Check for updates

Received: 26 April 2026 | Accepted: 17 July 2026

DOI: 10.1002/ajb2.70267

RESEARCH ARTICLE

Botany

Placing very long branch taxa in the plant tree of life: A case study with Dioscoreales mycoheterotrophs

Marybel Soto Gomez[1,2,3] | Qianshi Lin[1,2,4] | Nathaniel J. Klimpert[1,2] | Vivienne K. Y. Lam[1,2] | Juan Viruel[3,5,6] | Vincent S. F. T. Merckx[7,8] | Paula J. Rudall[3] | Sean W. Graham[1,2]

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

[2]Biodiversity Research Centre, University of British Columbia, 2212 Main Mall, Vancouver, BC V6T 1Z4, Canada

[3]Jodrell Laboratory, Royal Botanic Gardens, Kew, Richmond TW9 3DS, UK

[4]Department of Biology, Lakehead University, 955 Oliver Rd, Thunder Bay, ON P7B 5E1, Canada

[5]Technological College, University of Zaragoza, Crta Cuarte sn, Huesca 22071, Spain

[6]Institute for Biocomputation and Physics of Complex Systems (BIFI), Zaragoza 50018, Spain

[7]Naturalis Biodiversity Center, Leiden 2332 AA, the Netherlands

[8]Department of Evolutionary and Population Biology, Institute for Biodiversity and Ecosystem Dynamics, University of Amsterdam, Amsterdam 1098 XH, the Netherlands

Correspondence
Marybel Soto Gomez, Jodrell Laboratory, Royal Botanic Gardens, Kew, Richmond, Surrey TW9 3DS, UK.
Email: marybel.soto@gmail.com

Sean W. Graham, Department of Botany, University of British Columbia, 6270 University Boulevard, Vancouver, BC, V6T 1Z4, Canada.
Email: swgraham@mail.ubc.ca

Funding information
Natural Sciences and Engineering Research Council of Canada

Abstract

Premise: The inclusion of heterotrophic plants in broader systematic studies is challenging due to substantial morphological modification, gene loss, and accelerated substitution rates following photosynthesis loss. In particular, inclusion of very long branch (VLB) taxa with highly elevated rates may lead to phylogenomic misinference.

Methods: We explored how rapidly evolving plastomes behave in model-based phylogenetic inferences in the yam order Dioscoreales, which experienced convergent photosynthesis losses. Our taxon sampling included mycoheterotrophic lineages displaying a broad range of substitution rates and all photosynthetic genera.

Results: When only moderately rapidly evolving heterotrophs are included, relationships within Dioscoreales are congruent with recent mitochondrial analysis, and generally strongly supported. Photosynthetic *Stenomeris* (represented by only three genes) is moderately supported as sister to Dioscoreaceae. Notably, Burmanniaceae and Thismiaceae, two mycoheterotrophic families including highly reduced plastomes, are strongly supported as distantly related within Dioscoreales. Constrained analysis rejects combining Thismiaceae with Burmanniaceae, conflicting with current angiosperm classification. Including a subset of VLB mycoheterotrophs in phylogenetic inference results in localized or broad reduction in branch support—substantially so for Afrothismiaceae, a family placing in disparate locations in Dioscoreales (and beyond) in variant analyses, with no ability to differentiate among its alternative placements.

Conclusions: Inclusion of highly rate-elevated taxa in analyses of Dioscoreales phylogeny can depress branch support. Inferred relationships in the order are otherwise congruent with studies based on mitochondrial data; the family-level classification needs updating. The extraordinarily rapidly evolving family Afrothismiaceae places inconsistently in diverse plastid analyses, pointing to probable analytical limits for plastid-based phylogenomic analyses of VLB heterotrophs.

KEYWORDS

across-site compositional heterogeneity, *Afrothismia*, Dioscoreaceae, long branch attraction (LBA), loss of photosynthesis, monocot classification, monocot phylogeny, plastid phylogenomics, Stenomeridaceae

Phylogenomic data sets based on the plastid (chloroplast) genome have facilitated well-supported inferences of higher-level relationships both within and among most orders of monocots (e.g., Steele et al., 2012; Henriquez et al., 2014; Logacheva et al., 2014; Comer et al., 2015; Mennes et al., 2015; Ross et al., 2015; Sass et al., 2016; Givnish et al., 2010, 2015, 2016, 2018; Lam et al., 2018; Soto Gomez et al., 2020; Li et al., 2019, 2021; Zuntini et al., 2021;

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.
©2026 The Author(s). American Journal of Botany published by Wiley Periodicals LLC on behalf of Botanical Society of America.

Barrett et al., 2013, 2014, 2016, 2024a). Phylogenomic approaches have not yet been attempted on an extensive scale for the yam order Dioscoreales, although Garrett et al. (2023) employed plastid genomic data for a constituent family (Thismiaceae), and Givnish et al. (2018) and Lam et al. (2018) included several exemplar taxa in broader phylogenetic studies. Nonetheless, this order is perhaps the last major monocot clade in need of detailed attention using phylogenomic data (e.g., Trias-Blasi et al., 2015). A major complicating factor in addressing higher-order relationships in Dioscoreales is the repeated origin of fully mycoheterotrophic (FM), non-photosynthetic taxa in the order (Merckx et al., 2008a, 2010, 2013). These lineages, like heterotrophic plants in general, are often highly modified morphologically and can display substantially elevated substitutional rates that can be problematic in phylogenetic or phylogenomic inference (e.g., Neyland and Hennigan, 2003; Nickrent et al., 2004; Barkman et al., 2007; Merckx et al., 2009b; Su et al., 2015; Bellot et al., 2016; Lam et al., 2016, 2018; Naumann et al., 2016; Shepeleva et al., 2020—and see Felsenstein, 1978; Hendy and Penny, 1989, on the general problem of long branch attraction).

Fully mycoheterotrophic plants obtain fixed carbon from root-associated fungal partners rather than through photosynthesis (Leake, 2004; Merckx et al., 2009a). Therefore, the strong selective constraints that normally operate in genes associated with photosynthetic function can be completely lost, leading to rapid gene loss (e.g., Barrett et al., 2014; Graham et al., 2017). Other retained genes with non-photosynthetic functions can experience relaxed purifying selection (e.g., Garrett et al., 2023), which can contribute to elevated DNA substitution rates. However, the generally elevated substitution rates in heterotrophs likely relate to higher overall rates of mutation that result from a reduction in the efficiency of plastid DNA replication and repair machinery (e.g., Wicke et al., 2016). The mitochondrial genome is known to be less affected by substitutional rate elevation and gene loss in heterotrophic plants, and it has occasionally been applied to phylogenetic inference in these lineages using either small sets of genes (e.g., Merckx et al., 2009b; Mennes et al., 2013) or genome-scale sets of genes (Soto Gomez et al., 2020; Lin et al., 2022; Barrett et al., 2024b).

Plastid sequences still underlie much of our knowledge of higher-order plant relationships, highlighting their continuing importance for integrating previously unsequenced or problematic lineages into the plant tree of life. The photosynthesis-related genes *atpB* and *rbcL* are two of the three major loci that broadly underpin the Angiosperm Phylogeny Group classification systems (Angiosperm Phylogeny Group, 1998, 2003, 2009, 2016). However, both genes are generally missing from the plastid genomes of fully heterotrophic plants (e.g., Graham et al., 2017; Lam et al., 2018; Wicke and Naumann, 2018). Nonetheless, multiple phylogenomic studies have shown that retained plastid genes unrelated to photosynthesis in heterotrophic plants can be employed effectively in phylogenetic analyses, despite sometimes extensive gene loss and often substantial rate elevation in the included heterotrophic lineages (e.g., Delannoy et al., 2011; Logacheva et al., 2014; Givnish et al., 2015, 2018; Lam et al., 2015, 2018; Mennes et al., 2015; Bodin et al., 2016; Feng et al., 2016; Barrett et al., 2018, 2019, 2024a, 2024b; Lee et al., 2020; Soto Gomez et al., 2020; Yudina et al., 2021; Klimpert et al., 2022; Garrett et al., 2023). Indeed, a broad survey by Lam et al. (2018) to place the major lineages of mycoheterotrophs on the angiosperm tree of life demonstrated that taxa with elevated rates—and often relatively few retained genes—can be successfully incorporated in plastid phylogenomic analyses. Nonetheless, their study and others also provided evidence of strong long branch artifacts for a subset of mycoheterotrophic monocots displaying more extreme rates of DNA substitution, including members of Burmanniaceae (Lam et al., 2018) and Thismiaceae (Garrett et al., 2023) in Dioscoreales. Here we refer to these potentially problematic lineages as very long branch (VLB) taxa, which we define quantitatively below. It is still unclear how well VLB lineages can be accommodated in phylogenomic inference using plastid data, which is a major challenge in their systematics given the numerous losses of photosynthesis that have taken place in autotrophic lineages across the tree of life (e.g., Merckx and Freudenstein, 2010; Figueroa-Martinez et al., 2015; Dorrell at al., 2019; Nickrent, 2020; DeShaw et al., 2022).

Family-level circumscription in Dioscoreales remains broadly unsettled, with three to seven families recognized in previous studies (Caddick et al., 2002a, b; Merckx et al., 2009b, 2010; Merckx and Smets, 2014; Givnish et al., 2018; Lam et al., 2016, 2018; Cheek et al., 2024). The most recent Angiosperm Phylogeny Group scheme (Angiosperm Phylogeny Group, 2016) did not update the classification of the order (recognizing only Burmanniaceae, Dioscoreaceae, and Nartheciaceae), although they also acknowledged that there are problems with this status quo. One challenging classification issue is whether the mycoheterotrophic lineages in the order should be recognized under a single very broadly defined family (Burmanniaceae), as in recent angiosperm-wide classification schemes (Angiosperm Phylogeny Group, 2003, 2009, 2016), or whether Thismiaceae should be recognized as a separate family (Merckx et al., 2009b, 2010; Hunt et al., 2014; Merckx and Smets, 2014; Givnish et al., 2018; Lam et al., 2016, 2018; Shepeleva et al., 2020). Recognizing both Thismiaceae and Burmanniaceae may also require acceptance of several photosynthetic lineages at the family level, and/or additional re-circumscription of existing families (see below). For example, the fully mycoheterotrophic genus *Afrothismia* was placed in Thismiaceae until recently (Schlecter, 1921; Jonker, 1938; Maas et al., 1986; Maas-van de Kamer, 1998), but has not been included in plastid-based phylogenetic studies to date. However, few-gene phylogenetic inferences using mitochondrial and nuclear markers indicate that *Afrothismia* represents an additional independent loss of photosynthesis in Dioscoreales (Merckx and Bidartondo, 2008; Merckx et al., 2009b, 2010; Merckx and Smets, 2014; Shepeleva et al., 2020), a finding that is well supported in a mitochondrial phylogenomic study (Lin et al., 2022). Based on these phylogenetic findings and evidence of its morphological

15372197, 0, Downloaded from https://bsapubs.onlinelibrary.wiley.com/doi/10.1002/ajb2.70267, Wiley Online Library on [08/10/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

distinctness from Thismiaceae and related photosynthetic taxa, *Afrothismia* was recently recognized at the family level as Afrothismiaceae (Cheek et al., 2024).

The precise relationships of photosynthetic taxa within individual Dioscoreales families have also proved challenging to reconstruct. Most autotrophs have been assigned to the strictly photosynthetic families Dioscoreaceae and Nartheciaceae (e.g., Angiosperm Phylogeny Group, 2003, 2009, 2016), although *Burmannia*, a member of Burmanniaceae, includes both photosynthetic and FM species. The composition of Nartheciaceae has been relatively stable in angiosperm classification (e.g., Angiosperm Phylogeny Group, 2003, 2009, 2016) following the discovery that the genus *Isidrogalvia* (previously included in the family by Tamura, 1998 and Tamura et al., 2004) belongs in Alismatales as part of Tofieldiaceae (Azuma and Tobe, 2005, 2011). However, relationships among the five genera that now comprise Nartheciaceae have varied in previous studies based on a few genes or phylogenomic data sets (Merckx et al., 2008b; Fuse et al., 2012; Zhao et al., 2012; Garrett et al., 2023) and therefore need attention. The circumscription of the yam family Dioscoreaceae has also remained in flux, as two photosynthetic taxa, *Tacca* and *Trichopus*, have either been recognized as part of it (e.g., Caddick et al., 2002a, b; Angiosperm Phylogeny Group, 2003, 2009, 2016), or as two distinct monogeneric families, Taccaceae and Trichopodaceae, respectively (Huber, 1998; Kubitzki, 1998; Chase et al., 2006; Merckx et al., 2010; Merckx and Smets, 2014; Givnish et al., 2018; Lam et al., 2016, 2018; Lin et al., 2022). In addition, *Stenomeris* has been placed in Dioscoreaceae in recent angiosperm classification schemes (Angiosperm Phylogeny Group, 1998, 2003, 2009, 2016), but has also been recognized in its own family, Stenomeridaceae (reviewed by Caddick et al., 2002a). Only three plastid genes are available for *Stenomeris*, which may have contributed to its inconsistent local placement in the few phylogenetic studies that have included the genus (Caddick et al., 2002a; Wilkin et al., 2005; Merckx et al., 2009b, 2010; Hsu et al., 2013; Viruel et al., 2016). A recent nuclear phylogenomic data set recovered *Stenomeris* as sister to a clade comprising *Dioscorea*, *Tacca*, and *Trichopus*, although with substantial gene-tree conflict for this arrangement (Zuntini et al., 2024).

Eight of the 22 genera in Dioscoreales have yet to be included in plastid phylogenomic analyses. A broadly representative plastid phylogenomic sampling would be useful to address unsettled higher-order relationships and family-level circumscriptions in the order. The degree of impact of VLB lineages on plastid-based phylogenetic inference in the order is also unclear. Here we present a plastid phylogenomic and exemplar-based survey of Dioscoreales to address its genus- and family-level relationships, with a special focus on placing its fully mycoheterotrophic lineages. We specifically address (1) genus-level relationships in each family—our updated sampling now includes all of the photosynthetic genera and approximately half of the FM genera in Dioscoreales, totalling three quarters of the genera; (2) the validity of recent angiosperm classification schemes (e.g., Angiosperm Phylogeny Group, 2016) in which all mycoheterotrophic taxa are lumped in Burmanniaceae; and (3) the general impact of including VLB mycoheterotrophic lineages on phylogenomic inference, and whether strong negative effects can be ameliorated. As a backdrop to our study, we first present analyses of a "core" taxon set that includes most photosynthetic taxa and the least rapidly evolving mycoheterotrophic lineages examined here (i.e., *Haplothismia exannulata* in Thismiaceae; *Burmannia itoana* and *Campylosiphon congestus* in Burmanniaceae); the latter taxa may be least affected by long branch artifacts. We discuss a series of taxon-addition experiments that are mostly focused on different individual VLB mycoheterotrophic lineages to reduce the chances of long branch taxa attracting each other in phylogenetic inference (see Lam et al., 2018). We also explore several model-based approaches to attempt to reduce the effects of long branch artifacts in phylogenomic inference.

MATERIALS AND METHODS

Taxon sampling and library preparation

We used genome skims (Straub et al., 2012) to generate a new phylogenomic data set comprising the full complement of plastid protein-coding genes and rDNA genes for seven Dioscoreales species (Appendix S1). We added these data to published sequences for 29 other taxa in the order, along with 62 additional monocots as outgroups, to obtain a sampling that includes representatives of all 12 monocot orders (sensu Givnish et al., 2018; see Appendix S1 for source information of all sampled taxa). The 36 taxa sampled in Dioscoreales represent all nine strictly photosynthetic genera in the order and seven of 13 mycoheterotrophic genera for which we were able to retrieve plastid data. Six of the seven latter genera are fully mycoheterotrophic (FM), and *Burmannia* includes both photosynthetic and FM taxa (Table 1). The species *Afrothismia kupensis* Cheek & S.A. Williams, sampled here, is distinct from *A. gesnerioides* H. Maas (Cheek et al., 2019), a name recently applied to the material we used (Sainge et al., 2017; see Appendix S1). We were unable to obtain DNA of sufficient quality to be sequenced for photosynthetic *Stenomeris* (which is currently placed in Dioscoreaceae; Angiosperm Phylogeny Group, 1998, 2003, 2009, 2016), despite numerous attempts using a range of herbarium and fixed material, and so we represented this genus with three available plastid genes obtained from GenBank (Appendix S1). Genome skims for nearly all newly sequenced Dioscoreales taxa were produced following approaches outlined by Lam et al. (2015), except that we prepared genomic libraries using two additional library preparation kits (Bioo NEXTflex Rapid DNA Sequencing Kit, Bioo Scientific Corp. [Austin, TX, USA] and NEBNext Ultra II DNA Library Prep Kit [New England Biolabs, Ipswich, MA, USA]). We sequenced these libraries as 100-bp, 125-bp, or 300-bp paired-end reads, respectively, on multiplexed lanes of HiSeq 2000, HiSeq 2500, or MiSeq platforms (Illumina, San Diego, CA, USA). We obtained plastid data for *Trichopus sempervirens* as genome-skim "by-catch" from a nuclear target enrichment gene panel (Soto Gomez et al., 2019) using approaches outlined by Viruel et al.

