Molecular Phylogenetics and Evolution 139 (2019) 106540

Contents lists available at ScienceDirect

Molecular Phylogenetics and Evolution

journal homepage: www.elsevier.com/locate/ympev

Phylogenomics of Orchidaceae based on plastid and mitochondrial genomes

Yun-Xia Li[a,g,1], Zhang-Hai Li[a,1], Andre Schuiteman[b], Mark W. Chase[b,c], Jian-Wu Li[d], Wei-Chang Huang[e], Arief Hidayat[f], Sha-Sha Wu[g], Xiao-Hua Jin[a,h,*]

[a] State Key Laboratory of Systematic and Evolutionary Botany, Institute of Botany, Chinese Academy of Sciences, Beijing, China
[b] Royal Botanic Gardens, Kew, Richmond, Surrey TW9 3AB, UK
[c] School of Plant Biology, University of Western Australia, Crawley, Perth 6009, Australia
[d] Xishuangbanna Tropical Botanical Garden, Chinese Academy of Sciences, Mengla, Yunnan 666303, China
[e] Chenshan Shanghai Botanical Garden, Shanghai, Songjiang, China
[f] Research Center for Plant Conservation and Botanic Gardens, Indonesian Institute of Sciences (LIPI), Cibinong 16911, Indonesia
[g] Fujian Agriculture and Forestry University, Shangxia Street 15, Changshan District, Fuzhou, Fujian, China
[h] Southeast Asia Biodiversity Research Institute, Chinese Academy of Science (CAS-SEABRI), China

ARTICLE INFO

Keywords:
Cypripedioideae
Mitochondrial genome
Mycoheterotrophic orchids
Orchidaceae
Plastid genome
Vanilloideae

ABSTRACT

To advance our knowledge of orchid relationships and timing of their relative divergence, we used 76 protein-coding genes from plastomes (ptCDS) and 38 protein-coding genes from mitochondrial genomes (mtCDS) of 74 orchids representing the five subfamilies and 18 tribes of Orchidaceae, to reconstruct the phylogeny and temporal evolution of the Orchidaceae. In our results, the backbone of orchid tree well supported with both datasets, but there are conflicts between these trees. The phylogenetic positions of two subfamilies (Vanilloideae and Cypripedioideae) are reversed in these two analyses. The phylogenetic positions of several tribes and subtribes, such as Epipogiinae, Gastrodieae, Nerviliinae, and Tropidieae, are well resolved in mtCDS tree. Thaieae have a different position among higher Epidendroideae, instead of sister to the higher Epidendroideae. Interrelationships of several recently radiated tribes within Epidendroideae, including Vandeae, Collabieae, Cymbidieae, Epidendreae, Podochileae, and Vandeae, have good support in the ptCDS tree, but most are not resolved in the mtCDS tree. Conflicts between the two datasets may be attributed to the different substitution rates in these two genomes and heterogeneity of substitution rate of plastome. Molecular dating indicated that the first three subfamilies, Apostasioideae, Cypripedioideae and Vanilloideae, diverged relatively quickly, and then there was a longer period before the last two subfamilies, Orchidoideae and Epidendroideae, began to radiate. Most mycoheterotrophic clades of Orchidaceae evolved in the last 30 million years with the exception of Gastrodieae.

1. Introduction

A strongly supported phylogenetic reconstruction is crucial for understanding the biogeography, phylogenetic classification, conservation and drivers of clade diversification. Advances in next-generation sequencing not only provide tremendous opportunities for inferring phylogeny using numerous plastid, mitochondrial and nuclear loci (Delsuc et al., 2005; Egan et al., 2012; Givnish et al., 2010; Givnish et al., 2018; Hosner et al., 2015; Jarvis et al., 2014; Mirarab et al., 2014; Moore et al., 2010; Prum et al., 2015), but such analyses also present computational challenges (Warnow, 2017). Increased sampling of taxa and loci can optimistically be expected to improve phylogenetic inference but often also present researchers with unexpected problems. Large datasets decrease incongruence resulting from stochastic error (Philippe et al., 2011; Philippe et al., 2005; Salamin et al., 2005), but exacerbate systematic error (Braun and Kimball, 2002; Kim, 1998; Kuhner and Felsenstein, 1994; Kumar et al., 2011; Sanderson et al., 2000) in other cases. Furthermore, phylogenetic inference based on multi-locus matrices, especially if these represent different genomes, often produce contradictory topologies, each perhaps with good support (Dunn et al., 2008; Nosenko et al., 2013; Rokas et al., 2003; Schierwater et al., 2009).

