in myco-heterotrophs and debunking the 'saprophytic' plant myth. *Mycologist* **19**: 113–122.

Lemaire B, Huysmans S, Smets E, Merckx VSFT. 2011. Rate accelerations in nuclear 18S rDNA of mycoheterotrophic and parasitic angiosperms. *Journal of Plant Research* **124**: 561–576.

Li L, Wang B, Liu Y, Qiu Y-L. 2009. The complete mitochondrial genome sequence of the hornwort *Megaceros aenigmaticus* shows a mixed mode of conservative yet dynamic evolution in early land plant mitochondrial genomes. *Journal of Molecular Evolution* **68**: 665–678.

Li Y-X, Li Z-H, Schuiteman A, Chase MW, Li J-W, Huang W-C, Hidayat A, Wu S-S, Jin X-H. 2019. Phylogenomics of Orchidaceae based on plastid and mitochondrial genomes. *Molecular Phylogenetics and Evolution* **139**: 106540.

Lim GS, Barrett CF, Pang CC, 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**: 1129–1137.

Liu ZW, Wang ZH, Zhou J, Peng H. 2011. Phylogeny of Pyroleae (Ericaceae): implications for character evolution. *Journal of Plant Research* **124**: 325–337.

Logacheva MD, Schelkunov MI, Nuraliev MS, Samigullin TH, Penin AA. 2014. The plastid genome of mycoheterotrophic monocot *Petrosavia stellaris* exhibits both gene losses and multiple rearrangements. *Genome Biology and Evolution* **6**: 238–246.

Maas PJM, Maas-van de Kamer H, van Benthem J, Snelders HCM, Rübsamen T. 1986. Burmanniaceae. In: Organization for Flora Neotropica, ed. *Flora Neotropica, vol. 42*. New York, NY, USA: New York Botanical Garden, 1–189.

Maas-van de Kamer H. 1998. Burmanniaceae. In: Kubitzki K, ed. *The families and genera of vascular plants, vol. III, flowering plants. monocotyledons: Lilianae (except Orchidaceae)*. Berlin, Germany: Springer-Verlag, 154–164.

Maddison WP. 1997. Gene trees in species trees. *Systematic Biology* **46**: 523–536.

Maddison WP, Maddison DR. 2018. *Mesquite: a modular system for evolutionary analysis. v.3.51*. [WWW document] URL http://www.mesquiteproject.org [accessed 1 September 2021].

Mennes CB, Lam VKY, Rudall PJ, Lyon SP, Graham SW, Smets EF, Merckx VSFT. 2015. Ancient Gondwana break-up explains the distribution of the mycoheterotrophic family Corsiaceae (Liliales). *Journal of Biogeography* **42**: 1123–1136.

Merckx VSFT. 2013. Mycoheterotrophy: an introduction. In: Merckx VSFT, ed. *Mycoheterotrophy: the biology of plants living on fungi*. New York, NY, USA: Springer-Verlag, 1–18.

Merckx VSFT, Bakker FT, Huysmans S, Smets E. 2009a. Bias and conflict in phylogenetic inference of myco-heterotrophic plants: a case study in Thismiaceae. *Cladistics* **25**: 64–77.

Merckx VSFT, Bidartondo MI, Hynson NA. 2009b. Myco-heterotrophy: when fungi host plants. *Annals of Botany* **104**: 1255–1261.

Merckx VSFT, Chatrou LW, Lemaire B, Sainge MN, Huysmans S, Smets EF. 2008. Diversification of myco-heterotrophic angiosperms: evidence from Burmanniaceae. *BMC Evolutionary Biology* **8**: 178.

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

Merckx VSFT, Gomes SIF, Wapstra M, Hunt C, Steenbeeke G, Mennes CB, Walsh N, Smissen R, Hsieh TH, Smets EF *et al.* 2017. The biogeographical history of the interaction between mycoheterotrophic *Thismia* (Thismiaceae) plants and mycorrhizal *Rhizophagus* (Glomeraceae) fungi. *Journal of Biogeography* **44**: 1869–1879.

Merckx VSFT, Mennes CB, Peay KG, Geml J. 2013. Evolution and diversification. In: Merckx VSFT, ed. *Mycoheterotrophy: the biology of plants living on fungi*. New York, NY, USA: Springer-Verlag, 222–226.

Merckx VSFT, Smets EF. 2014. *Thismia americana*, the 101st anniversary of a botanical mystery. *International Journal of Plant Sciences* **175**: 165–175.

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.