15372197, 0, Downloaded from https://bsapubs.onlinelibrary.wiley.com/doi/10.1002/ajb2.70267, Wiley Online Library on [08/10/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

TABLE 1 Nutritional status and approximate number of species in the 22 genera that comprise Dioscoreales. Sampled genera are shown in bold font; the remaining genera are marked as not available (N/A) in the column showing the number of species sampled. Note that *Burmannia* (*) includes both photosynthetic and fully mycoheterotrophic species, and that monotypic *Marthella* (†) may be extinct.

Genus	Family	Nutritional status		No. of species	
		Photosynthetic	Fully mycoheterotrophic	Sampled	Total
Afrothismia	Afrothismiaceae		X	4	12[a]
Apteria	Burmanniaceae		X	1	1[a]
Burmannia*	Burmanniaceae	X	X	5	56[a]
Campylosiphon	Burmanniaceae		X	1	2[a]
Dictyostega	Burmanniaceae		X	N/A	1[a]
Gymnosiphon	Burmanniaceae		X	1	32[a]
Hexapterella	Burmanniaceae		X	N/A	2[a]
Marthella†	Burmanniaceae		X	N/A	1[a]
Miersiella	Burmanniaceae		X	N/A	1[a]
Dioscorea	Dioscoreaceae	X		7	629[b]
Stenomeris	Dioscoreaceae	X		1	2[c]
Aletris	Nartheciaceae	X		3	24[d]
Lophiola	Nartheciaceae	X		1	1[d]
Metanarthecium	Nartheciaceae		X	1	1[d]
Narthecium	Nartheciaceae	X		1	8[d]
Nietneria	Nartheciaceae	X		1	2[d]
Tacca	Taccaceae	X		2	17[b]
Haplothismia	Thismiaceae		X	1	1[a]
Oxygyne	Thismiaceae		X	N/A	6[e]
Thismia	Thismiaceae		X	2	45[a]
Tiputinia	Thismiaceae		X	N/A	1[a]
Trichopus	Trichopodaceae	X		2	2[b]

[a]Merckx et al. (2013).
[b]Govaerts et al. (2021).
[c]Caddick and Wilkin (1998).
[d]Govaerts (2021).
[e]Cheek et al. (2018).

(2019), which we sequenced as 150-bp paired-end reads in a multiplexed lane of a HiSeq X platform.

Contig assembly and gene annotation

We used methods described by Lam et al. (2015) to assemble contigs from genome-skim data and to retrieve sets of plastid protein-coding and rDNA genes after filtering for possible mitochondrial and nuclear contigs, with minor modifications noted here. We used CLC Genomics Workbench (v.6.5.1; CLC Bio, Aarhus, Denmark) to assemble contigs for all but one taxon (*Trichopus sempervirens*, represented here using genome-skim "bycatch" data), for which we used NOVOPlasty v. 2.7.2 (Dierckxsens et al., 2017) with the *rbcL* plastid gene sequence of *Dioscorea elephantipes* (GenBank accession NC_009601.1) as a seed input, and otherwise default settings. All contigs assembled here have at least 30× average coverage regardless of assembly program. We identified plastid gene sets from *T. sempervirens* using GeSeq (Tillich et al., 2017) with default settings. For FM taxa with substantially elevated rates (i.e., *Afrothismia* sp., *A. kupensis*, *A. hydra*, *A. winkleri*, *Apteria aphylla*, *Gymnosiphon longistylus*, *Haplothismia exannulata*, *Thismia rodwayi*, and *T. tentaculata*), we additionally ran pairwise BLASTP searches (Altschul et al., 1990) on individual genes to retain regions with a default expect (e)-value threshold of 10 relative to homologous genes of photosynthetic taxa. We used genes from photosynthetic *Tacca leontopetaloides* as a query sequence for searches of mycoheterotrophic *Afrothismia*, *Haplothismia*, and *Thismia;* we used photosynthetic *Burmannia capitata* as a query for mycoheterotrophic *Apteria* and *Gymnosiphon*. Most of the

15372197, 0, Downloaded from https://bsapubs.onlinelibrary.wiley.com/doi/10.1002/ajb2.70267, Wiley Online Library on [08/10/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

recovered protein-coding genes for *Afrothismia* sp., *A. hydra*, and *A. winkleri* contain internal stop codons at one to eight different positions across individual genes and were slightly or moderately truncated compared to other sampled taxa. These internal stop codons may be undocumented RNA-edited sites (e.g., Takenaka et al., 2013; Bell et al., 2020) or instances of pseudogenization. By contrast, the retained genes of *A. kupensis* were completely uninterrupted, although most were either truncated or slightly longer relative to other sampled taxa. We included all of the *Afrothismia* sequences in phylogenetic analyses described below. We also quantified nucleotide composition for individual taxa using AMAS (Borowiec, 2016).

Sequence alignment and concatenated matrix construction

We assembled a 98-taxon matrix by adding species that were newly sequenced here to published data retrieved from GenBank and two published matrices (Lam et al., 2018; Garrett et al., 2023). To do so, we generated 82 individual plastid gene files, representing 78 protein-coding genes and four rDNA genes, each with 98 taxon terminals. We produced DNA-based alignments for each gene, including manual adjustments, and prepared a concatenated matrix as described by Soto Gomez et al. (2020) for the plastid data set in that study. To help with manual adjustments for aligning FM taxa, we additionally used the results from pairwise BLASTP searches described above. We applied procedures used by Lam et al. (2015) to check for matrix compilation errors; none were found. Photosynthetic taxa are represented by all or most genes found in plastid genomes of photosynthetic angiosperms, except *Stenomeris borneensis*, for which we only included three available protein-coding genes (i.e., *atpB*, *matK*, *rbcL*; Appendix S1). Fully mycoheterotrophic taxa are represented by 6–24 protein-coding genes (counting putative pseudogenes in *Afrothismia*) and 2–4 rDNA genes, due to gene loss (Figure 1). We coded missing genes (i.e., those lost in FM taxa or the few that were not recovered for photosynthetic taxa) as missing data. The gene sets retrieved for each of the FM taxa are likely complete or nearly so, based on preliminary plastid genome assemblies to be presented elsewhere. We retrieved the *ycf1* locus for most photosynthetic taxa and for the FM species *Campylosiphon congestus* and *Burmannia itoana*, but did not include it in the final matrix due to difficulties aligning it across the full taxon set. The final DNA-based matrix is 80,390-bp long (derived from 67,156 bp of unaligned plastid sequence data in *Dioscorea membranacea*, for reference). We also translated this into a matrix comprising 24,980 amino-acid residues. Data matrices are available on Figshare (https://figshare.com/s/8470257673242e6bf43f).

Characterization of substitution rate differences in Dioscoreales

We characterized relative differences in the substitution rates of mycoheterotrophic and photosynthetic taxa using a Bayesian framework implemented in BEAST v. 2.6.3 (Bouckaert et al., 2019). We prepared a reduced and unpartitioned data set to keep these computationally demanding analyses tractable in terms of the number of genes and taxa. We selected seven plastid genes (*accD*, *rpl2*, *rps12*, *rps*4, *rps8*, *rrn16*, *rrn23*) that were available for a broad taxon sampling. After removing taxa for which these genes were not available, we obtained a 62-taxon data set that includes at least one representative of each of the sampled

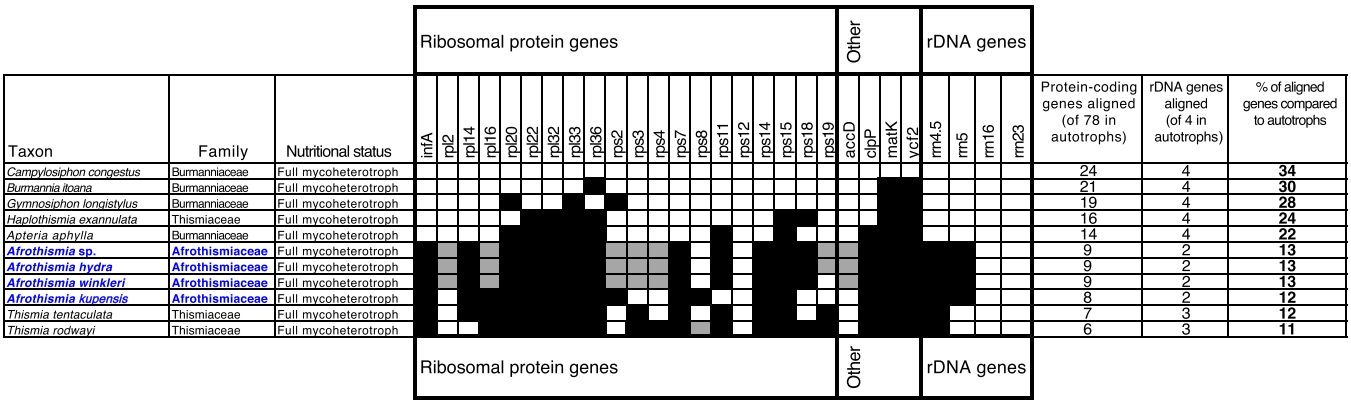


FIGURE 1 Fully mycoheterotrophic Dioscoreales sampled retain only a subset of plastid genes, summarized for loci included in alignments here (i.e., excluding *ycf1* and tRNA genes). Species are ordered vertically by number of genes retained, and genes horizontally by functional complex; gene nomenclature follows Wicke et al. (2011). Newly sequenced taxa are shown in blue. White box = gene present, black = gene lost (or possibly not retrieved from taxa with preliminary plastid genome assemblies), grey = genes containing internal stop codons. Note that *ycf1* was recovered in the Burmanniaceae species *C. congestus* and *B. itoana* (Lam et al., 2016) but not in Thismiaceae (Garrett et al., 2023) or Afrothismiaceae. A range of tRNA genes (excluding pseudogenized copies) were recovered from Burmanniaceae species (*C. congestus*: 23, *B. itoana*: 11, *G. longistylis*: 5, *A. aphylla*: 3; Lam et al., 2016), from Thismiaceae species (*H. exannulata*: 7, *T. tentaculata*: 2, *T. rodwayi*: 1; Garrett et al., 2023) and from Afrothismiaceae species (*Afrothismia* sp., *A. hydra*, *A. winkleri*, *A. kupensis*: 1).

15372197, 0, Downloaded from https://bsapubs.onlinelibrary.wiley.com/doi/10.1002/ajb2.70267, Wiley Online Library on [08/10/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

mycoheterotrophic genera. To focus on characterizing rate variation, we constructed an ultrametric input tree with a fixed topology. We based the topology on the best likelihood tree recovered in the mitochondrial data inferences of Lin et al. (2022) for Dioscoreales; for the remaining taxa, we based it on the best likelihood tree recovered in plastid-based inferences of the core taxon set here. We used the chronos function in the R package ape v. 5.6-2 (Paradis et al., 2004; Popescu et al., 2012; Paradis and Schliep, 2019) with default settings to make the tree ultrametric. Our analysis focuses on relative rates (e.g., Lam et al., 2018), and so we did not specify fossil constraints or attempt to date individual nodes, which may lead to artifactual results when highly rate-elevated taxa are included (e.g., Iles et al., 2015; Soto Gomez et al., 2020). We used the ModelFinder function (Kalyaanamoorthy et al., 2017) in IQ-TREE 1.6.12 (Nguyen et al., 2015) to find an optimal substitution model for the entire reduced (unpartitioned) data set, considering invariant (I) and gamma (G) rate heterogeneity across sites (-m TESTMERGEONLY). We obtained GTR with four categories as the best model, which we used in BEAST, and additionally specified for substitution rates and empirical frequencies to be estimated under a birth–death model and a random local clock (Drummond and Suchard, 2010); we otherwise used default settings. We ran six independent analyses, each for 400 million generations, sampling every 2000 trees, and visually assessed stationarity and convergence amongst them using Tracer 1.7.1 (Rambaut et al., 2018). We used the LogCombiner program in BEAST to combine five analyses that converged and considered combined posterior ESS values > 200 to be acceptable, which was the case for all parameters. We used the TreeAnnotator program to summarize the resulting sampling of trees after discarding 20% as burn-in and resampling at a lower frequency to obtain 10,000 trees that were then used to make a single combined tree.

Phylogenetic analyses using partitioned likelihood models

We conducted most phylogenetic analyses using a maximum likelihood approach with partitioned data as implemented in RAxML-NG (Kozlov et al., 2019). We ran multiple partitioned analyses using variant versions of the data matrix. However, the first two of these analyses were based on a core 89-taxon version of the matrix that includes (1) most photosynthetic taxa in Dioscoreales, including those in Burmanniaceae (i.e., *Burmannia bicolor*, *B. capitata*, *B. coelestis*, and *B. disticha*) and (2) three of the least rapidly evolving FM taxa sampled here (i.e., *Burmannia itoana* and *Campylosiphon congestus* in Burmanniaceae; *Haplothismia exannulata* in Thismiaceae). Thus, this core 89-taxon data set excludes all VLB taxa (see below). It also excludes *Stenomeris borneensis*, a data-deficient photosynthetic taxon. We selected the three least rapidly evolving FM taxa after assessing rate elevation in an analysis of rate variation (see Figure 2 and below) and by visual inspection of branch-length estimates in likelihood analyses here. We ran partitioned DNA- and AA-based analyses on this core 89-taxon set and used the resulting trees as references to examine the effect of nine potentially problematic taxa on phylogenetic inference.

Eight of the nine problematic taxa are highly rate-elevated mycoheterotrophs in Afrothismiaceae, Burmanniaceae, and Thismiaceae (see below); the ninth is the data-deficient photosynthetic taxon (*Stenomeris*), for which only three plastid genes are available. We assessed the placement of the nine potentially problematic taxa by adding one or more of these excluded species to the core 89-taxon data set (in separate DNA-based partitioned likelihood analyses). The taxon changes to the core 89-taxon matrix included adding only (independently), in turn: (1) the data-deficient photosynthetic taxon, *S. borneensis*, resulting in a 90-taxon matrix; (2) two VLB representatives of Burmanniaceae, *Apteria aphylla* and *Gymnosiphon longistylus*, resulting in a 91-taxon matrix; (3) two VLB representatives of Thismiaceae, *Thismia rodwayi* and *T. tentaculata*, resulting in a second 91-taxon matrix; (4) four VLB species of Afrothismiaceae—by far the most rate-elevated lineage that we surveyed—resulting in a 93-taxon matrix; (5) the only sampled member of Afrothismiaceae lacking internal stop codons in protein-coding genes, *A. kupensis*, resulting in a second 90-taxon matrix; and (6) all Dioscoreales taxa with molecular data here, resulting in a "full" taxon set comprising 98 species. We additionally analyzed the full 98-taxon matrix in a partitioned analysis based on amino-acid data.

For these analyses, we partitioned DNA alignments by considering gene and codon position (a "G × C" scheme, e.g., Lam et al., 2015) for the 78 protein-coding genes, with additional starting partitions for each of the four rDNA genes, resulting in 238 initial partitions. The amino-acid (AA) matrix had 78 initial partitions, corresponding to one partition per protein-coding gene. We used the ModelFinder function (Kalyaanamoorthy et al., 2017) in IQ-TREE 1.6.12 (Nguyen et al., 2015) to combine partitions that did not have substantially different substitution models, employing the relaxed hierarchical clustering algorithm to examine the top 10% partition merging schemes (-rcluster 10), the corrected Akaike information criterion (AICc) to merge partitions, and considered invariant (I) and gamma (G) rate heterogeneity across sites (-m TESTMERGEONLY). We obtained a range of optimal DNA and AA substitution models for the resulting partitioning schemes (Appendix S2), which we used in subsequent likelihood analyses.