Both plastids and mitochondria are cytoplasmic organelles in angiosperms. The plastid genomes of higher plants are usually from 120 to 160 kb, including approximately 80 unique protein-coding genes, 30 tRNA genes and 4 rRNA genes (Douglas, 1998; Li et al., 2019; Wicke

* Corresponding author at: State Key Laboratory of Systematic and Evolutionary Botany, Institute of Botany, Chinese Academy of Sciences, Beijing, China.
E-mail address: xiaohuajin@ibcas.ac.cn (X.-H. Jin).
[1] These authors contributed equally to this work.

https://doi.org/10.1016/j.ympev.2019.106540
Received 18 March 2019; Received in revised form 5 June 2019; Accepted 18 June 2019
Available online 25 June 2019
1055-7903/

et al., 2011). Mitochondrial genomes of seed plants are usually in the range of 200–700 kb, including approximately 41 protein-coding genes, 15–21 tRNA genes and 3 rRNA genes (Gray et al., 1999; Gualberto and Newton, 2017; Richardson et al., 2013). Previous studies of seed plants have indicated that the synonymous substitution rate of plastid genes is intermediate between those of genes in the mitochondrial and nuclear genomes, the synonymous substitution rate being approximately 1:3:16, respectively, with few exceptions (Drouin et al., 2008; Mower et al., 2007; Sloan et al., 2012; Wolfe et al., 1987). Here we use protein-coding genes of the plastid and mitochondrial genomes to reconstruct the phylogeny of Orchidaceae.

Orchidaceae are one of the two largest families of angiosperms, consisting of approximately 750 genera and 27,000 species, of which are about 235 species in 32 genera are mycoheterotrophic (Chase et al., 2003; Chase et al., 2015; Dressler, 1981; Dressler, 1993; Rothacker, 2007). Some early ideas about orchid classification did not agree with each other (Burns-Balogh and Funk, 1986; Dressler, 1981; Dressler, 1993; Freudenstein and Rasmussen, 1999), but only the first and last of these actually used formal phylogenetic analyses of morphological data, the others were based on an intuitive interpretation of morphology; the results of these studies did not agree. Recent molecular phylogenetic analyses have generally supported recognition of five subfamilies (Cameron, 2004; Cameron et al., 1999; Chase et al., 2003; Chase et al., 2015; Gorniak et al., 2010). The same is true of two recently comprehensive analyses based on concatenated analyses of the plastome (Givnish et al., 2015) and nuclear genome (Deng et al., 2015). However, these two phylogenomic analyses were based on limited taxon sampling (39 and 10 species, respectively), although the plastome analysis was augmented by a supermatrix approach that added 144 species to extend the plastid phylogenomic backbone. In previous studies, plastome-scale results placed with strong support some mycoheterotrophic groups in positions congruent with nuclear and mitochondrial sequence results (Givnish et al., 2018; Lam et al., 2018), but phylogenetics of many mycoheterotrophic species has been problematic due to their elevated substitution rates (Barrett and Davis, 2012; Delannoy et al., 2011; Feng et al., 2016; Givnish et al., 2018; Lam et al., 2018; Logacheva et al., 2011; Wicke et al., 2016), and absence of some standard DNA markers in the highly modified plastid genomes of these taxa; we included eight mycoheterotrophic species from Vanilloideae and Epidendroideae in this study.

In the present study, we analyzed 76 plastid protein-coding genes (CDS) and 38 mitochondrial CDS from 74 orchid species representing five subfamilies and 18 tribes of Orchidaceae (following (Chase et al., 2015)), and our aims were to (1) reconstruct the phylogeny of Orchidaceae with improved sampling; (2) determine if there is conflict between the plastid and mitochondrial datasets; and (3) improve the understanding of the temporal evolution of orchids.

2. Materials and methods

2.1. Taxon sampling, DNA extraction and sequencing

In order to match ptCDS with mtCDS of each species, 74 species representing 18 tribes in all five orchid subfamilies (Table S1) were selected based on the framework of Chase et al. (2015). Four tribes, i.e. Codonorchideae, Triphoreae, Xerorchideae and Wullschlaegelieae were not included due to lack of suitable material. Total DNA was extracted from silica gel-dried leaves using a modified CTAB method (Doyle and Doyle, 1987). Total DNA was sonicated to produce 500 bp fragments with a Covaris M220 machine with appropriate parameters. Bands of interest were excised from agarose gels, and then retrieved with the E.Z.N.A. Gel Extraction Kit. Library construction was then performed following the manufacturer's protocol of NEBNext Ultra DNA Library Prep Kit. Paired-end sequencing was performed at the Institute of Botany, Chinese Academy of Sciences, on the Illumina Hiseq 2500 platform, which generated at least 2 GB of raw data for each sample.

2.2. Assembly, extraction and annotation

The raw sequence reads were first trimmed and filtered using NGSQCTOOLKIT v2.3.3 (Patel and Jain, 2012). Bases with a PHRED quality lower than 20 were trimmed. All trimmed reads shorter than 70 bp were discarded. The forward and reverse sequences were paired using the set paired reads option prior to assembly. The plastome of *Calanthe triplicata* (NC_024544.1) and mitochondrial genome of *Gastrodia elata* (MF070084-MF070102) and *Phoenix dactylifera* (NC_016740) downloaded from NCBI were employed as reference species in the subsequent steps. Paired reads of individual accessions were mapped onto the plastome of *Calanthe triplicata* and mitochondrial genome of *Gastrodia elata* to obtain the plastid (ptCDS) and mitochondrial (mtCDS) protein coding sequences, respectively.