One Thousand Plant Transcriptomes Initiative. 2019. One thousand plant transcriptomes and the phylogenomics of green plants. *Nature* **574**: 679–685.

Perez-Lamarque B, Selosse M, Opik M, Morlon H, Martos F. 2020. Cheating in arbuscular mycorrhizal mutualism: a network and phylogenetic analysis of mycoheterotrophy. *New Phytologist* **226**: 1822–1835.

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

Petersen G, Cuenca A, Zervas A, Ross GT, Graham SW, Barrett CF, Davis JI, Seberg O. 2017. Mitochondrial genome evolution in Alismatales: size reduction and extensive loss of ribosomal protein genes. *PLoS ONE* **12**: e0177606.

Petersen G, Darby H, Lam VKY, Pedersen HÆ, Merckx VSFT, Zervas A, Seberg O, Graham SW. 2019. Mycoheterotrophic *Epirixanthes* (Polygalaceae) has a typical angiosperm mitogenome but unorthodox plastid genomes. *Annals of Botany* **124**: 791–807.

Qiu YL, Li L, Hendry TA, Li R, Taylor DW, Issa MJ, Ronen AJ, Vekaria ML, White AM. 2006. Reconstructing the basal angiosperm phylogeny: evaluating information content of mitochondrial genes. *Taxon* **55**: 837–856.

Qiu YL, Li L, Wang B, Xue J-Y, Hendry TA, Li R-Q, Brown JW, Liu Y, Hudson GT, Chen Z-D. 2010. Angiosperm phylogeny inferred from sequences of four mitochondrial genes. *Journal of Systematics and Evolution* **48**: 391–425.

Schlechter R. 1921. Die Thismieae. *Notizblatt des Botanischen Gartens Berlin-Dahlem* **8**: 31–45.

Schwery O, Onstein RE, Bouchenak-Khelladi Y, Xing Y, Carter RJ, Linder HP. 2015. As old as the mountains: the radiations of the Ericaceae. *New Phytologist* **207**: 355–367.

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

Shepeleva EA, Schelkunov MI, Hroneš M, Sochor M, Dančák M, Merckx VSFT, Kikuchi IABS, Chantanaorrapint S, Suetsugu K, Tsukaya H *et al.* 2020. Phylogenetics of the mycoheterotrophic genus *Thismia* (Thismiaceae: Dioscoreales) with a focus on the Old World taxa: delineation of novel natural groups and insights into the evolution of morphologial traits. *Botanical Journal of the Linnean Society* **193**: 287–315.

Shimodaira H. 2002. An approximately unbiased test of phylogenetic tree selection. *Systematic Biology* **51**: 492–508.

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

Silvestro D, Michalak I. 2012. RAxMLGUI: a graphical front-end for RAxML. *Organisms, Diversity and Evolution* **12**: 335–337.

Skippington E, Barkman TJ, Rice DW, Palmer JD. 2015. Miniaturized mitogenome of the parasitic plant *Viscum scurruloideum* is extremely divergent and dynamic and has lost all *nad* genes. *Proceedings of the National Academy of Sciences, USA* **112**: E3515–E3524.

Skippington E, Barkman TJ, Rice DW, Palmer JD. 2017. Comparative mitogenomics indicates respiratory competence in parasitic *Viscum* despite loss of complex I and extreme sequence divergence, and reveals horizontal gene transfer and remarkable variation in genome size. *BMC Plant Biology* **17**: 49.

Small ID, Schallenberg-Rüdinger M, Takenaka M, Mireau H, Ostersetzer-Biran O. 2019. Plant organellar RNA editing: what 30 years of research has revealed. *The Plant Journal* **101**: 1040–1056.

Soltis PS, Soltis DE. 2003. Applying the bootstrap in phylogeny reconstruction. *Statistical Science* **18**: 256–267.

Soto Gomez M. 2020. *Phylogenomic studies of the monocot sister orders Pandanales and Dioscoreales*. PhD thesis, The University of British Columbia, Vancouver, BC, Canada.

Soto Gomez M, Lin Q, Da Silva Leal E, Gallaher T, Scherberich D, Mennes CB, Smith SY, Graham SW. 2020. A bi-organellar phylogenomic study of Pandanales: inference of higher-order relationships and unusual rate-variation. *Cladistics* **36**: 481–504.

Stamatakis A. 2006. RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics* **22**: 2688–2690.

Steele PR, Hertweck KL, Mayfield D, McKain MR, Leebens-Mack J, Pires JC. 2012. Quality and quantity of data recovered from massively parallel sequencing: examples in Asparagales and Poaceae. *American Journal of Botany* **99**: 330–348.