The DNA- and AA-based partitioned likelihood analyses performed using RAxML-NG included 40 independent searches for the best tree, starting from 20 parsimony-based and 20 random trees. We estimated branch support using bootstrap analysis (Felsenstein, 1985) by applying the extended majority rule (MRE) bootstrap stopping test

15372197, 0, Downloaded from https://bsapubs.onlinelibrary.wiley.com/doi/10.1002/ajb2.70267, Wiley Online Library on [08/10/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

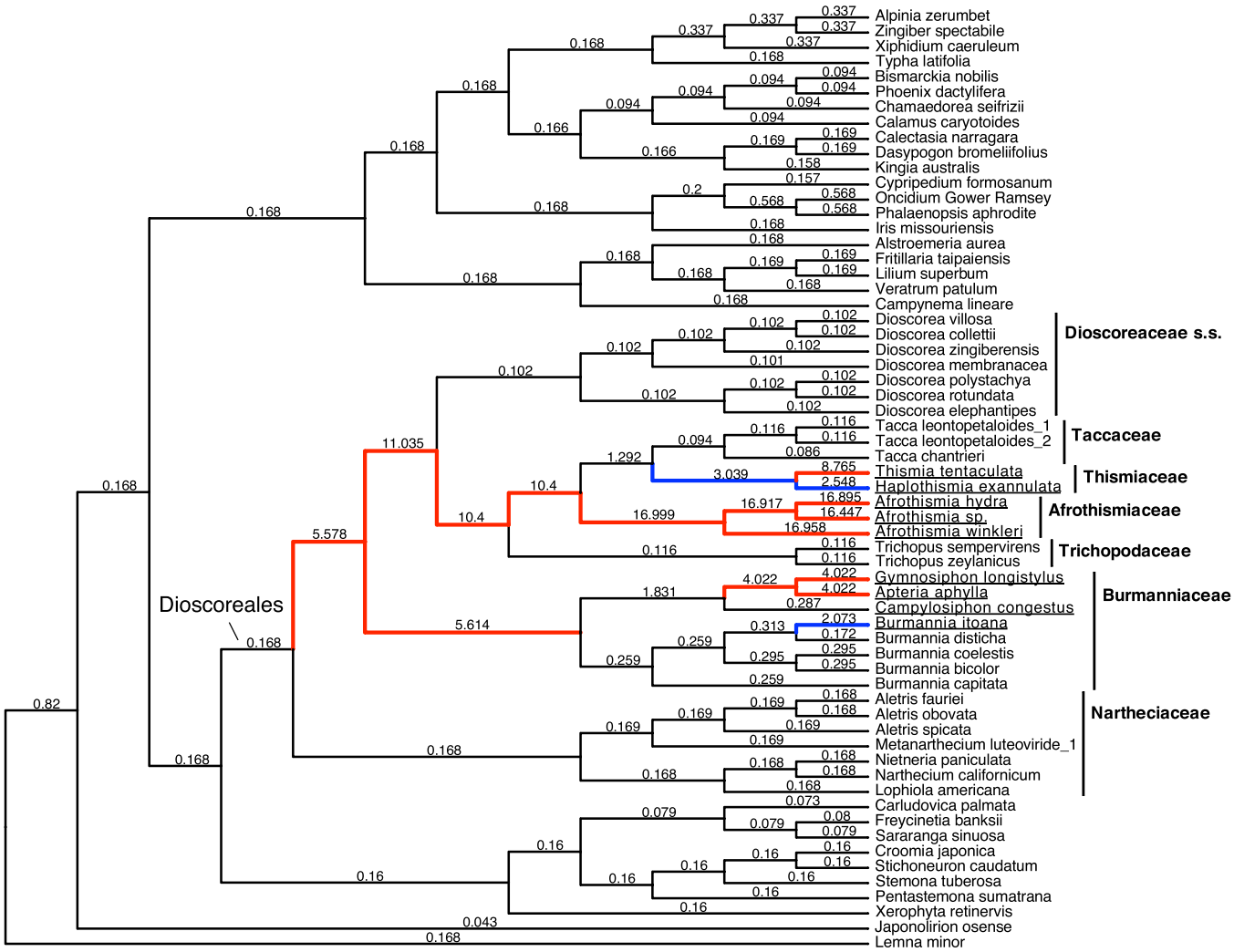


FIGURE 2 Fully mycoheterotrophic members of Dioscoreales display a range of elevated DNA substitutional rates in plastid genes, as determined in Bayesian analyses of relative rates using a random local clock model for a reduced data set of monocots (seven plastid genes available for 62 taxa, using a fixed topology based on Figure 4 of Lin et al., 2022). The estimated number of substitutions per site is shown for individual branches; the names of fully mycoheterotrophic taxa are underlined; see Appendices S15 and S16 for a summary of 95% highest posterior density values of relative rates. Blue branches indicate intermediate rates that are two–four times faster than those of photosynthetic taxa; red branches indicate rates that are over four times faster. We refer to the latter as very long branch taxa.

(Pattengale et al., 2010) every 50 replicates. This test consists of randomly splitting the set of bootstrap replicates into two equal halves (with 1000 permutations per test), computing the majority-rule consensus trees for each half, and using the weighted Robinson–Foulds (WRF) criterion to assess dissimilarity between the two resulting consensus trees. Bootstrapping is stopped when the relative WRF between sets is ≤1% for at least 990 permutations (--bs-cutoff 0.01). Here this resulted in 100–1000 standard/slow bootstrap replicates across the eight variant likelihood analyses. Each bootstrap analysis used the same partitioning schemes and substitution models as the corresponding searches for best trees. For all bootstrap analyses performed here, we considered weakly and strongly supported branches to respectively have <70% and ≥95% support (e.g., Soltis and Soltis, 2003).

Phylogenetic analyses using site-heterogeneous models

We also analyzed AA data using two model-based approaches thought to ameliorate the effects of long branch artifacts (Lartillot et al., 2007; Quang et al., 2008; Wang et al., 2018). We implemented these using IQ-TREE 1.6.12 (Nguyen et al., 2015) for likelihood inference and PhyloBayes MPI 1.9 (Lartillot et al., 2013) for Bayesian inference. For IQ-TREE analyses, we used a posterior mean site frequency (PMSF) model with the most complex empirical mixture model implemented in that program, which includes 60 AA mixture classes, empirical state frequencies, and a discrete gamma distribution with four rate categories (-m LG + C60 + F + G). In PMSF analyses, AA profiles are estimated for each site in an alignment based on a guide tree

15372197, 0, Downloaded from https://bsapubs.onlinelibrary.wiley.com/doi/10.1002/ajb2.70267, Wiley Online Library on [08/10/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

(estimated here using a simpler model in IQ-TREE, -m LG + F + G) and an input mixture model (estimated in the first phase of the analysis). The PMSF model has been shown to ameliorate the effects of long branch artifacts in simulated and empirical phylogenomic analyses (Wang et al., 2018), while providing a rapid approximation of the more computationally intensive profile mixture models (Quang et al., 2008). We conducted 1000 ultrafast bootstrap replicates to estimate branch support for this approach.

For AA-based analyses using PhyloBayes MPI, we implemented the CAT-GTR model (using a discretized gamma distribution with four categories), which captures variation in AA preferences across sites using a Bayesian nonparametric approach (Lartillot, 2020). We ran two independent chains until they reached ~10,200 points (i.e., iterations) and quantitatively assessed their convergence and mixing using the commands tracecomp and bpcomp, respectively, after setting a burn-in of 1000. We considered ESS values >100 and discrepancies among chains <0.3 to represent acceptable runs (Lartillot, 2020), which was the case for all model parameters. We obtained a consensus tree with posterior probabilities for branch support by pooling the trees produced by the two chains.

Constraint tests of alternative phylogenetic hypotheses

We performed Shimodaira–Hasegawa (SH) and approximately unbiased (AU) tests (Shimodaira and Hasegawa, 2001) to assess variant placements for VLB mycoheterotrophs. Specifically, we tested whether several alternative arrangements of these mycoheterotrophs were significantly worse than optimal trees from analysis of the core 89-taxon set (i.e., alternative hypothesis 1 below; compare Figure 3 and Appendices S3–S5) or from analysis of an expanded 93-taxon set that included Afrothismiaceae (i.e., alternative hypotheses 2–5 below; compare Appendix S6 to Appendices S7–S10). The minimal constraints used in each test (see below) are indicated with red branches in Appendices S5 and S7–S10. We used family-level boundaries based on narrow-sense circumscriptions (as shown in Table 1, following Merckx et al., 2010; Givnish et al., 2018; Lam et al., 2018; Cheek et al., 2024), and we follow these throughout the study, except where noted. We included *Haplothismia* as a slowly evolving mycoheterotroph to represent Thismiaceae in these tests and treated Afrothismiaceae as a lineage distinct from the rest of Thismiaceae when included (following the treatment by Cheek et al., 2024). *Stenomeris*, a photosynthetic taxon with uncertain family placement represented here by only three genes, was excluded from all constrained analyses.

For alternative hypothesis 1, we constrained all members of Burmanniaceae and Thismiaceae to form a clade (note that we excluded Afrothismiaceae from consideration in this case to remove the possible distorting effect of its extreme rate elevation, e.g., Appendices S6, S11). Thus, ignoring Afrothismiaceae, this constraint is consistent with the current broad circumscription of Burmanniaceae in recent angiosperm classifications that incorporates Thismiaceae (Angiosperm Phylogeny Group, 2003, 2009, 2016). We compared the resulting tree to a null hypothesis from the corresponding unconstrained search, in which Burmanniaceae and Thismiaceae are inferred to be distantly related clades in Dioscoreales (see Figure 3; Appendices S3, S4). The remaining tests consider possible placements of Afrothismiaceae by constraining clades that comprise (2) Afrothismiaceae and Dioscoreaceae; (3) Afrothismiaceae and Burmanniaceae, the latter narrowly defined; (4) Afrothismiaceae and Thismiaceae (= Thismiaceae s.l.); and (5) Afrothismiaceae, Thismiaceae, and Taccaceae, with the last two families forced to be sister taxa to reflect the topology recovered by Lin et al. (2022). Hypotheses 2–5 each assess different possible arrangements of Afrothismiaceae in Dioscoreales that we observed in phylogenetic analyses here (i.e., hypotheses 2 and 3; see Figure 4, Appendices S11–S14), or that are consistent with a possible broad circumscription of Thismiaceae to include Afrothismiaceae (i.e., hypothesis 4; see Maas-van de Kamer, 1998), or with a possible placement of Afrothismiaceae as a close relative of Thismiaceae and Taccaceae (i.e., hypothesis 5; see Merckx et al., 2009b, 2010; Lin et al., 2022). We compared these four alternative hypotheses to a main hypothesis recovered in phylogenetic analyses here, where Afrothismiaceae forms a clade with the family Poaceae in the order Poales (see Appendix S6).

In all cases, we tested whether we could reject trees recovered under constrained hypotheses compared to the unconstrained optimal trees for that taxon set. For tree searches using constrained branches, we excluded data-deficient and VLB taxa that were not relevant to the hypothesis. Thus, we employed the core 89-taxon set for test 1, and a modified version of the 93-taxon set including only two species of Afrothismiaceae for tests 2–5. We did the latter reduction because three of the four Afrothismiaceae taxa are nearly identical at the sequence level. To find trees consistent with these five hypotheses, we made constraint trees for tests 1–4 by constraining one branch to obtain two taxon partitions, one comprising the species of interest, and the other comprising the remaining taxa (highlighted branches in Appendices S5, S7–S9). For test 5, we constrained two branches to simultaneously obtain two taxon partitions, one comprising Thismiaceae-Taccaceae, and the other comprising Afrothismiaceae-Thismiaceae-Taccaceae (in this case effectively placing Afrothismiaceae close to the other two families, but not allowing it to be inside or sister to either family individually; see highlighted branches in Appendix S10). We recovered the best likelihood trees that satisfied each of the five constraints using the tree-constraint function in RAxML-NG, with default settings. We then performed the SH and AU tests in CONSEL 0.20 (Shimodaira and Hasegawa, 2001), using site likelihoods from partitioned likelihood analyses to address whether the fit of the data was significantly different for unconstrained vs. constrained best trees.

15372197, 0, Downloaded from https://bsapubs.onlinelibrary.wiley.com/doi/10.1002/ajb2.70267, Wiley Online Library on [08/10/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

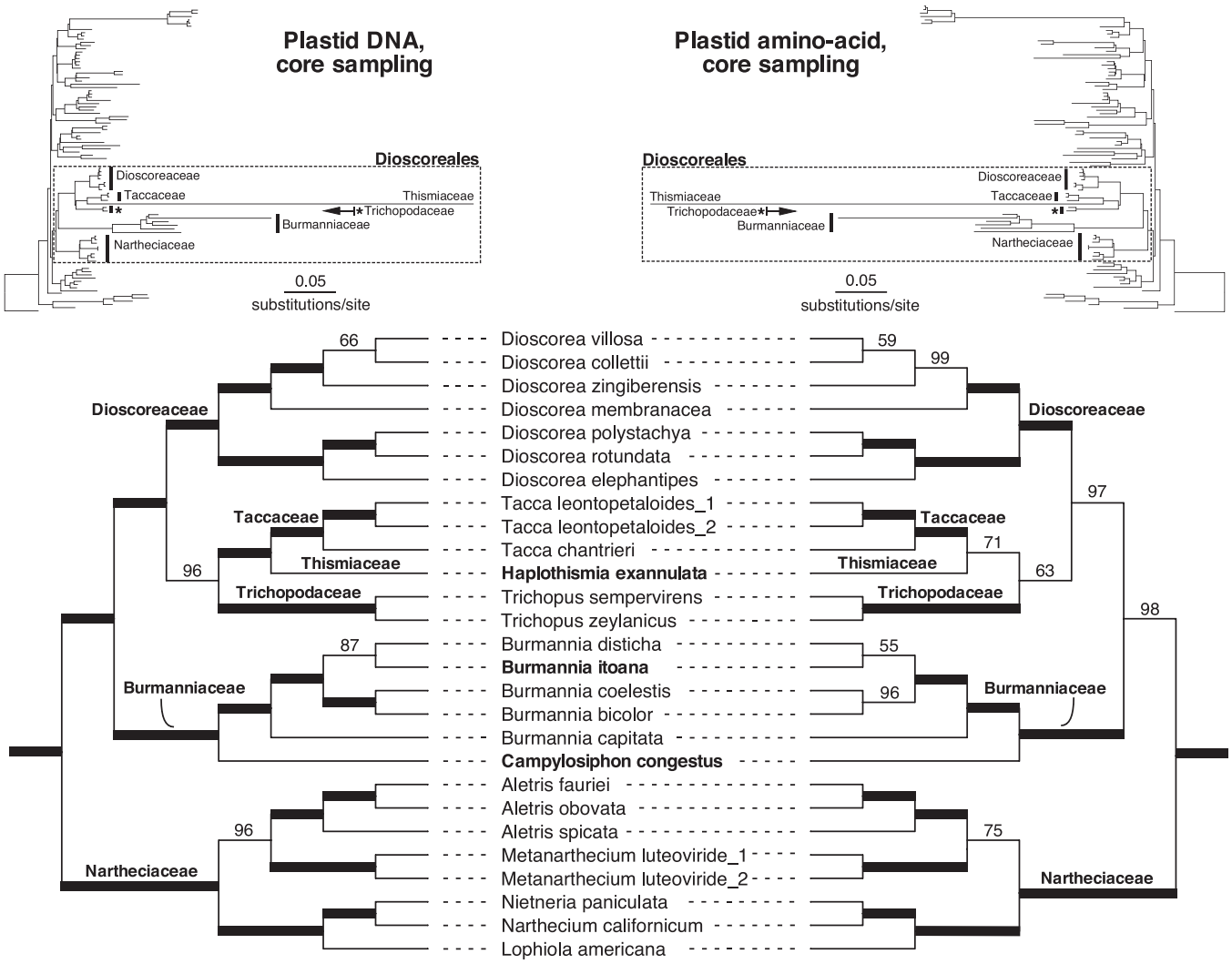


FIGURE 3 Highly congruent phylogenetic estimates recovered for Dioscoreales based on DNA vs. amino-acid (AA)-based likelihood analyses of a core 89-taxon set. Three relatively slowly evolving fully mycoheterotrophic taxa in Burmanniaceae and Thismiaceae were included in this core set (noted here in bold font, and see relative rates depicted in Figure 2). Both analyses exclude very long branch mycoheterotrophs and a data-deficient photosynthetic species. Thick lines indicate 100% bootstrap support; values <100% are shown above branches. Upper phylograms depict full samplings of each analysis (see Appendices S3, S4); scale bars indicate estimated substitutions per site.