In total, 78 ptCDS and 38 mtCDS were retrieved and annotated for per species (Table S1). In a few cases, some ptCDS could not be recovered, perhaps due to gene loss or transfer to the nuclear genome, particularly in plastomes of mycoheterotrophic species (Feng et al., 2016; Lin et al., 2017).These absent ptCDS were treated as missing data in subsequent analyses. The boundaries of CDS or exons for each species were adjusted manually to make sure that all protein-coding sequences were maintained as open reading frames. All of the above was implemented in Geneious v. 10.1.9 (http://www.geneious.com) followed by checking in Sequin and finally export as FASTA files.

2.3. Alignment, saturation test, and concatenation

All protein-coding sequences exported from Geneious were aligned with MAFFT (Katoh and Standley, 2013), and subsequently adjusted manually in Bioedit v. 5.0.9 (Hall, 1999). Saturation was examined for all CDS using DAMBE5 (Xia, 2013). Two plastid genes, *psbF* and *psbL*, were removed due to potential saturation (Tables S2, S3). Three datasets, i.e., matrix 1 (M1, all 76 species with all 76 ptCDS; Table S4), 2 (M2, all 76 species with all 38 mtCDS; Table S5) and 3 (M3, all 76 species with all 114 loci; Table S6) were used for phylogeny inference.

Aligned sequences were made into multi-gene supermatrices using Sequencematrix (Vaidya et al., 2011). Matrices in NEXUS or PHYLIP formats were exported for further analysis. These phylogenomic matrices have been deposited in the Treebase (http://www.treebase.org).

2.4. Phylogenetic analyses

Phylogenetic analyses were performed for each matrix using maximum likelihood (ML), maximum parsimony (MP) and Bayesian inference (BI). PartitionFinder2 and IQTree were used to find the most appropriate partitioning scheme for maximum likelihood (ML) and Bayesian inference (BI), respectively, for the three matrices. We also performed unpartitioned analyses for supplemental analysis to determine if there is conflict between the plastid and mitochondrial datasets (Farris et al., 1994; Hipp et al., 2004). The ILD test was implemented in PAUP v4.0b10 using 100 and 1000 replicates with random addition sequences and tree-bisection reconnection (TBR) branch swapping. The partition homogeneity test for ptCDS and mtCDS (M1, M2) shows incongruence ($P = 0.01$), and visual inspection indicated that there were some "hard" conflicts between these two results (as Mason-Gamer and Kellogg, 1996), such as relative positive of Cypripedioideae and Vanilloideae.

Maximum parsimony analyses were performed in PAUP v. 4.0b10 (Swofford, 2003) with the following heuristic search options: 10 random-stepwise-addition replicates with TBR branch swapping and 1000 bootstrap replicates. Bootstrap percentages (BP) of internal nodes were used as indicators of node reliability. Bayesian inference was conducted using MrBayes v3.2 on XSEDE in the CIPRES Science Gateway (Miller, 2010). Configuration files defined linked branchlengths, available models in MrBayes3.2, AICc model selection approach and a greedy heuristic search algorithm were prepared before

analyses (Lanfear et al., 2016). Two independent Markov-chain Monte Carlo (MCMC) analyses were performed, starting with a random trees and proceeding for 10,000,000 generations, sampling trees every 1000 generations. An average standard deviation of the split frequencies (ASDSF < 0.01) was determined to indicate a stationary distribution. The tree topology and posterior probabilities were computed after discarding the burn-in phase trees (the first 25% of sampled trees).

Maximum-likelihood (ML) phylogenetic trees were performed using IQtree (Nguyen et al., 2014). Prior to inference, all CDS were individually evaluated for the with the most appropriate model with the minimum BIC score computed by ModelFinder embeded in IQtree (Kalyaanamoorthy et al., 2017). All CharPartition commands were added to the Charset block manually. The final partitioned NEXUS file with the model for each partition specified for the individual CDSs was analyzed by ML with 2000 ultrafast bootstrap (UFBoot) (Chernomor et al., 2016; Minh et al., 2013; Nguyen et al., 2014) replicates. We consider bootstrap percentages greater than 95 as well supported.

2.5. Molecular dating

The Bayesian tree for matrix 2 (76 species, 38 mtCDS) was used as a topological constraint and calibrated using BEAST v. 2.1.3 (Bouckaert et al., 2014). Due to the limitation of computing facilities (including online CIPRES Science Gateway) and BEAST, only three commonly used mtCDS, *atp*, *matR,* and *nad5* (Cameron, 2009; Qiu et al., 2010), were used in molecular dating. Two orchid fossils were used as calibration points for crown clades: (1) *Dendrobium*, 23.2 million years ago (Ma) (mean: 1.0, sigma: 1.25 (Conran et al., 2009); (2) Goodyerinae, 15 Ma (mean: 1.0, sigma: 1.25) (Ramírez et al., 2007). Priors were also placed on the stem node of Orchidaceae and the monocots (offset: 112 Ma, sigma: 1.0) and the most recent common ancestor of all extant orchids (offset: 90 Ma, sigma: 1.0) based on previous results (Givnish et al., 2015; Givnish et al., 2016b; Xiang et al., 2016). Two runs of MCMC searches were performed for 100 million generations and sampled every 10,000 generations, and typically four non-independent chains were used for each run. A Yule process was chosen for the tree prior. Log files were monitored using Tracer v1.6 (Rambaut et al., 2014). The first 10%of trees saved from the first run and the first 8% of trees saved from the second run were discarded, and the remaining trees were combined in Logcombiner v. 2.3.0. Convergence was determined by the effective sample sizes (ESSs) of all parameters assessed as more than 100. A maximum clade credibility (MCC) chronogram was generated in TreeAnnotator v. 1.8.0 (Drummond and Rambaut, 2007; Drummond et al., 2012) with median heights for node ages.