Stern DB, Palmer JD. 1986. Tripartite mitochondrial genome of spinach: physical structure, mitochondrial gene mapping, and locations of transposed chloroplast DNA sequences. *Nucleic Acids Research* **14**: 5651–5666.

Straub SCK, Parks M, Weitemier K, Fishbein M, Cronn RC, Liston A. 2012. Navigating the tip of the genomic iceberg: next-generation sequencing for plant systematics. *American Journal of Botany* **99**: 349–364.

Susko E, Roger AJ. 2021. Long branch attraction biases in phylogenetics. *Systematic Biology* **70**: 838–843.

Swofford DL. 2002. *PAUP*. Phylogenetic analysis using parsimony (* and other methods). v.4*. Sunderland, MA, USA: Sinauer Associates.

Swofford DL, Olsen GJ, Wadell PJ, Hillis DM. 1996. Phylogenetic inference. In: Hillis DM, Moritz C, Mable BK, eds. *Molecular systematics, 2nd edn.* Sunderland, MA, USA: Sinauer, 407–514.

Tedersoo L, Pellet P, Kõljalg U, Selosse MA. 2007. Parallel evolutionary paths to mycoheterotrophy in understorey Ericaceae and Orchidaceae: ecological evidence for mixotrophy in Pyroleae. *Oecologia* **151**: 206–217.

Wolfe KH, Li WH, Sharp PM. 1987. Rates of nucleotide substitution vary greatly among plant mitochondrial, chloroplast, and nuclear DNAs. *Proceedings of the National Academy of Sciences, USA* **84**: 9054–9058.

Yang Z. 2006. *Computational molecular evolution*. Oxford, UK: Oxford University Press.

Yokoyama J, Koizumi Y, Yokota M, Tsukaya H. 2008. Phylogenetic position of *Oxygyne shinzatoi* (Burmanniaceae) inferred from 18S rDNA sequences. *Journal of Plant Research* **121**: 27–32.

Yurina NP, Odintsowa WS. 2016. Mitochondrial genome structure of photosynthetic eukaryotes. *Biochemistry* **81**: 101–113.

Zhang C, Rabiee M, Sayyari E, Mirarab S. 2018. ASTRAL-III: polynomial time species tree reconstruction from partially resolved gene trees. *BMC Bioinformatics* **19**: 153.

Supporting Information

Additional Supporting Information may be found online in the Supporting Information section at the end of the article.

Fig. S1 Summary of bootstrap support for phylogenetic placement of mycoheterotrophic monocots inferred from partitioned and unpartitioned amino-acid maximum likelihood analyses.

Fig. S2 Phylogenetic placement of mycoheterotrophic monocots inferred from the unpartitioned maximum likelihood analysis of DNA data, without predicted RNA editing sites.

Fig. S3 Summary of bootstrap support for phylogenetic placement of mycoheterotrophic Ericaceae inferred from partitioned and unpartitioned amino-acid maximum likelihood analyses.

Fig. S4 Phylogenetic placement of mycoheterotrophic Ericaceae inferred from the unpartitioned maximum likelihood analysis of DNA data, without predicted RNA editing sites.

Fig. S5 Phylogenetic placement of mycoheterotrophic Ericaceae inferred from the unpartitioned maximum likelihood analysis of DNA data, without RNA editing sites identified in *Hypopitys monotropa*.

Fig. S6 Relative substitution rates between green and mycoheterotrophic monocot lineages.

Fig. S7 Relative substitution rates between green and mycoheterotrophic Ericaceae lineages.

Methods S1 DNA isolation, library preparation and sequencing.

Methods S2 Characterising effect of RNA edit sites on DNA sequence data.

Methods S3 Characterising GC content, codon usage and rate elevation in mycoheterotrophic lineages.

Methods S4 Topological constraint tests.

Table S1 Specimen and sequence source information.

Table S2 Status of 37 protein-coding mitochondrial genes in newly sequenced monocot taxa in this study.

Table S3 Status of 38 protein-coding mitochondrial genes in Ericales taxa.

Table S4 Optimal substitution models for unpartitioned analyses and for final partitions from 'G × C' (gene-by-codon) partitioning schemes.

Table S5 GC content and fraction of codon usage.

Table S6 Location of predicted RNA editing sites in monocot DNA matrix.

Table S7 Location of predicted RNA editing sites in Ericales DNA matrix.

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.