RESULTS

Rate elevation in fully mycoheterotrophic lineages of Dioscoreales

The Bayesian analysis characterizing relative substitution rates found moderate to very high levels of rate elevation in the nine included FM taxa (*Afrothismia* sp., *A. hydra*, and *A. winkleri* in Afrothismiaceae; *Haplothismia exannulata* and *Thismia tentaculata* in Thismiaceae; *Apteria aphylla*, *Burmannia itoana*, *Campylosiphon congestus*, and *Gymnosiphon longistylus* in Burmanniaceae) compared to photosynthetic relatives (Figure 2). Our initial likelihood-based phylogenetic inference using the core taxon sampling included three of the nine FM species characterized in Bayesian analyses of relative substitution rates (Figure 3: *C. congestus* and *B. itoana* in Burmanniaceae; *H. exannulata* in Thismiaceae). Two of these FM taxa (*B. itoana* and *H. exannulata*) exhibited moderate but statistically significant substitutional elevation, which we defined as rates up to around two to four times higher in their corresponding terminal lineages (or on nearby subtending branches) compared to photosynthetic relatives (Figure 2; Appendices S15, S16). The rate of *C. congestus* was comparable to that of photosynthetic taxa. The six remaining FM species, which we defined as very long branch (VLB) taxa and therefore excluded from the core taxon set, all exhibited substantially higher substitution rate elevation, with the very highest rates (> 16-fold) in Afrothismiaceae taxa (Figure 2). The rates of these six FM species were significantly higher than those of photosynthetic taxa. However, only the rates of the Afrothismiaceae taxa and *T. tentaculata* were significantly higher than those of FM taxa with moderate rates because the 95% highest posterior densities (HPDs) of *G.*

15372197, 0, Downloaded from https://bsapubs.onlinelibrary.wiley.com/doi/10.1002/ajb2.70267, Wiley Online Library on [08/10/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

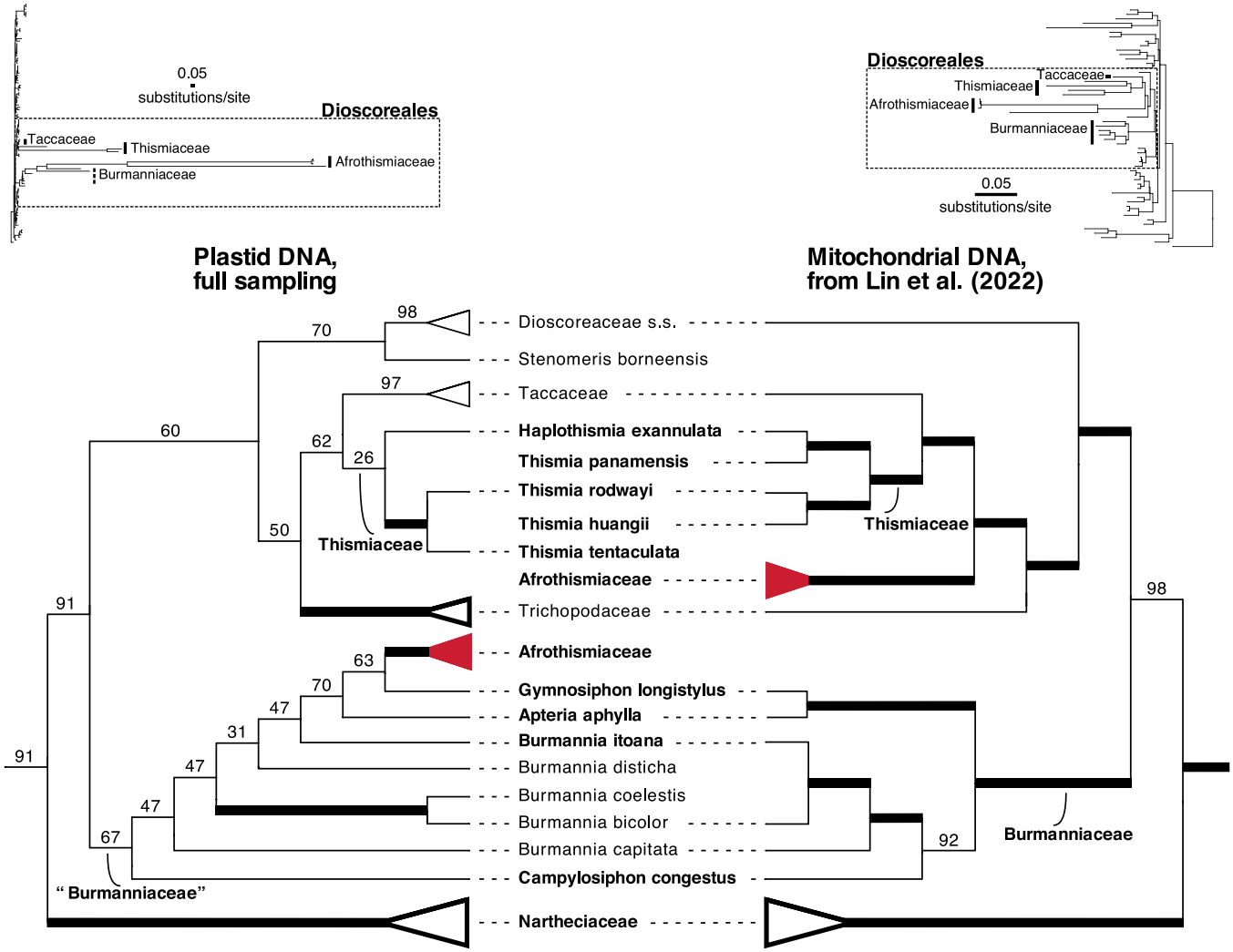


FIGURE 4 Highly divergent placements of Afrothismiaceae (red clade) in plastid vs. mitochondrial phylogenomic analysis (left vs. right trees, respectively). Both trees were inferred using partitioned DNA-based maximum likelihood analyses. The plastid tree was inferred using the full 98-taxon sampling of Dioscoreales, which included very long branch mycoheterotrophs and a data-deficient photosynthetic species (*Stenomeris*), based on 82 plastid genes (7–28 genes for fully mycoheterotrophic taxa, in bold font). The mitochondrial tree was inferred by Lin et al. (2022, their Fig. 4) based on 37 mitochondrial genes. Nartheciaceae and all monogeneric families represented by more than one species (Afrothismiaceae, Dioscoreaceae, Taccaceae, Trichopodaceae) are depicted here as triangles. Thick lines indicate 100% bootstrap support; values < 100% are shown above branches. Upper phylograms depict full samplings of each analysis; the family Burmanniaceae, inferred to be nonmonophyletic in the DNA analysis, is indicated in the upper left phylogram with a vertical dotted line (see Appendix S12 here and Fig. 4 of Lin et al., 2022 for further detail). Scale bars indicate estimated substitutions per site in the phylograms.

longistylus and *A. aphylla* overlapped with those of *B. itoana* and *H. exannulata* (Figure 2; Appendices S15, S16). Nonetheless, to be conservative, we categorized *G. longistylus* and *A. aphylla* as VLB taxa too, based on their mean predicted rates (Figure 2) and relatively longer branches in likelihood-based phylogenetic inference (e.g., Appendix S17).

Curiously, several internal branches with high rates were inferred to occur near the origin of Dioscoreales. This may occur if the algorithm wrongly inferred shared rate elevation in the common ancestors of lineages that independently evolved high rates. Although these estimates have wide 95% HPDs that mostly overlap with those of photosynthetic taxa, it may imply that rates in some terminal branches were underestimated in the analysis (Figure 2; Appendices S15 and S16; see also the plastid-based phylograms in Figures 3 and 4 and Appendices S6, S11–S14, S17, S18). Regardless, based on the high rate of elevation inferred here for multiple VLB taxa in Afrothismiaceae, Thismiaceae, and Burmanniaceae (Figure 2), we implemented taxon-addition phylogenetic experiments to sequentially (and independently) add these lineages to the core taxon set comprising photosynthetic species and more slowly evolving mycoheterotrophs (Appendices S6, S11, S17, S18).

Phylogenetic analyses of the core taxon set (excluding VLB taxa and *Stenomeris*)

The trees inferred from DNA- and AA-based likelihood analyses of the core 89-taxon data set were congruent with each other and were also generally very well supported

15372197, 0, Downloaded from https://bsapubs.onlinelibrary.wiley.com/doi/10.1002/ajb2.70267, Wiley Online Library on [08/10/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

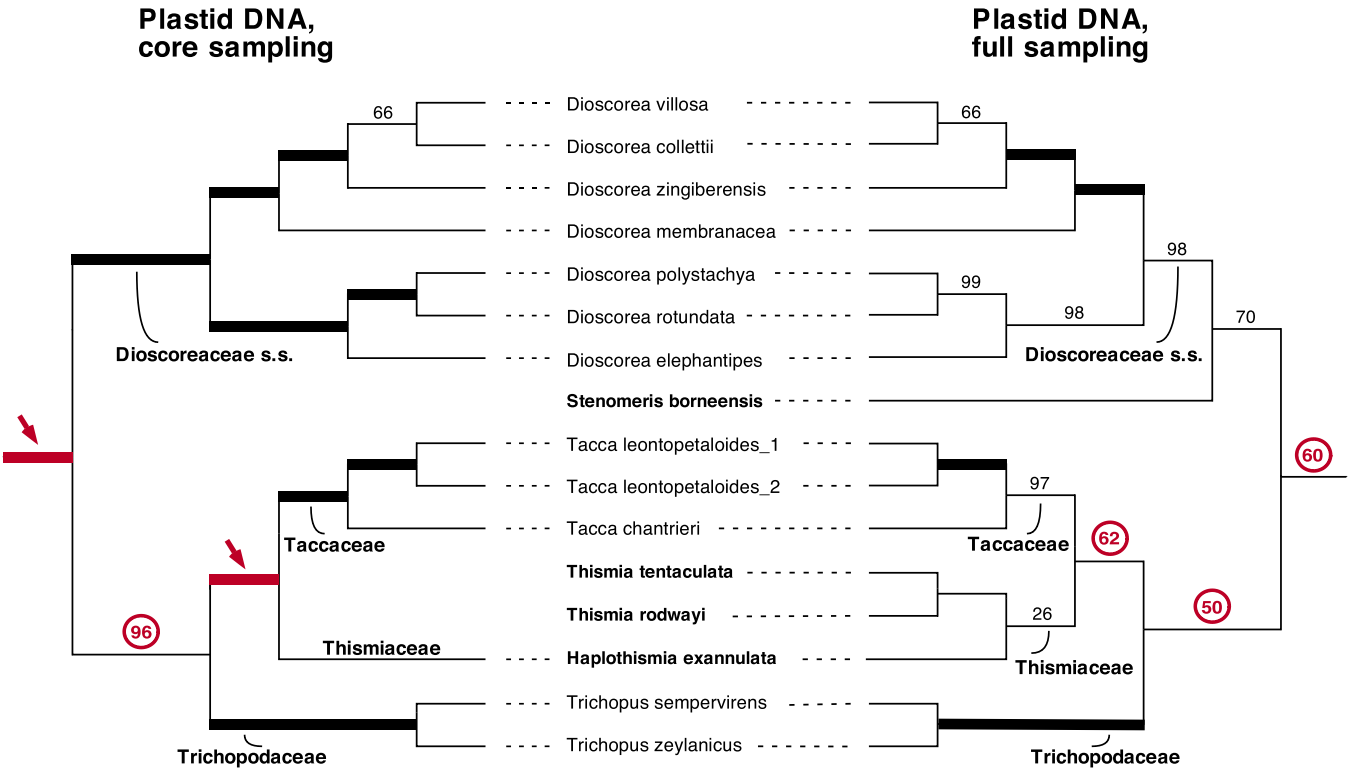


FIGURE 5 Strong depression of bootstrap support values, with consistent placement of Thismiaceae, in DNA-based maximum likelihood analysis, following inclusion of very long branch (VLB) mycoheterotrophs in Dioscoreales. Relationships were inferred in partitioned likelihood analyses of plastid DNA using either the core (89-taxon) vs. full (98-taxon) samplings (left vs. right trees, respectively). Support values are indicated on branches (thick lines indicate 100% bootstrap support); red lines/arrows/circles indicate branches that are well supported in analysis of the core taxon set that have a substantial depression in support in the analysis of the full taxon set. Fully mycoheterotrophic taxa and the data-deficient photosynthetic taxon *Stenomeris* are shown in bold font. The full-taxon analysis includes all VLB mycoheterotrophs and *Stenomeris*, in addition to the taxa included in the core taxon set (see Materials and Methods: Phylogenetic analyses using partitioned likelihood models). Two Dioscoreales families, Burmanniaceae and Nartheciaceae, are omitted for clarity but were included in the analyses (see Appendices S3 and S12 for full relationships and Figure 3 for a view of the core taxon sample across Dioscoreales).

regarding higher-level relationships in Dioscoreales (Figure 3; see Appendices S3 and S4 for monocot-wide relationships from DNA and AA inferences, respectively). Well-supported relationships, which were consistent between DNA and AA analyses, include the placements of the three most slowly evolving FM lineages from Burmanniaceae (*Burmannia itoana* and *Campylosiphon congestus*) and Thismiaceae (*Haplothismia exannulata*) (Figure 3; Appendices S3, S4). The FM family Thismiaceae—represented by *Haplothismia* alone in the core set here—was part of a clade with Taccaceae, and both families were in turn sister to Trichopodaceae among sampled taxa (Figure 3; Appendices S3, S4). This arrangement had strong support from DNA data and moderate support from AA data. These three families (Taccaceae, Thismiaceae, and Trichopodaceae) were strongly supported as part of a larger clade with Dioscoreaceae in both analyses. Thus, the family Thismiaceae was not associated with Burmanniaceae, and there was strong bootstrap support for their separation. Burmanniaceae, represented in the core set by two FM and four photosynthetic taxa, were instead sister to all families of Dioscoreales except Nartheciaceae (Figure 3; Appendices S3, S4). These family-level relationships in Dioscoreales all had moderate or strong support from DNA and AA analyses, as did most genus-level relationships (Figure 3; Appendices S3, S4). Outside Dioscoreales, the monocot-wide relationships were also both congruent and generally well supported across analyses of the core taxon set (Appendices S3, S4). For example, Dioscoreales was consistently strongly supported as the sister group of Pandanales, with 100% bootstrap support for this relationship (Appendices S3, S4).

Effect of including a data-deficient photosynthetic species

The placement of photosynthetic *Stenomeris borneensis* was consistent across analyses that included this species. It is represented here by only three available plastid genes from GenBank (Figure 5; Appendices S12–S14, S19). We recovered *Stenomeris* as the sister group of *Dioscorea* (Dioscoreaceae) with poor to moderate support (67–70%) in DNA analyses when added individually to the core taxon

15372197, 0, Downloaded from https://bsapubs.onlinelibrary.wiley.com/doi/10.1002/ajb2.70267, Wiley Online Library on [08/10/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

set, or when included as part of the full 98-taxon set (Figure 5, right-hand tree; Appendices S12, S19). There was at best moderate support (80–81%) for its placement in AA analyses of the full taxon set using site-heterogeneous models in both likelihood and Bayesian frameworks (Appendices S13, S14).