3. Results

3.1. Data sets

For this study, ptCDS of 64 species and mtCDS of 73 species were newly obtained (Table S1). The ratio of lengths of ptCDS and mtCDS is 2:1. The combined alignment of matrix 1 of ptCDS comprised 63246 bp with 38.3%GC content, of which 12640 bp were potentially parsimony informative. The concatenated matrix 2 of mtCDS alignment comprised 33867 bp with 44.1%GC content, of which 3514 bp were potentially parsimony informative. The combined ptCDS and mtCDS matrix (matrix 3) comprised 97113 bp, of which 16,154 were potentially informative.

3.2. Phylogenetic analyses of ptCDS

Results for matrix 1, including Bayesian inference (BI), maximum likelihood (ML) and MP (maximum parsimony), are similar with a few exceptions for support of some nodes and relative positions of some tribes (Fig. 1). Results of four partitioned and the nonpartitioned ML analyses are almost identical with the exception of support for a few nodes. We use the BI results with for partitioned analyses in further discussion because BI results are the best supported.

The backbone of Orchidaceae is strongly supported (PP > 0.95, BP_{ML} > 95 or BP_{MP} > 85) with the exception of five nodes in Epidendroideae (Fig. 1). Orchidaceae are subdivided into five subfamilies with high support. Apostasioideae, Cypripedioideae and Vanilloideae are successively sister to the two largest subfamilies, Orchidoideae and Epidendroideae, which are sister to each other. Most tribes and subtribes have high support, and some tribes or subtribes, such as Epipogiinae, Gastrodieae, Pogonieae and Vanilleae, have long branches (inset in Fig. 1). Conversely, several tribes of Epidendroideae, including Collabieae, Cymbidieae, Epidendreae, Podochileae, and Vandeae, have much shorter branches along the backbone (inset in Fig. 1).

In Orchidoideae, Orchideae are sister to well supported Diurideae plus Cranichideae. In Epidendroideae, Neottieae and Sobralieae are successive sisters to the rest of Epidendroideae with high support. Thaieae are weakly supported as sister to higher Epidendroideae (Fig. 1). Mycoheterotrophic Gastrodieae are sister to Epipogiinae, and the embedding of Gastrodieae in Nervilieae is strongly supported. The position of Gastroidieae/Nervilieae is weakly supported. Two tribes, Arethuseae and Malaxideae, were resolved as successive sisters to the remaining of higher Epidendroideae (Fig. 1). In Arethuseae, Arethusinae are sister to Coelogyninae. Collabieae and Podochileae form a clade well supported as sister to Cymbidieae, Epidendreae plus Vandeae (Fig. 1). *Risleya* falls with Calypsoinae (Epidendreae) instead of in the expected position in Collabieae. Vandeae are sister to a moderate supported clade of Cymbidieae and Epidendreae. In Neottieae (including two mycoheterotrophic accessions) interrelationships were not resolved.

3.3. Phylogenetic analyses of mtCDS

Topologies from the phylogenetic analyses of 38 mtCDS genes (matrix 2) (Fig. 2) are generally congruent in terms of the clades identified with that produced from the ptCDS (matrix 1) but with many exceptions in which relative positions of subfamilies, tribes and subtribes shift (Fig. 3). Support for some nodes also differs. The backbone of Orchidaceae is strongly supported with the exception of some nodes of higher Epidendroideae (Fig. 2). Orchidaceae can be subdivided into five well-supported clades that correspond to the previously recognized subfamilies (Fig. 2). Apostasioideae, Cypripedioideae and Vanilloideae are successively sister to the rest of Orchidaceae (Fig. 2).

In Orchidoideae, Diurideae and Cranichideae are successively sister to Orchideae with strong support. In Epidendroideae, Neottieae and Sobralieae are successively sister to the rest of Epidendroideae with high support. Gastrodieae are sister to Nervilieae, and these two tribes form a highly supported clade along the backbone of Orchidaceae. Tropidieae are sister to the rest of Epidendroideae with high support (Fig. 2), after which Arethuseae, and Thaieae are successively sister to the rest with high support. *Epigeneium amplum* is distantly related to other species of Dendrobiinae as sister to the remaining of higher Epidendroideae but with no support (PP = 0.73). Cymbidieae are sister to Vandeae (minus *Polystachya*) with high support. *Polystachya* is nested within Malaxideae but with no support (PP = 0.84), far away from its usual position as sister to the rest of Vandeae. The interrelationships of six tribes, Collabieae, Podochileae, Epidendroideae, Vandeae, Malaxideae and Cymbidieae (all higher Epidendroideae) are resolved but with low to medium support (PP < 0.95, BP_{ML} < 95, BP_{MP} < 85). In Neottieae, *Neottia* is sister to remaining of tribe, including *Aphyllorchis, Cephalanthera,* and *Neottia*. In Arethuseae, *Arundina* is sister to the rest of Coelogyninae.

3.4. Phylogenetic analyses based on the combined dataset of ptCDS and mtCDS

Results of matrix 3, including BI, ML and MP, are similar to the

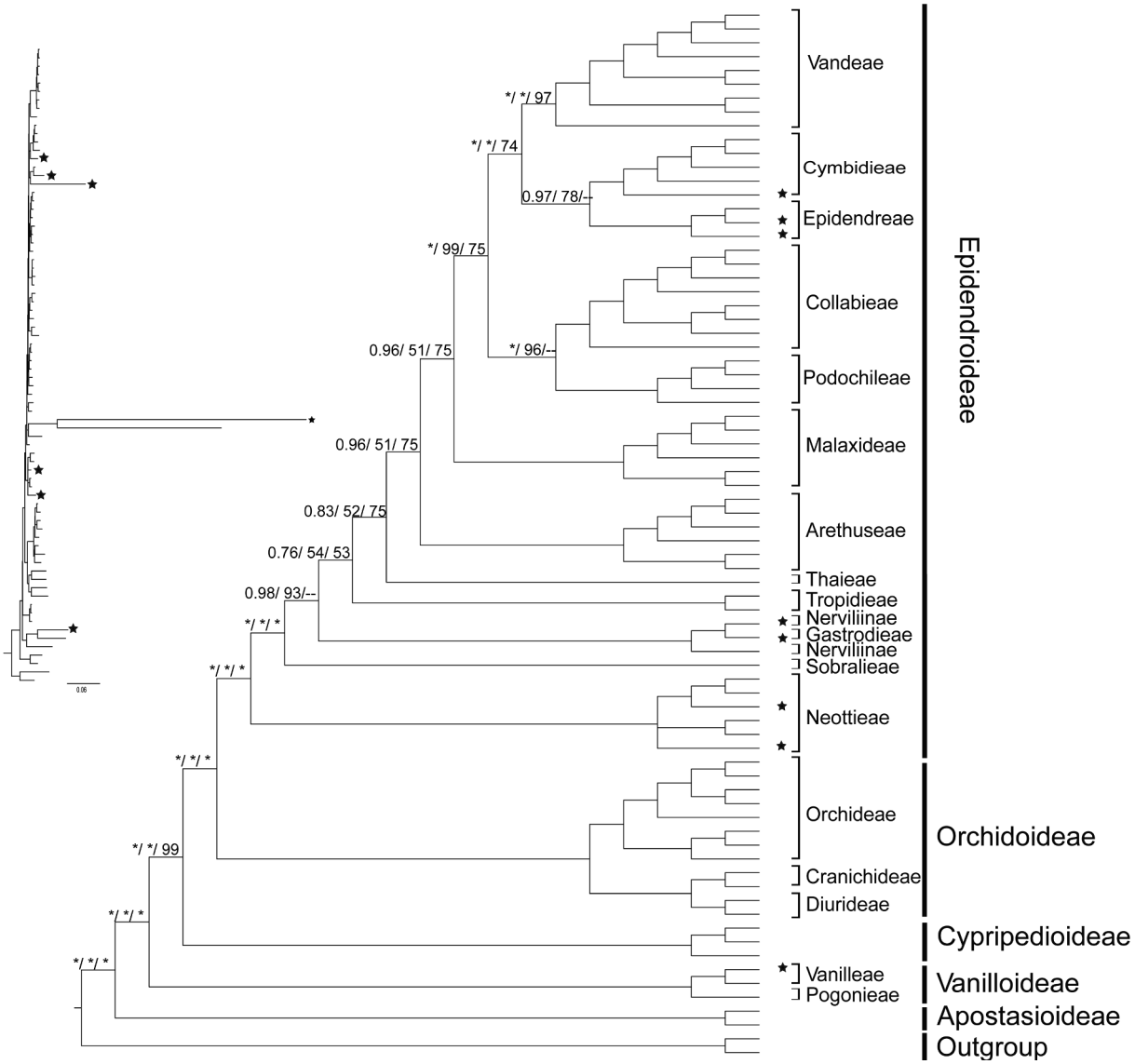

Fig. 1. Phylogram from 76 protein-coding genes of plastid genome for 76 taxa, including 74 orchids and two outgroups, inset is the outline of Bayesian inference tree showing the long branches associated with the mycoheterotrophic taxa. Asterisks indicate mycoheterotrophic taxa. Numbers above branch are Bayesian posterior probabilities/bootstrap percentages of maximum likelihood bootstrap percentages/bootstrap percentages of maximum parsimony. "*" indicates BS = 100% or PP = 1.00. "–" indicates BS < 50% or PP < 0.50. Orchid tribe and subfamily are shown for each placeholder.

matrix 1 tree (plastome) with few exceptions involving support of some nodes and relative position of few groups (Fig. 3). The backbone of Orchidaceae is fully supported (PP > 0.95, BP_{ML} > 95 or BP_{MP} > 95). Orchidaceae are subdivided into the five subfamilies with high support (Fig. 3). All sampled tribes and subtribe were recovered with high support. In Neottieae, *Neottia* is sister to remaining members of the tribe.