Taxon-addition experiments with VLB mycoheterotrophic lineages

Adding individual (or related sets of) VLB mycoheterotrophs in Burmanniaceae and Thismiaceae to analyses yielded generally congruent relationships in Dioscoreales and across monocots, relative to the core taxon set (compare Figure 3 and Appendices S3, S4 to Appendices S17, S18). Despite this, most branches neighboring points of attachment of VLB lineages to the tree—and often more distant ones—experienced reductions in bootstrap support compared to the DNA-based analysis of the core taxon set (Table 2, compare Figure 3 to Appendices S17, S18). In a taxon-addition experiment in which only the two VLB mycoheterotrophs in Burmanniaceae (*Apteria* and *Gymnosiphon*) were added to the core set, these two taxa formed a strongly supported clade that was in turn weakly supported (52%) as the sister group of *Campylosiphon* (Table 2, Appendix S17). Three of the five remaining branches in Burmanniaceae then exhibited substantial depression in bootstrap support of 41%–47%, compared to DNA-based analysis of the core taxon set (Table 2, compare Figure 3 to Appendix S17). In a taxon-addition experiment in which only the two VLB mycoheterotrophs in Thismiaceae (*Thismia rodwayi* and *T. tentaculata*) were added to the core set, both *Thismia* and *Haplothismia* were inferred to be closely related to *Tacca* in a poorly supported grade (rather than a clade), with 44% bootstrap support (Table 2, Appendix S18). The addition of these two species did not depress branch support substantially (Table 2, compare Figure 3 to Appendix S18).

When Afrothismiaceae were included, this lineage appeared to be attracted to distantly related sets of taxa and broadly reduced bootstrap support across the tree (Table 2, compare Figure 3 and Appendices S3, S4 to Appendices S6, S11). First, in a taxon-addition experiment in which only the four sampled members of Afrothismiaceae were added to the core set, these taxa were inferred to be the sister group of Poaceae in the order Poales—a clade with one of the longest branches among photosynthetic lineages—although with very poor bootstrap support (14%) for this arrangement (Table 2, Appendix S6). There were also substantial drops in bootstrap support along the entire "backbone" of the inferred monocot phylogeny, including within Burmanniaceae (compare Figure 3 and Appendices S3, S4 to Appendix S6). Adding only *Afrothismia gesneriodes* to the core set appeared to be less disruptive to support values across the monocot phylogeny, including within Burmanniaceae (compare Figure 3 and Appendices S3, S4 to Appendix S11). In this taxon-addition experiment, Afrothismiaceae instead placed within *Dioscorea*, again with poor support (35%; Table 2, Appendix S11).

C60 and CAT-GTR approaches compared to partitioned ML analyses

Likelihood and Bayesian approaches designed to ameliorate composition-related effects for challenging long branch taxa did not yield substantially different placements of problematic VLB mycoheterotrophs compared to a standard partitioned ML analysis that included all 98 taxa sampled here (Appendices S12–S14). In all of these AA-based analyses, members of Thismiaceae formed either a clade (partitioned ML analysis and PhyloBayes CAT-GTR analysis; Appendices S12, S14) or a grade that was closely related to Taccaceae (ML C60 analysis; Appendix S13), with moderate to poor bootstrap support for these arrangements. Afrothismiaceae were moderately to weakly supported as nested in Burmanniaceae (sister to the longest branch taxon there, *Gymnosiphon*), in all three analyses.

Constraint tests of alternative placements of Thismiaceae and Afrothismiaceae

We can strongly reject alternative hypothesis 1. This hypothesis included a constrained clade comprising Burmanniaceae and Thismiaceae (as suggested in Angiosperm Phylogeny Group, 2016; see Appendix S5), compared to the unconstrained (best) likelihood tree (Figure 3), in which these two lineages are recovered as distant relatives in Dioscoreales ($P = 0.001$ for both AU and SH tests; Table 3). We cannot reject any of the other constrained hypotheses (2–5; Appendices S7–S10), which tested the placement of Afrothismiaceae in various positions compared to its unconstrained (presumably artifactual) placement as the sister group of Poaceae (Appendix S6), using either AU or SH tests. These included hypothesis 2: an alternative position of Afrothismiaceae within or as the sister group of Dioscoreaceae ($P = 0.483$ and $P = 0.477$ for AU and SH tests, respectively), as suggested in Appendix S11; hypothesis 3: Afrothismiaceae within or as the sister group of Burmanniaceae ($P = 0.446$ and $P = 0.442$), as suggested in Appendices S12–S14; hypothesis 4: Afrothismiaceae within (or as the sister group of) Thismiaceae, represented by *Haplothismia* ($P = 0.126$ and $P = 0.145$), as suggested in previous studies (e.g., Maas-van de Kamer, 1998); and hypothesis 5: Afrothismiaceae as the sister group of a clade comprising Taccaceae and Thismiaceae ($P = 0.15$ and $P = 0.155$), as suggested in previous studies (Merckx et al., 2009b and Lin et al., 2022) (see Table 3 for a summary of the five hypotheses and their constraint tests). Thus, these tests indicate that plastid phylogenomic data lack power to place Afrothismiaceae in the monocot tree of life.

DISCUSSION

Our study provides the first plastome-based, large-scale survey of relationships in Dioscoreales. It uses full genomes for photosynthetic taxa and reduced genomes for fully

15372197, 0, Downloaded from https://bsapubs.onlinelibrary.wiley.com/doi/10.1002/ajb2.70267, Wiley Online Library on [08/10/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

TABLE 2 Summary of phylogenetic experiments to sequentially add very long branch (VLB) mycoheterotrophs to a core 89-taxon set, which includes only moderately rapidly evolving mycoheterotrophs (i.e., *Burmannia itoana* and *Campylosiphon congestus* in Burmanniaceae; *Haplothismia exannulata* in Thismiaceae). The level of branch support for the placement of added taxa is indicated with respect to phylogenetic inferences documented in individual appendices.

VLB mycoheterotrophic taxa added	Type of analysis	Placement of taxa in analysis	Support for placement	Disruption of support in nearby branches[a]	Tree figures
Gymnosiphon + *Apteria* (Burmanniaceae)	Partitioned ML (DNA)	–Sister to *Campylosiphon* (in Burmanniaceae)	Weak	Yes	Appendix S17
Thismia tentaculata + *T. rodwayi* (Thismiaceae)	Partitioned ML (DNA)	–Sister to Taccaceae (Thismiaceae paraphyletic)	Weak	No	Appendix S18
Afrothismiaceae	Partitioned ML (DNA)	–Sister to Poaceae	Weak	Yes	Appendix S6
Afrothismia kupensis (Afrothismiaceae)	Partitioned ML (DNA)	–Nested within *Dioscorea* (Dioscoreaceae)	Weak	Yes	Appendix S11
Gymnosiphon, *Apteria*, *Thismia tentaculata*, *T. rodwayi*, and Afrothismiaceae	Partitioned ML (DNA)	–Afrothismiaceae sister to *Gymnosiphon* –(*Gymnosiphon*, *Apteria*, Afrothismiaceae) sister to *B. itoana* –*Thismia* sister to *Haplothismia*	All weak	Yes	Appendix S12
Gymnosiphon, *Apteria*, *Thismia tentaculata*, *T. rodwayi*, and Afrothismiaceae	C60 mixture model ML (amino-acid)	–Afrothismiaceae sister to *Gymnosiphon* –(*Gymnosiphon*, *Apteria*, Afrothismiaceae) sister to *B. itoana* –*Thismia* sister to Taccaceae	All weak	Yes	Appendix S13
Gymnosiphon, *Apteria*, *Thismia tentaculata*, *T. rodwayi*, and Afrothismiaceae	CAT-GTR mixture model Bayesian (amino-acid)	–Afrothismiaceae sister to *Gymnosiphon* –(*Gymnosiphon*, *Apteria*, Afrothismiaceae) sister to *Burmannia* –*Thismia* sister to *Haplothismia*	All weak	No	Appendix S14

[a]At least 20% reduction in branch support (posterior probabilities as percentages) compared to DNA-based analysis of the core 89-taxon set (Figure 3, Appendix S3).

15372197, 0, Downloaded from https://bsapubs.onlinelibrary.wiley.com/doi/10.1002/ajb2.70267, Wiley Online Library on [08/10/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

TABLE 3 Likelihood-based tests showing that an alternative arrangement of Thismiaceae with Burmanniaceae can be ruled out (hypothesis 1), while multiple alternative placements of Afrothismiaceae cannot (hypotheses 2–5); see Results section: Constraint tests of alternative placements of Thismiaceae and Afrothismiaceae (and see appendices noted below for the topological constraints used). Results (*P* values) from approximately unbiased (AU) and Shimodaira–Hasegawa (SH) tests compare constrained vs. unconstrained optimal trees for the five alternative hypotheses. Asterisks (*) indicate significant *P*-values (NS = nonsignificant).

Alternative hypothesis	Constraint used	AU *P*-value	SH *P*-value
1 (Appendix S5)	Burmanniaceae and Thismiaceae (ignoring *Afrothismia*)	0.001*	0.001*
2 (Appendix S7)	Dioscoreaceae and *Afrothismia*	0.483 (NS)	0.477 (NS)
3 (Appendix S8)	Burmanniaceae and *Afrothismia*	0.446 (NS)	0.442 (NS)
4 (Appendix S9)	Thismiaceae s.l. (including *Afrothismia*)	0.126 (NS)	0.145 (NS)
5 (Appendix S10)	Thismiaceae and Taccaceae plus *Afrothismia*	0.15 (NS)	0.155 (NS)

mycoheterotrophic taxa, an approach that has broadly worked well elsewhere for plastid-based studies including fully mycoheterotrophic lineages (e.g., Lam et al., 2018; and see Figure 3 here). Here we found that adding the most rapidly evolving lineages can lead to local depression in branch support and thus minor uncertainties in taxon placement—as shown in the taxon-addition experiments for VLB taxa in Thismiaceae and Burmanniaceae (e.g., Appendices 17, 18)—compared to the core taxon sampling that included more slowly evolving mycoheterotrophic lineages (Figure 3; relative rates of evolution summarized in Figure 2). Our overall understanding of higher-order relationships in Dioscoreales based on plastid data is therefore now highly consistent with a comparable study based on mitochondrial data by Lin et al. (2022; summarized in Figure 4). The main exception concerns Afrothismiaceae (e.g., Figure 4): this extremely fast evolving lineage is recovered in multiple weakly supported positions in monocot phylogeny across the variant analyses where we added different sets of taxa from the family (Appendices 6, 11), or using different likelihood and Bayesian approaches (e.g., Appendices S12–S14). Our results therefore demonstrate that most, but not all heterotrophic/VLB lineages can be accommodated in phylogenetic analyses of plastid genome data, with Afrothismiaceae as the main remaining problem.

Dioscoreales relationships are now largely well understood

Our taxon sampling is complete at the genus level for photosynthetic Dioscoreales, and most higher-order relationships among these taxa are now resolved here (Figure 3). A residual uncertainty concerns the earliest splits in Nartheciaceae, specifically the relative positions of *Aletris* and *Metanarthecium*—both to each other and to a clade comprising the three other genera in the family. These relationships were also uncertain in previous few-gene studies of Nartheciaceae (Merckx et al., 2008b; Fuse et al., 2012; Zhao et al., 2012) and may be associated with short deep branches that reflect rapid early diversification (e.g., Appendices S3, S4). An *Aletris-Metanarthecium* clade is strongly supported in our DNA analyses of the core taxon set (Figure 3, Appendix S3), which is completely consistent with moderately supported inferences from mitochondrial phylogenomic data (Figure 4; Lin et al., 2022). However, this arrangement conflicts with the strongly supported placement of *Metanarthecium* as sister to the remaining Nartheciaceae genera in a recent DNA-based analysis of similar plastid phylogenomic data to that employed here (Garrett et al., 2023). The reason for this clash is unclear, although the latter analysis included numerous VLB taxa in Thismiaceae that we did not include here.

A second uncertainty for the photosynthetic taxa concerns the placement of *Stenomeris*, for which plastid phylogenomic data are currently lacking. The genus is either placed in Dioscoreaceae (e.g., Angiosperm Phylogeny Group, 2016), or has been recognized as its own family, Stenomeridaceae (reviewed by Caddick et al., 2002a). Here *Stenomeris* is recovered as the sister group of *Dioscorea* (Dioscoreaceae) with poor to moderate support based on the inclusion of only three available plastid genes in our phylogenomic analyses (Appendices S12–S14, S19). The placement of *Stenomeris* was also uncertain in previous phylogenetic studies that used a few plastid markers or morphology (Caddick et al., 2002a; Wilkin et al., 2005; Hsu et al., 2013; Viruel et al., 2016). A specimen of *Stenomeris dioscoreifolia* that was nested within *Dioscorea* with strong support in analyses by Merckx et al. (2009b, 2010) likely was mislabeled or contaminated (V. Merckx, unpublished data). A recent study based on nuclear phylogenomic data recovered *Stenomeris* as the sister group to a clade comprising *Dioscorea*, *Tacca*, and *Trichopus* with strong support (Zuntini et al., 2024; see supplementary fig. 2 in their paper), an arrangement that has not been recovered in previous studies. However, 50% of the ~353 genes used in the coalescence-based analysis conflict with this arrangement, and so its position may still be considered to be ambiguous.

Using plastid data to place mycoheterotrophs with high to very high substitution rates

Analyses of a core 89-taxon set here, which includes the least rapidly evolving full mycoheterotrophs sampled in

15372197, 0, Downloaded from https://bsapubs.onlinelibrary.wiley.com/doi/10.1002/ajb2.70267, Wiley Online Library on [08/10/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Dioscoreales (Figure 2), provide well supported inferences across most branches in the order (Figure 3; Appendices S3, S4). The placements of the three FM taxa included in this set (i.e., *Burmannia itoana* and *Campylosiphon congestus* in Burmanniaceae; *Haplothismia exannulata* in Thismiaceae) and the remaining higher-order relationships in Dioscoreales are not sensitive to the use of DNA or AA data, although the latter yielded lower bootstrap support for some branches (Figure 3; Appendices S3, S4). We obtained these results despite extensive gene loss in the three FM taxa (Figure 1), suggesting that their moderately high substitution rates and patchy gene recovery do not strongly interfere with phylogenetic inference. This finding is consistent with (1) a previous study showing that phylogenetic analyses can benefit from inclusion of taxa with highly incomplete gene sets, as even small amounts of data can help to subdivide long branches (e.g., Wiens and Tiu, 2012); and (2) phylogenomic studies that successfully place other mycoheterotrophic lineages, including Corsiaceae (Bodin et al., 2016; Givnish et al., 2016; Lam et al., 2018), *Geosiris* (Iridaceae; Joyce et al., 2018; Lam et al., 2018), *Petrosavia* (Petrosaviaceae; Logacheva et al., 2014; Lam et al., 2018), Triuridaceae (Lam et al., 2015, 2018; Soto Gomez et al., 2020), multiple Orchidaceae taxa (e.g., *Aphyllorchis*, *Cephalanthera*, *Corallorhiza*, *Cymbidium*, *Cyrtosia*, *Danxiaorchis*, *Degranvillea*, *Dipodium*, *Hetaeria*, *Lecanorchis*, *Neottia*, *Risleya*; Delannoy et al., 2011; Givnish et al., 2015; Feng et al., 2016; Lam et al., 2018; Kim et al., 2020; Li et al., 2020; Barrett et al., 2024a), and the eudicot families Ericaceae, Gentianaceae, and Polygalaceae (Braukmann and Stefanović, 2012; Braukmann et al., 2017; Lam et al., 2018; Petersen et al., 2019).

The monophyly of the mycoheterotrophic family Burmanniaceae s.s. remains strongly supported in this analysis (setting aside the problematic behavior of Afrothismiaceae, e.g., Table 2, Appendices S12–S14). However, we were unable to robustly place the *Apteria-Gymnosiphon* clade precisely within Burmanniaceae because the simultaneous addition of these two genera results in substantial reductions (≥20%) in bootstrap support for most branches in the family compared to the core taxon set (Table 2, compare Figure 3 to Appendix S17). This support reduction may represent a mild long branch artifact. Analyses using mitochondrial phylogenomic data (Lin et al., 2022) also inferred a strongly supported *Apteria-Gymnosiphon* clade, in turn moderately supported as the sister group of *Burmannia* plus *Campylosiphon* (Figure 4). The latter arrangement is at odds with an older parsimony analysis using three nuclear and mitochondrial loci, which obtained a strongly supported *Apteria-Burmannia* clade, in turn sister to a strongly supported *Gymnosiphon-Hexapterella* clade (Merckx et al., 2008a). Although *Hexapterella* was not included here or by Lin et al. (2022), the contrasting results of Merckx et al. (2008a) may represent another long branch artifact in our studies. Broader samplings of Burmanniaceae are needed in future studies (including mitochondrial- and plastid-based, for the more slowly evolving taxa) to continue clarifying genus-level relationships and characterizing the exact number of origins of full mycoheterotrophy in this family, currently estimated to be around eight such transitions (Merckx et al., 2013).