3.5. Molecular dating

Our results based on the mitochondrial results indicate that Cypripedioideae diverged from the remainder of Orchidaceae at 86 million years ago (Ma) and Vanilloideae at 81 Ma. The two largest subfamilies, Orchidoideae and Epidendroideae, diverged from each other 67 Ma (see Fig. 4 for ages plus 95% confidence intervals), near the K-T boundary. Within Vanilloideae, Pogonieae diverged from Vanilleae 59 Ma. The four major clades at the base of Epidendroideae diverged from the rest between 60 and 44 Ma. The seven tribes of higher Epidendroideae diverged from each rapidly between 37 and 22 Ma (Fig. 4, Table S6). Gastrodieae separated from Nervilieae about 35 Ma, whereas other six mycoheterotrophic taxa diverged from their sister groups within last 30 Ma.

4. Discussion

4.1. Conflict between ptCDS and mtCDS

There appear to be some conflicts between the mitochondrial and plastid trees. The backbone of Orchidaceae is strongly supported in both with the exception of a few nodes. The results of ptCDS agree with most previous analyses, including those based on ptCDS and nuclear genes (Deng et al., 2015; Givnish et al., 2015). The phylogenetic positions of several tribes and subtribes, such as Epipogiinae, Gastrodieae, Nervilinae, and Tropidieae, unresolved or weakly supported in previous analyses (Freudenstein and Chase, 2015; Givnish et al., 2015; Givnish et al., 2018; Lam et al., 2018), are still unresolved in our analyses of