We expected to observe strong long branch artifacts in the taxon-addition experiments for Thismiaceae because the two added VLB species, *Thismia rodwayi* and *T. tentaculata*, have higher substitution rates than those of VLB in Burmanniaceae (see Figure 2 for *T. tentaculata* and Appendix S18 for both species). However, our results from this experiment are broadly consistent with those from analyses of the core taxon set that only included the moderately rate-elevated *Haplothismia* as a Thismiaceae representative (compare Figure 3 to Appendix S18). Although we recovered Thismiaceae as a grade in two analyses (Appendices S13, S18), members of this FM family still form a larger clade with Taccaceae that is in turn inferred to be sister to Trichopodaceae, with only a slight reduction (<20%) in bootstrap support (Table 2, compare Figure 3 to Appendix S18). Thus, the arrangement recovered here for these three families is also largely consistent with recent plastid- and mitochondrial-based phylogenomic inferences (Lin et al., 2022; Garrett et al., 2023).

Addressing the challenge of placing VLB mycoheterotrophs using plastid data

In general, model-based methods like maximum likelihood may be expected to be less prone to long branch artifacts than nonparametric methods such as parsimony (Felsenstein, 1981; Huelsenbeck, 1995; Swofford et al., 2001; for examples, see Nickrent et al., 2004; Merckx et al., 2009b; Lam et al., 2015, 2016, 2018). However, model-based analysis is not immune to long branch problems, and so mixture models implemented in likelihood or Bayesian frameworks may help ameliorate the effects of long branches in some cases (e.g., Lartillot et al., 2007; Quang et al., 2008; Wang et al., 2018). Here, we show that these alternative model-based approaches can still be misled with respect to VLB taxa. This appears to be the case for Afrothismiaceae, a family with very severe rate elevation in its retained plastid genes (Figure 2; Appendices S6, S11). We recovered multiple highly disjunct and conflicting placements for Afrothismiaceae across the variant likelihood and Bayesian analyses performed here (Figure 4; Appendices S6, S11–S14) (see also Caddick et al., 2002a; Merckx and Bidartondo, 2008; Yokoyama et al., 2008; Merckx et al., 2006, 2009b, 2010; Shepeleva et al., 2020; Lin et al., 2022). The optimal (and absurd) placement of Afrothismiaceae here as the sister group of Poaceae (grasses) in the taxon-addition experiment that included all sampled Afrothismiaceae taxa (Appendix S6) is especially peculiar and is very likely a long branch artifact because grasses represent one of the faster evolving photosynthetic lineages of angiosperms (Bousquet et al., 1992; De La Torre et al., 2017, and see Appendices S3, S4 here). Heterotrophic plant plastomes may be characterized by an increase in AT content, a feature that could cause sequence convergence among these taxa and

15372197, 0, Downloaded from https://bsapubs.onlinelibrary.wiley.com/doi/10.1002/ajb2.70267, Wiley Online Library on [08/10/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

therefore their incorrect inference as strongly supported "clades" despite being distantly related, as documented for the parasitic species *Rhopalocnemis phalloides* (Balanophoraceae, Santanales; Schelkunov et al., 2019). Nucleotide compositional bias may be involved in the inferred placement of Afrothismiaceae nested among FM species of Burmanniaceae here because these lineages have an AT content of 69.4% and 64.6%, respectively, which is higher than in the autotrophic plants sampled here (60.5%). However, nucleotide composition is not likely involved in the inferred placement of Afrothismiaceae within Poales because the latter group has an AT-content of 59.8%, which is comparable to that of the remaining autotrophic taxa sampled here.

The different placements that we recovered for Afrothismiaceae also disagree with previous studies based on a single nuclear locus and one to two relatively slowly evolving mitochondrial genes (Merckx and Bidartondo, 2008; Merckx et al., 2010; Merckx and Smets, 2014) and with analyses based on a full complement of protein-coding mitochondrial genes (Lin et al., 2022, and see Figure 4). In these nonplastid-based studies, Afrothismiaceae is deeply nested within Dioscoreales close to Taccaceae, Trichopodaceae and the rest of Thismiaceae, but is also consistently outside these three clades, with varying levels of branch support for these arrangements, including strong support from mitochondrial genomic data (Lin et al., 2022 and see Figure 4 here). These results helped justify the recent recognition of Afrothismiaceae as a distinct family (Cheek et al., 2024).

The inclusion of Afrothismiaceae in plastid-based analyses here also has a strong depressive effect on support values for branches near its point of attachment in the monocot tree (Table 2; Appendices S6, S11–S14). This type of long branch artifact occurred whether Afrothismiaceae was added individually in lineage-specific phylogenetic experiments using DNA-based likelihood analyses (Appendices S6, S11), or in experiments that include the full complement of photosynthetic and FM taxa sampled here using likelihood analyses of DNA and AA data and using Bayesian analyses of AA data (Appendices S12–S14). Finally, none of the placements of Afrothismiaceae observed here for plastid data can be rejected by AU or SH tests (Table 3, Appendices S7–S10). The position of Afrothismiaceae in Dioscoreales can therefore be considered as unresolved by using plastid genomic data alone (but resolved by mitochondrial data; Lin et al., 2022 and see Figure 4 here).

Our results may therefore point to a limit to plastid-based phylogenetic analyses that incorporate extremely long branch taxa. In addition, methods intended to reduce composition-related effects for long branch taxa (Lartillot et al., 2007; Quang et al., 2008; Wang et al., 2018) do not appear to help here (Appendices S13, S14). Barrett et al. (2024a) showed that analyses that take account of heterotachy (lineage-specific variation in evolutionary rates) for long branch lineages can yield well-supported inferences, based on a study of a rapidly evolving FM orchid (*Degranvillea dermaptera*). In a study of *Wullschlaegelia*, another rapidly evolving FM orchid lineage (Barrett et al., 2024b), they also pointed to possible limits to this approach related to model parameterization, considering retained genes and taxon sampling. While we did not explore the issue of heterotachy here, we agree with Barrett et al. (2024a) that additional research is needed to compare the efficacy of different approaches—including comparisons with partitioned analyses and methods addressing composition-related effects—for handling exceptionally rapidly evolving lineages such as Afrothismiaceae.

Options for updating the higher-order classification of Dioscoreales

The current circumscription of photosynthetic Nartheciaceae is unproblematic, and the removal of Thismiaceae from Burmanniaceae leaves the latter as a well-defined family of autotrophic and mycoheterotrophic lineages. Separating Thismiaceae from Burmanniaceae in angiosperm classification schemes is strongly supported by our plastid-based analysis considering: (1) AU and SH tests, which reject constraining members of Burmanniaceae and Thismiaceae together (Table 3, Appendix S5), and (2) the strongly supported intervening branches between these two lineages in analyses of the core taxon set (Figure 3). Moreover, the Taccaceae-Thismiaceae clade that we inferred was also recovered in previous analyses based on plastid data sets comprising fewer Dioscoreales taxa (Givnish et al., 2018; Lam et al., 2018) or genes (Lam et al., 2016) or focusing on specific lineages in the order (i.e., Thismiaceae; Garrett et al., 2023). It is also consistent with inferences based on mitochondrial and nuclear data sets (Yokoyama et al., 2008; Merckx et al., 2009b, 2010; Merckx and Smets, 2014; Shepeleva et al., 2020; Lin et al., 2022).

There are several options for future adjustment to the family-level classification of Dioscoreales to accommodate these results. First, based on consideration of mitochondrial data and morphology, Afrothismiaceae were recently formalized as a new family in the order (Cheek et al., 2024). Alternative solutions that recognize fewer families ought to also be considered. One option would be to lump all lineages outside Nartheciaceae and Burmanniaceae (defined narrowly to exclude Thismiaceae) into a single large family—Dioscoreaceae would have priority—resulting in a reorganized three-family classification. However, this approach is undesirable in our view because it sweeps considerable morphological diversity (e.g., Caddick et al., 2002a) into a single taxon, making it morphologically highly heterogeneous and difficult to diagnose. A second solution would be to recognize multiple smaller families, including Taccaceae, Trichopodaceae, Thismiaceae, and the recently described Afrothismiaceae. This scheme is in practice what some recent workers have used (Merckx et al., 2010; Lin et al., 2022) and is the one that we employed here. A third solution would be to recognize Dioscoreaceae as including only *Dioscorea* (and possibly *Stenomeris*) and to recognize an

15372197, 0, Downloaded from https://bsapubs.onlinelibrary.wiley.com/doi/10.1002/ajb2.70267, Wiley Online Library on [08/10/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

expanded Taccaceae that incorporates additional families. For example, Afrothismiaceae and Thismiaceae could be lumped in Taccaceae, and perhaps Trichopodaceae too. This solution would yield a four- or five-family classification scheme for the order (depending on whether Trichophodaceae are recognized as a separate family).

All three of these options are consistent with a more narrowly defined Burmanniaceae (i.e., excluding members of Thismiaceae and Afrothismiaceae) and also maintain recognition of Nartheciaceae and Dioscoreaceae. However, an adequate classification of Dioscoreales depends in part on resolving the placement of photosynthetic *Stenomeris* within the order, a crucial priority for future research on the order. Deciding among these competing classification schemes also depends on additional considerations, such as the need to (1) minimize redundancy (e.g., lumping monotypic families that are sister to larger families), (2) maximize support for monophyly, and (3) facilitate straightforward identification by botanists (e.g., Backlund and Bremer, 1998). Regarding the latter criteria, we are currently engaged in a study to update our understanding of morphological evolution in Dioscoreales as a whole.

Some conclusions, future work, and advice for other studies

Our study provides some guidance for researchers attempting to place very rapidly evolving lineages as part of phylogenomic analysis. Here we used plastid genomic data to address the placement of VLB taxa within the yam order Dioscoreales, which is relevant to studies of its classification and evolution. For our workflow, we first characterized the range of nucleotide substitution rates in mycoheterotrophic lineages using a Bayesian relative rates analysis (Figure 2). We then used these results to select three of the less rapidly evolving FM lineages (i.e., *Burmannia itoana* and *Campylosiphon congestus* in Burmanniaceae; *Haplothismia exannulata* in Thismiaceae) for inclusion in initial analyses with photosynthetic Dioscoreales. Our phylogenetic inferences of this "core" taxon set, which comprises all photosynthetic taxa in Dioscoreales, including several species of *Burmannia*—along with the three moderately rapidly evolving FM lineages—are in general (1) well supported and (2) highly congruent with studies based on mitochondrial data (Figures 3, 4). The more slowly evolving mycoheterotrophic taxa may provide "anchors" that help to place the more rapidly evolving VLB taxa in our subsequent taxon-addition experiments. While the inclusion of VLB taxa causes moderate to substantial reductions in branch support, this phenomenon is only localized in most cases (e.g., Figure 5, Table 2). In addition, all added VLB taxa (setting aside Afrothismiaceae) are recovered in positions consistent with earlier studies (e.g., Lin et al., 2022; Garrett et al., 2023). Our study unambiguously confirms that Thismiaceae and Burmanniaceae evolved in distant parts of Dioscoreales and thus experienced distantly convergent losses of photosynthesis. This result is inconsistent with current classification (Angiosperm Phylogeny Group, 2016), which should be updated accordingly. We review several possible solutions for an updated classification, which we recommend should be postponed pending careful re-examination of the morphological support for different options for family reorganization. We strongly advise against recognizing a three-family order with only (1) Nartheciaceae, (2) Burmanniaceae (defined narrowly to exclude Thismiaceae) and (3) the remaining photosynthetic and non-photosynthetic taxa lumped into a single large family Dioscoreaceae, due to the general unwieldiness of that circumscription (see above) and uncertainties about the placement of some taxa, such as photosynthetic *Stenomeris*.

We also documented how inclusion of members of Afrothismiaceae, a VLB taxon, substantially disrupts phylogenetic inference using plastid data (Figure 5). Fortunately, the position of Afrothismiaceae in Dioscoreales phylogeny is now well understood, following phylogenetic analysis of the much more slowly evolving mitochondrial genome (e.g., Lin et al., 2022). Afrothismiaceae may be a "challenge too far" for plastid-based phylogenomic analyses at this point, even after using likelihood and Bayesian approaches designed to ameliorate composition-related effects for challenging long branch taxa (e.g., C60 and CAT-GTR; Lartillot et al., 2007; Quang et al., 2008; Wang et al., 2018). We recovered Afrothismiaceae, recently described as a family (Cheek et al., 2024), in highly disjunct positions across the monocot tree of life, based on a range of analyses using essentially the same plastid data set, and found no ability to distinguish among these placements, using likelihood-based tests of alternative trees. The inclusion of members of Afrothismiaceae also leads to widespread and severe reduction in branch support in plastid-based analyses. Plastid data sets that include Afrothismiaceae thus represent a potentially useful test case for the development of analytical approaches to deal with phylogenetic inference artifacts associated with very long branch taxa (see also Barrett et al., 2024a, b).

AUTHOR CONTRIBUTIONS

M.S.G., Q.L., V.K.Y.L., J.V., V.S.F.T.M. extracted and sequenced DNA. M.S.G., N.J.K., V.K.Y.L. assembled genome skim data and extracted gene sets. M.S.G. performed analyses. M.S.G. and S.W.G. conceived the study and wrote the paper, with contributions from all authors.

ACKNOWLEDGMENTS

The authors thank staff at the Royal Botanic Gardens, Kew, and the United States National Herbarium for providing material of *Stenomeris*; Sangjin Jo and Ki-Joong Kim (Korea University) for making plastome data for two sampled species of *Burmannia* available on GenBank; current reviewers and previous reviewers of an earlier version of this study; and Peter Stevens for discussion of Dioscoreales classification. This research was enabled in part by WestGrid (www.westgrid.ca), Compute Canada (www.computecanada.ca),

15372197, 0, Downloaded from https://bsapubs.onlinelibrary.wiley.com/doi/10.1002/ajb2.70267, Wiley Online Library on [08/10/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

and Research Computing at the James Hutton Institute ("UK's Crop Diversity Bioinformatics HPC"; BBSRC grants BB/S019669/1 and BB/X019683/1). This research was supported by an NSERC Postgraduate Fellowship, UBC Four-Year Fellowship, and Kew Future Leader Fellowship to M.S.G.; an NSERC Discovery Grant to S.W.G.; and RYC2023-042611-I by MCIU/AEI/10.13039/501100011033 and FSE+ grants to J.V.

CONFLICT OF INTEREST STATEMENT

Sean W. Graham is the Editor-in-Chief of the *American Journal of Botany* but took no part in the peer-review and decision-making processes for this paper; Pamela Diggle acted as Editor-in-Chief, and an Associate Editor handled the peer-review process.

DATA AVAILABILITY STATEMENT

All data matrices are available in Figshare: https://figshare.com/s/8470257673242e6bf43f.

ORCID

Marybel Soto Gomez https://orcid.org/0000-0003-1812-7416
Qianshi Lin https://orcid.org/0000-0001-5477-8000
Nathaniel J. Klimpert https://orcid.org/0000-0003-0015-4045
Vivienne K. Y. Lam https://orcid.org/0009-0005-0713-0240
Juan Viruel https://orcid.org/0000-0001-5658-8411
Vincent S. F. T. Merckx https://orcid.org/0000-0002-3959-8623
Paula J. Rudall https://orcid.org/0000-0002-4816-1212
Sean W. Graham https://orcid.org/0000-0001-8209-5231

REFERENCES

Altschul, S. F., W. Gish, W. Miller, E. W. Myers, and D. J. Lipman. 1990. Basic local alignment search tool. *Journal of Molecular Biology* 215: 403–410.

Angiosperm Phylogeny Group. 1998. An ordinal classification for the families of flowering plants. *Annals of the Missouri Botanical Garden* 85: 531–553.

Angiosperm Phylogeny Group. 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.

Angiosperm Phylogeny Group. 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.

Angiosperm Phylogeny Group. 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.

Azuma, H., and H. Tobe. 2005. Biogeography of arctic-subarctic plants during the glacial period: molecular phylogeny of Tofieldiaceae. *In* XVII International Botanical Congress: Abstracts, abstract P0777, 2005, Vienna, Austria.