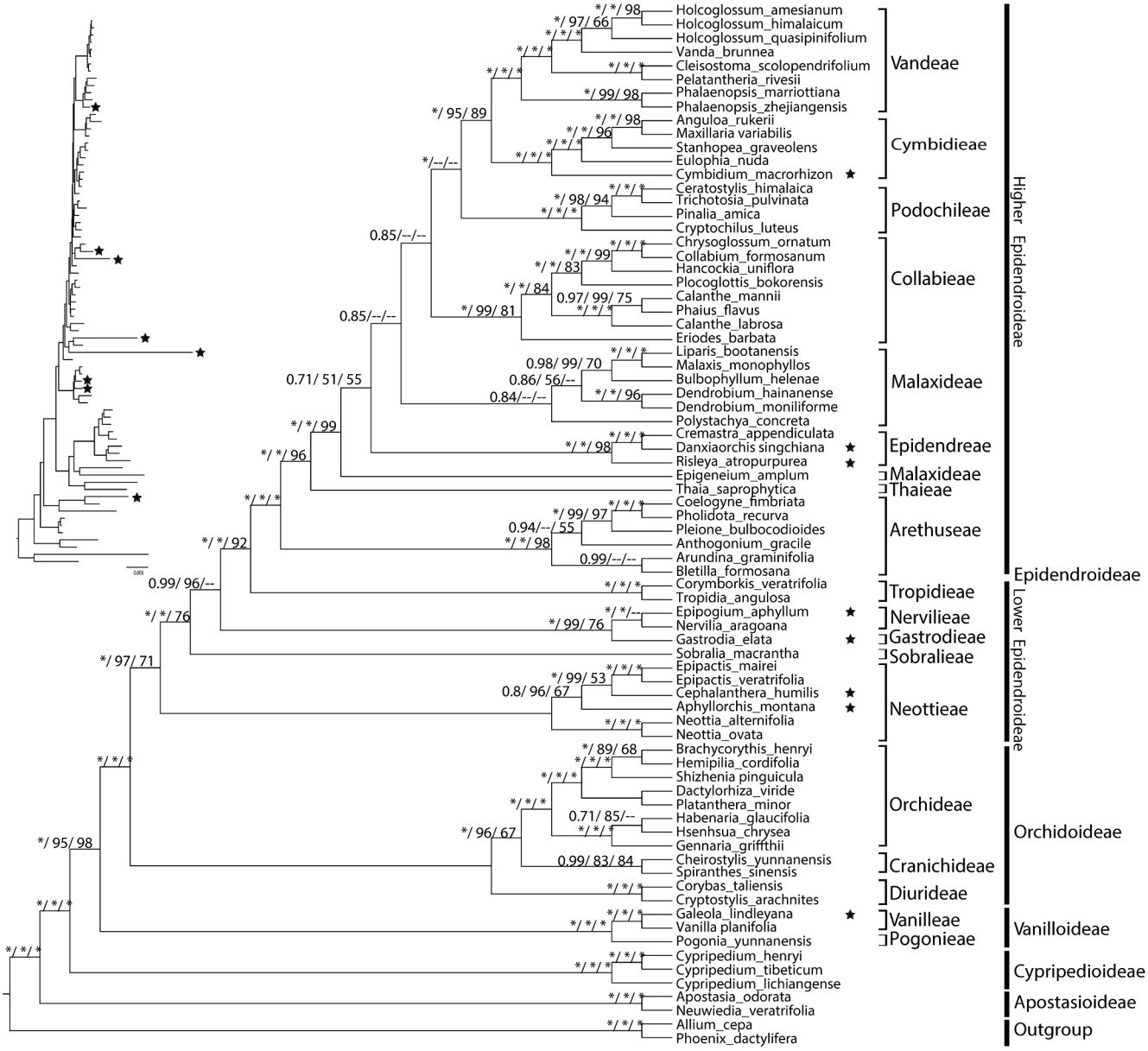

Fig. 2. Phylogram from 38 protein-coding genes of mitochondrial genome for 76 taxa, including 74 orchids and two outgroups; inset is outline of the Bayesian inference tree showing the long branches associated with the mycoheterotrophic taxa. Asterisks indicate mycoheterotrophic taxa. Numbers above branch are Bayesian posterior probabilities/maximum likelihood bootstrap percentages/maximum parsimony bootstrap percentages. "*" indicates BS = 100% or PP = 1.00. "–" indicates BS < 50% or PP < 0.50. Orchid tribe and subfamily are shown for each placeholder.

ptCDS (Fig. 1).

The mtCDS results, however, are different in some respects from that of ptCDS (Fig. 5) and all previous analyses of nuclear regions (rDNA and low-copy nuclear genes) and plastid DNA markers. The relative positions of two subfamilies, Vanilloideae and Cypripedioideae, are reversed in the trees based on these two datasets, but in the combined analysis the order found with the plastid data is recovered, which might be expected, given that the mitochondrial matrix has many fewer informative positions (Fig. 5). It is noteworthy that the relative positions of Vanilloideae and Cypripedioideae in the mtCDS tree agree with some previous results based on plastid *rbcL* (Cameron et al., 1999), although in this case the result was poorly supported. The phylogenetic positions of some tribes and subtribes, such as Epipogiinae, Gastrodieae, Tropidieae, which have always been weakly supported or not been included in previous studies, are well resolved in analyses of the mtCDS data. The position of Thaieae within the higher Epidendroideae in the mtCDS analysis is instead sister to higher Epidendroideae with the ptCDS and combined analyses. The interrelationships of several recently radiated tribes, including Vandeae, Collabieae, Cymbidieae, Epidendreae, and Podochileae, have high support in ptCDS and combined trees but most of these not in that from the mtCDS data alone.

The apparent conflicts between the plastid and mitochondrial trees may be driven by the different substitution rates of these two genomes, heterogeneity of the plastid substitution rate and RNA editing in mitochondrial genome (Bowe and DePamphilis, 1996; Liu et al., 2014; Petersen et al., 2006; Wicke et al., 2016). In comparison with plastid genes, the mitochondrial genes are characterized by lower substitution rates and correspondingly lower homoplasy, which has been argued to make them more suitable for resolving relatively deep relationships (Qiu et al., 2010; Richardson et al., 2013). In contrast, plastid genes might be more useful for reconstructing phylogenetic relationships among taxa with relatively rapid/recent speciation events, in which phylogenetic signal is often low in mtCDS. Recent results indicated that an elevated substitution rate also has an adverse impact on