Azuma, H., and H. Tobe. 2011. Molecular phylogenetic analyses of Tofieldiaceae (Alismatales): family circumscription and intergeneric relationships. *Journal of Plant Research* 124: 349–357.

Backlund, A., and K. Bremer. 1998. To be or not to be—principles of classification and monotypic plant families. *Taxon* 47: 391–400.

Barkman, T. J., J. R. McNeal, S.-H. Lim, G. Coat, H. B. Croom, N. D. Young, and C. W. DePamphilis. 2007. Mitochondrial DNA suggests at least 11 origins of parasitism in angiosperms and reveals genomic chimerism in parasitic plants. *BMC Evolutionary Biology* 7: 248.

Barrett, C. F., W. J. Baker, J. R. Comer, J. G. Conran, S. C. Lahmeyer, J. H. Leebens-Mack, J. Li, et al. 2016. Plastid genomes reveal support for deep phylogenetic relationships and extensive rate variation among palms and other commelinid monocots. *New Phytologist* 209: 855–870.

Barrett, C. F., J. I. Davis, J. Leebens-Mack, J. G. Conran, and D. W. Stevenson. 2013. Plastid genomes and deep relationships among the commelinid monocot angiosperms. *Cladistics* 29: 65–87.

Barrett, C. F., J. V. Freudenstein, J. Li, D. R. Mayfield-Jones, L. Perez, J. C. Pires, and C. Santos. 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, C. F., M. C. Pace, and C. W. Corbett. 2024a. Plastid genome evolution in leafless members of the orchid subfamily Orchidoideae, with a focus on *Degranvillea dermaptera*. *American Journal of Botany* 111: e16370.

Barrett, C. F., M. C. Pace, C. W. Corbett, A. H. Kennedy, H. L. Thixton-Nolan, and J. V. Freudenstein, 2024b. Organellar phylogenomics at the epidendroid orchid base, with a focus on the mycoheterotrophic *Wullschlaegelia*. *Annals of Botany* 2024: mcae084.

Barrett, C. F., B. T. Sinn, and A. H. Kennedy. 2019. Unprecedented parallel photosynthetic losses in a heterotrophic orchid genus. *Molecular Biology and Evolution* 36: 1884–1901.

Barrett, C. F., S. Wicke, and C. Sass. 2018. Dense infraspecific sampling reveals rapid and independent trajectories of plastome degradation in a heterotrophic orchid complex. *New Phytologist* 218: 1192–1204.

Bell, D., Q. Lin, W. K. Gerelle, S. Joya, Y. Chang, Z. N. Taylor, C. J. Rothfels, 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: 1–25.

Bellot, S., N. Cusimano, S. Luo, G. Sun, S. Zarre, A. Gröger, E. Temsch, and S. S. Renner. 2016. Assembled plastid and mitochondrial genomes, as well as nuclear genes, place the parasite family Cynomoriaceae in the Saxifragales. *Genome Biology and Evolution* 8: 2214–2230.

Bodin, S. S., J. S. Kim, and J.-H. Kim. 2016. Phylogenetic inferences and the evolution of plastid DNA in Campynemataceae and the mycoheterotrophic *Corsia dispar* D.L. Jones & B. Gray (Corsiaceae). *Plant Molecular Biology Reporter* 34:192–210.

Borowiec, M. L. 2016. AMAS: a fast tool for alignment manipulation and computing of summary statistics. *PeerJ* 4: e1660.

Bouckaert, R., T. G. Vaughan, J. Barido-Sottani, S. Duchêne, M. Fourment, A. Gavryushkina, J. Heled, et al. 2019. BEAST 2.5: An advanced software platform for Bayesian evolutionary analysis. *PLoS Computational Biology* 15: e1006650.

Bousquet, J., S. H. Strauss, A. H. Doerksen, and R. A. Price. 1992. Extensive variation in evolutionary rate of *rbcL* gene sequences among seed plants. *Proceedings of the National Academy of Sciences, USA* 89: 7844–7848.

Braukmann, T. W. A., M. B. Broe, S. Stefanović, and J. V. Freudenstein. 2017. On the brink: the highly reduced plastomes of non-photosynthetic Ericaceae. *New Phytologist* 216: 254–266.

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

Caddick, L. R., P. J. Rudall, P. Wilkin, T. A. J. Hedderson, and M. W. Chase. 2002a. Phylogenetics of Dioscoreales based on combined analyses of morphological and molecular data. *Botanical Journal of the Linnean Society* 138: 123–144.

Caddick, L. R. and P. Wilkin. 1998. A revision of the genus *Stenomeris* (Dioscoreaceae). *Kew Bulletin* 53: 703–712.

Caddick, L. R., P. Wilkin, P. J. Rudall, T. A. J. Hedderson, and M. W. Chase. 2002b. Yams reclassified: a recircumscription of Dioscoreaceae and Dioscoreales. *Taxon* 51: 103–114.

Chase, M., M. Fay, D. Devey, and O. Maurin. 2006. Multigene analyses of monocot relationships: a summary. *Aliso* 1993: 63–75.

15372197, 0, Downloaded from https://bsapubs.onlinelibrary.wiley.com/doi/10.1002/ajb2.70267, Wiley Online Library on [08/10/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Cheek, M., M. Etuge, and S. A. Williams. 2019. *Afrothismia kupensis* sp. nov. (Thismiaceae), Critically Endangered, with observations on its pollination and notes on the endemics of Mt Kupe, Cameroon. *Blumea* 64: 158–164.

Cheek, M., M. Soto Gomez, S. W. Graham, and P. J. Rudall. 2024. Afrothismiaceae (Dioscoreales), a new fully mycoheterotrophic family endemic to tropical Africa. *Kew Bulletin* 79: 55–73.

Cheek, M., H. Tsukaya, P. J. Rudall, and K. Suetsugu. 2018. Taxonomic monograph of *Oxygyne* (Thismiaceae), rare achlorophyllous mycoheterotrophs with strongly disjunct distribution. *PeerJ* 6: e4828.

Comer, J. R., W. B. Zomlefer, C. F. Barrett, J. I. Davis, D. W. Stevenson, K. Heyduk, and J. H. Leebens-Mack. 2015. Resolving relationships within the palm subfamily Arecoideae (Arecaceae) using plastid sequences derived from next-generation sequencing. *American Journal of Botany* 102: 888–899.

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

De La Torre, A. R., Z. Li, Y. Van De Peer, and P. K. Ingvarsson. 2017. Contrasting rates of molecular evolution and patterns of selection among gymnosperms and flowering plants. *Molecular Biology and Evolution* 34: 1363–1377.

DeShaw, A. E., F. Figueroa-Martinez, T. Pröschold, M. Lorenz, A. M. Nedelcu, D. R. Smith, and A. Reyes-Prieto. 2022. The plastomes of *Hyalomonas oviformis* and *Hyalogonium fusiforme* evolved dissimilar architectures after the loss of photosynthesis. *Plant Direct*, 6: e454

Dierckxsens, N., P. Mardulyn, and G. Smits. 2017. NOVOPlasty: de novo assembly of organelle genomes from whole genome data. *Nucleic Acids Research* 45: e18.

Dorrell, R. G., T. Azuma, M. Nomura, G. Audren de Kerdrel, L. Paoli, S. Yang, C. Bowler, et al. 2019. Principles of plastid reductive evolution illuminated by nonphotosynthetic chrysophytes. *Proceedings of the National Academy of Sciences, USA* 116: 6914–6923.

Drummond, A. J., and M. A. Suchard. 2010. Bayesian random local clocks, or one rate to rule them all. *BMC Biology* 8: 114

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

Felsenstein, J. 1981. Evolutionary trees from DNA sequences: a maximum likelihood approach. *Journal of Molecular Evolution* 17: 368–376.

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

Feng, Y.-L., S. Wicke, J.-W. Li, Y. Han, C.-S. Lin, D.-Z. Li, T.-T. Zhou, et al. 2016. Lineage-specific reductions of plastid genomes in an orchid tribe with partially and fully mycoheterotrophic species. *Genome Biology and Evolution* 8: 2164–2175.

Figueroa-Martinez, F., A. M. Nedelcu, D. R. Smith, and A. Reyes-Prieto. 2015. When the lights go out: the evolutionary fate of free-living colorless green algae. *New Phytologist* 206: 972–982.

Fuse, S., N. S. Lee, and M. N. Tamura. 2012. Biosystematic studies on the family Nartheciaceae (Dioscoreales) I. Phylogenetic relationships, character evolution and taxonomic re-examination. *Plant Systematics and Evolution* 298: 1575–1584.

Garrett, N., J. Viruel, N. Klimpert, M. Soto Gomez, V. K. Y. Lam, V. S. F. T. Merckx, and S. W. Graham. 2023. Plastid phylogenomics and molecular evolution of Thismiaceae (Dioscoreales). *American Journal of Botany* 110: e16141.

Givnish, T. J., M. Ames, J. R. McNeal, M. R. McKain, P. R. Steele, C. W. DePamphilis, S. W. Graham, et al. 2010. Assembling the tree of the monocotyledons: plastome sequence phylogeny and evolution of Poales. *Annals of the Missouri Botanical Garden* 97: 584–616.

Givnish, T. J., D. Spalink, M. Ames, S. P. Lyon, S. J. Hunter, A. Zuluaga, W. J. D. Iles, et al. 2015. Orchid phylogenomics and multiple drivers of their extraordinary diversification. *Proceedings of the Royal Society B: Biological Sciences* 282: 20151553.

Givnish, T. J., A. Zuluaga, I. Marques, V. K. Y. Lam, M. Soto Gomez, W. J. D. Iles, M. Ames, et al. 2016. Phylogenomics and historical biogeography of the monocot order Liliales: out of Australia and through Antarctica. *Cladistics* 32: 581–605.

Givnish, T. J., A. Zuluaga, D. Spalink, M. Soto Gomez, V. K. Y. Lam, J. M. Saarela, C. Sass, 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: 1888–1910.

Govaerts, R. 2021. World checklist of Nartheciaceae [online]. Website https://checklistbuilder.science.kew.org/reportbuilder.do [accessed 6 April 2021].

Govaerts, R., P. Wilkin, L. Raz, and O. Téllez-Valdés. 2021. World checklist of Dioscoreaceae [online]. Website https://checklistbuilder.science.kew.org/reportbuilder.do [accessed 6 April 2021].

Graham, S. W., V. K. Y. Lam, and V. S. F. T. Merckx. 2017. Plastomes on the edge: the evolutionary breakdown of mycoheterotroph plastid genomes. *New Phytologist* 214: 48–55.

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

Henriquez, C. L., T. Arias, J. C. Pires, T. B. Croat, and B. A. Schaal. 2014. Phylogenomics of the plant family Araceae. *Molecular Phylogenetics and Evolution* 75: 91–102.

Hsu, K.-M., J.-L. Tsai, M.-Y. Chen, H.-M. Ku, S.-C. Liu. 2013. Molecular phylogeny of *Dioscorea* (Dioscoreaceae) in East and Southeast Asia. *Blumea-Biodiversity, Evolution and Biogeography of Plants* 58: 21–27.

Huber, H. 1998. Dioscoreaceae. In K. Kubitzki [ed.], The families and genera of vascular plants, Vol. 3: Monocotyledons, Lilianae (except Orchidaceae), 216–235. Springer, Berlin, Germany.

Huelsenbeck, J. P. 1995. Performance of phylogenetic methods in simulation. *Systematic Biology* 44: 17–48.

Hunt, C. A., G. Steenbeeke, and V. S. F. T. Merckx. 2014. *Thismia megalongensis* (Thismiaceae), a new species of *Thismia* from New South Wales. *Telopea* 16: 165–174.

Iles, W. J. D., S. Y. Smith, M. A. Gandolfo, and S. W. Graham. 2015. Monocot fossils suitable for molecular dating analyses. *Botanical Journal of the Linnean Society* 178: 346–374.

Jonker, F. P. 1938. Monograph of the Burmanniaceae. *Mededelingen van het Botanisch Museum en Herbarium van de Rijksuniversiteit te Utrecht* 51: 1–279.

Joyce, E. M., D. M. Crayn, V. K. Y. Lam, W. K. Gerelle, S. W. Graham, and L. Nauheimer. 2018. Evolution of *Geosiris* (Iridaceae): historical biogeography and plastid-genome evolution in a genus of non-photosynthetic tropical rainforest herbs disjunct across the Indian Ocean. *Australian Systematic Botany* 31: 504–522.

Kalyaanamoorthy, S., B. Q. Minh, T. K. F. Wong, A. von Haeseler, and L. S. Jermiin. 2017. ModelFinder: fast model selection for accurate phylogenetic estimates. *Nature Methods* 14: 587–589.

Kim, Y.-K., S. Jo, S.-H. Cheon, M.-J. Joo, J.-R. Hong, M. Kwak, and K.-J. Kim. 2020. Plastome evolution and phylogeny of Orchidaceae, with 24 new sequences. *Frontiers in Plant Science* 11: 22.

Klimpert, N. J., J. L. S. Mayer, D. S. Sarzi, F. Prosdocimi, F. Pinheiro, and S. W. Graham. 2022. Phylogenomics and plastome evolution of a Brazilian mycoheterotrophic orchid, *Pogoniopsis schenckii*. *American Journal of Botany* 109: 2030–2050.

Kozlov, A. M., D. Darriba, T. Flouri, B. Morel, and A. Stamatakis. 2019. RAxML-NG: a fast, scalable and user-friendly tool for maximum likelihood phylogenetic inference. *Bioinformatics* 35: 4453–4455.

Kubitzki, K. 1998. Taccaceae. In K. Kubitzki [ed.], The families and genera of vascular plants, Vol. 3, Monocotyledons, Lilianae (except Orchidaceae), 425–428. Springer, Berlin, Germany.

Lam, V. K. Y., H. Darby, V. S. F. T. Merckx, G. Lim, T. Yukawa, K. M. Neubig, J. R. Abbott, et al. 2018. Phylogenomic inference in extremis: a case study with mycoheterotroph plastomes. *American Journal of Botany* 105: 480–494.

Lam, V. K. Y., V. S. F. T. Merckx, and S. W. Graham. 2016. A few-gene plastid phylogenetic framework for mycoheterotrophic monocots. *American Journal of Botany* 103: 692–708.

Lam, V. K. Y., M. Soto Gomez, and S. W. Graham. 2015. The highly reduced plastome of mycoheterotrophic *Sciaphila* (Triuridaceae)

15372197, 0, Downloaded from https://bsapubs.onlinelibrary.wiley.com/doi/10.1002/ajb2.70267, Wiley Online Library on [08/10/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

is colinear with its green relatives and is under strong purifying selection. *Genome Biology and Evolution* 7: 2220–2236.

Lartillot, N. 2020. PhyloBayes: Bayesian phylogenetics using site-heterogeneous models. *In* C. Scornavacca, F. Delsuc, and N. Galtier [eds.], Phylogenetics in the genomic era, 1.5:1–1.5:16. Self-published by the editors, open-access e-book. Available at https://inria.hal.science/PGE/.

Lartillot, N., H. Brinkmann, and H. Philippe. 2007. Suppression of long-branch attraction artefacts in the animal phylogeny using a site-heterogeneous model. *BMC Evolutionary Biology* 7: 1–14.

Lartillot, N., N. Rodrigue, D. Stubbs, and J. Richer. 2013. PhyloBayes MPI: phylogenetic reconstruction with infinite mixtures of profiles in a parallel environment. *Systematic Biology* 62: 611–615.

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

Lee, S. Y., K. Meng, H. Wang, R. Zhou, W. Liao, F. Chen, S. Zhang, and Q. Fan. 2020. Severe plastid genome size reduction in a mycoheterotrophic orchid, *Danxiaorchis singchiana*, reveals heavy gene loss and gene relocations. *Plants* 9: 521.

Li, H. T., Y. Luo, L. Gan, P.-F. Ma, L.-M. Gao. J.-B. Yang, J. Cai, et al. 2021. Plastid phylogenomic insights into relationships of all flowering plant families. *BMC Biology* 19: 232.

Li, H. T., T. S. Yi, L.-M. Gao, P.-F. Ma, T. Zhang, J.-B. Yang, M. A. Gitzendanner, et al. 2019. Origin of angiosperms and the puzzle of the Jurassic gap. *Nature Plants* 5: 461–470.

Li, Z.-H., Y. Jiang, X. Ma, J.-W. Li, J.-B. Yang, J.-Y. Wu, and X.-H. Jin. 2020. Plastid genome evolution in the subtribe Calypsoinae (Epidendroideae, Orchidaceae). *Genome Biology and Evolution* 12: 867–870.

Lin, Q., T. W. A. Braukmann, M. Soto Gomez, J. L. S. Mayer, F. Pinheiro, V. S. F. T. Merckx, S. Stefanović, and S. W. Graham. 2022. Mitochondrial genomic data are effective at placing mycoheterotrophic lineages in plant phylogeny. *New Phytologist* 236: 1908–1921.

Logacheva, M. D., M. I. Schelkunov, M. S. Nuraliev, T. H. Samigullin, and A. A. Penin. 2014. The plastid genome of mycoheterotrophic monocot *Petrosavia stellaris* exhibits both gene losses and multiple rearrangements. *Genome Biology and Evolution* 6: 238–246.

Maas, P. J., H. Maas-van de Kamer, J. van Benthem, H. C. M. Snelders, and T. Rübsamen. 1986. Burmanniaceae. In T. A. Zanoni [ed.], Flora neotropica, 1–189. New York Botanical Garden, Bronx, NY, USA.

Maas-van de Kamer, H. 1998. Burmanniaceae. In K. Kubitzki [ed.], The families and genera of vascular plants, Vol. III, Flowering plants. Monocotyledons: Lilianae (except Orchidaceae), 154–164. Springer-Verlag, Berlin, Germany.

Mennes, C. B., V. K. Y. Lam, P. J. Rudall, S. P. Lyon, S. W. Graham, E. F. Smets, and V. S. F. T. Merckx. 2015. Ancient Gondwana break-up explains the distribution of the mycoheterotrophic family Corsiaceae (Liliales). *Journal of Biogeography* 42: 1123–1136.

Mennes, C. B., E. F. Smets, S. N. Moses, and V. S. F. T. Merckx. 2013. New insights in the long-debated evolutionary history of Triuridaceae (Pandanales). *Molecular Phylogenetics and Evolution* 69: 994–1004.

Merckx, V., M. I. Bidartondo, and N. A. Hynson. 2009a. Myco-heterotrophy: when fungi host plants. *Annals of Botany* 104: 1255–1261.

Merckx, V., L. W. Chatrou, B. Lemaire, M. N. Sainge, S. Huysmans, and E. F. Smets. 2008a. Diversification of myco-heterotrophic angiosperms: evidence from Burmanniaceae. *BMC Evolutionary Biology* 8: 178.

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

Merckx, V., P. Schols, K. Geuten, S. Huysmans, and E. Smets. 2008b. Phylogenetic relationships in Nartheciaceae (Dioscoreales), with focus on pollen and orbicule morphology. *Belgian Journal of Botany* 141: 64–77.

Merckx, V., P. Schols, H. Maas-van de Kamer, P. Maas, S. Huysmans, and E. Smets. 2006. Phylogeny and evolution of Burmanniaceae (Dioscoreales) based on nuclear and mitochondrial data. *American Journal of Botany* 93: 1684–1698.

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

Merckx, V. S. F. T., and M. I. Bidartondo. 2008. Breakdown and delayed cospeciation in the arbuscular mycorrhizal mutualism. *Proceedings of the Royal Society, B: Biological Sciences* 275: 1029–1035.

Merckx, V. S. F. T., S. Huysmans, and E. Smets. 2010. Cretaceous origins of myco-heterotrophic lineages in Dioscoreales. In O. Seberg, G. Petersen, A. S. Barfod, and J. I. Davis [eds.], Diversity, phylogeny, and evolution in the monocotyledons, 39–53. Aarhus University Press, Aarhus, Denmark.

Merckx, V. S. F. T., C. B. Mennes, K. G. Peay, and J. Geml. 2013. Evolution and diversification. *In* V. S. F. Merckx [ed.], Mycoheterotrophy: The biology of plants living on fungi, 215–244.

Merckx, V. S. F. T., and E. F. Smets. 2014. *Thismia americana*, the 101st anniversary of a botanical mystery. *International Journal of Plant Sciences* 175: 165–175.

Naumann, J., J. P. Der, E. K. Wafula, S. S. Jones, S. T. Wagner, L. A. Honaas, P. E. Ralph, 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.

Neyland, R., and M. Hennigan. 2003. A phylogenetic analysis of large-subunit (26S) ribosome DNA sequences suggests that the Corsiaceae are polyphyletic. *New Zealand Journal of Botany* 41: 1–11.

Nguyen, L.-T., H. A. Schmidt, A. von Haeseler, and B. Q. Minh. 2015. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Molecular Biology and Evolution* 32: 268–274.

Nickrent, D. L. 2020. Parasitic angiosperms: How often and how many? *Taxon* 69: 5–27.

Nickrent, D. L., A. Blarer, Y.-L. Qiu, R. Vidal-Russell, and F. E. Anderson. 2004. Phylogenetic inference in Rafflesiales: the influence of rate heterogeneity and horizontal gene transfer. *BMC Evolutionary Biology* 4: 1–17.

Paradis, E., J. Claude, and K. Strimmer. 2004. APE: analyses of phylogenetics and evolution in R language. *Bioinformatics* 20: 289–290.

Paradis, E., and K. Schliep. 2019. ape 5.0: An environment for modern phylogenetics and evolutionary analyses in R. *Bioinformatics* 35: 526–528.

Pattengale, N. D., M. Alipour, O. R. P. Bininda-Emonds, B. M. E. Moret, and A. Stamatakis. 2010. How many bootstrap replicates are necessary? *Journal of Computational Biology* 17: 337–354.

Petersen, G., H. Darby, V. K. Y. Lam, H. A. Pedersen, V. S. F. T. Merckx, A. Zervas, O. Seberg, and S. W. Graham. 2019. Mycoheterotrophic *Epirixanthes* (Polygalaceae) has a typical angiosperm mitogenome but unorthodox plastid genomes. *Annals of Botany* 124: 791–807.

Popescu, A.-A., K. T. Huber, and E. Paradis. 2012. ape 3.0: New tools for distance based phylogenetics and evolutionary analysis in R. *Bioinformatics* 28: 1536–1537.

Quang, L. S., O. Gascuel, and N. Lartillot. 2008. Empirical profile mixture models for phylogenetic reconstruction. *Bioinformatics* 24: 2317–2323.

Rambaut, A., A. J. Drummond, D. Xie, G. Baele, and M. A. Suchard. 2018. Posterior summarization in Bayesian phylogenetics using Tracer 1.7. *Systematic Biology* 67: 901–904.

Ross, T. G., C. F. Barrett, M. Soto Gomez, V. K. Y. Lam, C. L. Henriquez, D. H. Les, J. I. Davis, et al. 2015. Plastid phylogenomics and molecular evolution of Alismatales. *Cladistics* 32: 160–178.

Sainge, M. N., G. B. Chuyong, and A. T. Peterson. 2017. Endemism and geographic distribution of African Thismiaceae. *Plant Ecology and Evolution* 150: 304–312.

Sass, C., W. J. D. Iles, C. F. Barrett, S. Y. Smith, and C. D. Specht. 2016. Revisiting the Zingiberales: using multiplexed exon capture to resolve ancient and recent phylogenetic splits in a charismatic plant lineage. *PeerJ* 4: e1584.

Schelkunov, M. I., M. S. Nuraliev, and M. D. Logacheva. 2019. *Rhopalocnemis phalloides* has one of the most reduced and mutated plastid genomes known. *PeerJ* 7: e7500.

15372197, 0, Downloaded from https://bsapubs.onlinelibrary.wiley.com/doi/10.1002/ajb2.70267, Wiley Online Library on [08/10/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Schlecter, R. 1921. Die Thismieae. *Notizblatt des Königl. botanischen Gartens und Museums zu Berlin* 71: 31–45.

Shepeleva, E. A., M. I. Schelkunov, M. Hroneš, M. Sochor, M. Dančák, V. S. F. T. Merckx, I. A. B. S. Kikuchi, 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 morphological traits. *Botanical Journal of the Linnean Society* 193: 287–315.

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

Soltis, P. S., and D. E. Soltis. 2003. Applying the bootstrap in phylogeny reconstruction. *Statistical Science* 18: 256–267.

Soto Gomez, M., Q. Lin, E. da Silva Leal, T. J. Gallaher, D. Scherberich, C. B. Mennes, S. Y. Smith, and S. W. Graham. 2020. A bi-organellar phylogenomic study of Pandanales: inference of higher-order relationships and unusual rate-variation patterns. *Cladistics* 36: 481–504.

Soto Gomez, M., L. Pokorny, M. B. Kantar, F. Forest, I. J. Leitch, B. Gravendeel, P. Wilkin, et al. 2019. A customized nuclear target enrichment approach for developing a phylogenomic baseline for *Dioscorea* yams (Dioscoreaceae). *Applications in Plant Sciences* 7: e11254.

Steele, P. R., K. L. Hertweck, D. Mayfield, M. R. McKain, J. Leebens-Mack, and J. Chris Pires. 2012. Quality and quantity of data recovered from massively parallel sequencing: examples in Asparagales and Poaceae. *American Journal of Botany* 99: 330–348.

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

Su, H.-J., J.-M. Hu, F. E. Anderson, J. P. Der, and D. L. Nickrent. 2015. Phylogenetic relationships of Santalales with insights into the origins of holoparasitic Balanophoraceae. *Taxon* 64: 491–506.

Swofford, D. L., P. J. Waddell, J. P. Huelsenbeck, P. G. Foster, P. O. Lewis, and J. S. Rogers. 2001. Bias in phylogenetic estimation and its relevance to the choice between parsimony and likelihood methods. *Systematic Biology* 50: 525–539.

Takenaka, M., A. Zehrmann, D. Verbitskiy, B. Härtel, and A. Brennicke. 2013. RNA editing in plants and its evolution. *Annual Review of Genetics* 47: 335–352.

Tamura, M. N. 1998. Nartheciaceae. *In* K. Kubitzki [ed.], The families and genera of vascular plants, Vol. III, Flowering plants. Monocotyledons: Lilianae (except Orchidaceae), 381–392. Springer-Verlag, Berlin, Germany.

Tamura, M. N., S. Fuse, H. Azuma, and M. Hasebe. 2004. Biosystematic studies on the family Tofieldiaceae I. Phylogeny and circumscription of the family inferred from DNA sequences of *mat*K and *rbc*L. *Plant Biology* 6: 562–567.

Tillich, M., P. Lehwark, T. Pellizzer, E. S. Ulbricht-Jones, A. Fischer, R. Bock, and S. Greiner. 2017. GeSeq – versatile and accurate annotation of organelle genomes. *Nucleic Acids Research* 45: W6–W11.

Trias-Blasi, A., W. J. Baker, A. L. Haigh, D. A. Simpson, O. Weber, and P. Wilkin. 2015. A genus-level phylogenetic linear sequence of monocots. *Taxon* 64: 552–581.

Viruel, J., M. Conejero, O. Hidalgo, L. Pokorny, R. F. Powell, F. Forest, M. B. Kantar, et al. 2019. A target capture-based method to estimate ploidy from herbarium specimens. *Frontiers in Plant Science* 10: 937.

Viruel, J., J. G. Segarra-Moragues, L. Raz, F. Forest, P. Wilkin, I. Sanmartín, and P. Catalán. 2016. Late Cretaceous-Early Eocene origin of yams (*Dioscorea*, Dioscoreaceae) in the Laurasian Palaearctic and their subsequent Oligocene-Miocene diversification. *Journal of Biogeography* 43: 750–762.

Wang, H.-B., B.-Q. Minh, E. Susko, and A. J. Roger. 2018. Modeling site heterogeneity with posterior mean site frequency profiles accelerates accurate phylogenomic estimation. *Systematic Biology* 67: 216–235.

Wicke, S., K. F. Müller, C. W. dePamphlis, D. Quandt, S. Bellot, and G. M. Schneeweiss. 2016. Mechanistic model of evolutionary rate variation en route to a nonphotosynthetic lifestyle in plants. *Proceedings of the National Academy of Sciences, USA* 113: 9045–9050.

Wicke, S., and J. Naumann. 2018. Molecular evolution of plastid genomes in parasitic flowering plants. *Plastid Genome Evolution* 85: 315–347.

Wicke, S., G. M. Schneeweiss, C. W. dePamphilis, K. F. Müller, and D. Quandt. 2011. The evolution of the plastid chromosome in land plants: gene content, gene order, gene function. *Plant Molecular Biology* 76: 273–297.

Wiens, J. J., and J. Tiu. 2012. Highly incomplete taxa can rescue phylogenetic analyses from the negative impacts of limited taxon sampling. *PLoS One* 7: e42925.

Wilkin, P., P. Schols, M. W. Chase, K. Chayamarit, C. A. Furness, S. Huysmans, F. Rakotonasolo, et al. 2005. A plastid gene phylogeny of the yam genus, *Dioscorea*: roots, fruits and Madagascar. *Systematic Botany* 30: 736–749.

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

Yudina, S. V., M. I. Schelkunov, L. Nauheimer, D. Crayn, S. Chantanaorrapint, M. Hroneš, M. Sochor, et al. 2021. Comparative analysis of plastid genomes in the non-photosynthetic genus *Thismia* reveals ongoing gene set reduction. *Frontiers in Plant Science* 12: 1–13.

Zhao, Y.-M., W. Wang, and S.-R. Zhang. 2012. Delimitation and phylogeny of *Aletris* (Nartheciaceae) with implications for perianth evolution. *Journal of Systematics and Evolution* 50: 135–145.

Zuntini, A. R., T. Carruthers, O. Maurin, P. C. Bailey, K. Leempoel, G. E. Brewer, N. Epitawalage, et al. 2024. Phylogenomics and the rise of the angiosperms. *Nature* 629: 843–850.

Zuntini A. R., L. P. Frankel, L. Pokorny, F. Forest, and W. J. Baker. 2021. A comprehensive phylogenomic study of the monocot order Commelinales, with a new classification of Commelinaceae. *American Journal of Botany* 108: 1066–1086.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

Appendix S1: Specimen source information.

Appendix S2: Partitioning schemes for phylogenetic analyses.

Appendix S3: Partitioned DNA analysis of the core taxon set.

Appendix S4: Partitioned amino-acid analysis of the core taxon set.

Appendix S5: Constrained monocot phylogeny to test hypothesis 1.

Appendix S6: Partitioned DNA analysis of the core taxon set plus Afrothismiaceae.

Appendix S7: Constrained monocot phylogeny to test hypothesis 2.

Appendix S8: Constrained monocot phylogeny to test hypothesis 3.

Appendix S9: Constrained monocot phylogeny to test hypothesis 4.

Appendix S10: Constrained monocot phylogeny to test hypothesis 5.

Appendix S11: Partitioned DNA analysis of the core taxon set plus *Afrothismia kupensis*.

15372197, 0, Downloaded from https://bsapubs.onlinelibrary.wiley.com/doi/10.1002/ajb2.70267, Wiley Online Library on [08/10/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Appendix S12: Partitioned DNA analysis of the full taxon set.

Appendix S13: Partitioned amino-acid analysis of the full taxon set using model C60.

Appendix S14: Partitioned amino-acid analysis of the full taxon set using model CAT-GTR.

Appendix S15: Variance plot for the Bayesian relative rate estimates.

Appendix S16: Phylogeny with branch labels corresponding to the variance plot of Bayesian relative rate estimates.

Appendix S17: Partitioned DNA analysis of the core taxon set plus *Apteria aphylla* and *Gymnosiphon longistylis*.

Appendix S18: Partitioned DNA analysis of the core taxon set plus *Thisma rodwayi* and *T. tentaculata*.

Appendix S19: Partitioned DNA analysis of the core taxon set plus *Stenomeris borneensis*.

How to cite this article: Soto Gomez, M., Q. Lin, N. J. Klimpert, V. K. Y. Lam, J. Viruel, V. S. F. T. Merckx, P. J. Rudall, and S. W. Graham. 2026. Placing very long branch taxa in the plant tree of life: A case study with Dioscoreales mycoheterotrophs. *American Journal of Botany* 113(9): e70267. https://doi.org/10.1002/ajb2.70267

15372197, 0, Downloaded from https://bsapubs.onlinelibrary.wiley.com/doi/10.1002/ajb2.70267, Wiley Online Library on [08/10/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License