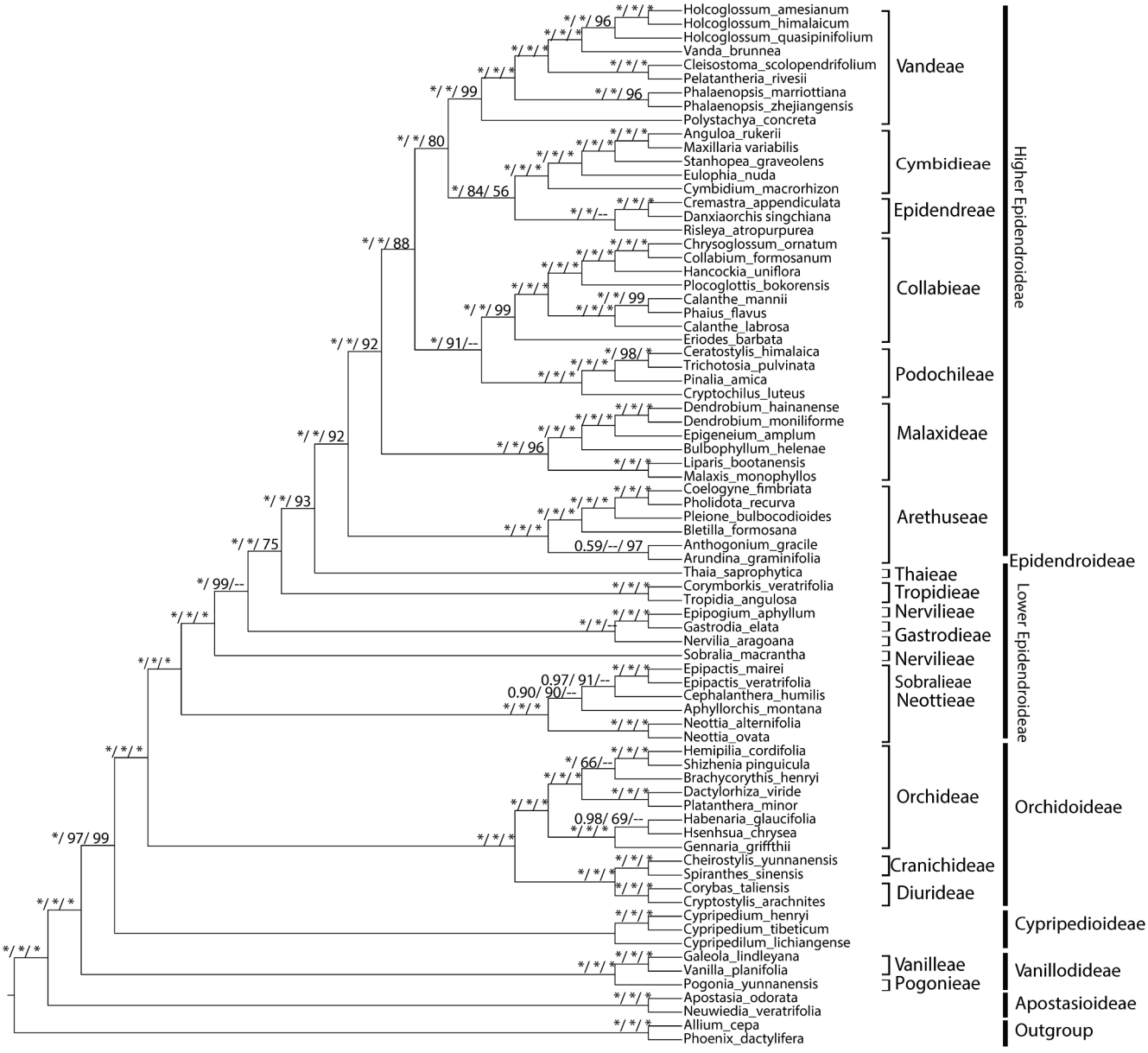

Fig. 3. Phylogram from the combined 38 mitochondrial genome and 76 plastid genome protein-coding genes for 76 taxa, including 74 orchids and two outgroup; inset is outline of Bayesian inference tree. Asterisks indicate mycoheterotrophic taxa. Numbers above branch are Bayesian posterior probabilities/ML bootstrap percentages/maximum parsimony bootstrap percentages. "*" indicates BS = 100% or PP = 1.00. "–" represents BS < 50% or PP < 0.50. Orchid tribe and subfamily are shown for each placeholder.

phylogenetic results (Xi et al., 2014). In particular, an elevated substitute rate has been identified for many plastid genes in mycoheterotrophic species (Barrett and Davis, 2012; Delannoy et al., 2011; Logacheva et al., 2011; Wicke et al., 2016). Fully mycoheterotrophic species are common in Orchidaceae (Merckx, 2013), particularly in some tribes (e.g. Gastrodieae and Wullschlaegelieae), subtribes (e.g. Epipogiinae) and genera (e.g. *Aphyllorchis* and *Lecanorchis*), all species of which are fully mycoheterotrophic (Givnish et al., 2015; Merckx, 2013). Degradation of parts of the plastid genome and an elevated substitute rate are common in these mycoheterotrophs. With few exceptions, the gene contents and substitution rates of mitochondrial genomes are more stable across the angiosperms (Drouin et al., 2008; Mower et al., 2007; Sloan et al., 2012; Wolfe et al., 1987). RNA editing sites are frequent in mitochondrial genes, but broad studies have concluded that RNA editing had no direct effect on reconstructions of phylogeny (Qiu et al., 2010). Nuclear paralogs of mitochondrial genes usually have high substitution rates, which might lead to the long branches and distorted results (Petersen et al., 2006). Our assembly of mitochondrial genes was based on reads with high coverage (coverage ranging from 50 to 100) in genome skimming data, which will reduce the likelihood of recovering nuclear paralogs.

Despite this apparent conflict, the combined analysis of these two datasets was also strongly supported (Figs. 3, 5). A similar result for the combined analysis of apparently incongruent mitochondrial and plastid genes was observed in a study of monocots (Davis et al., 2004).The results of combined matrix 3 were similar to those of matrix 1 (plastid CDS), but those of three weakly supported clades of lower Epidendroideae instead received high combined support. However, two of three nodes have slightly lower MP support along backbone of Orchidaceae (for positions of Gastrodieae + Nervilieae, $BP_{MP} < 50$; node of Tropidieae, $BP_{MP} = 75$), but these are still higher than with the ptCDS alone, and their posterior probabilities are 1.00 in all analyses. It is

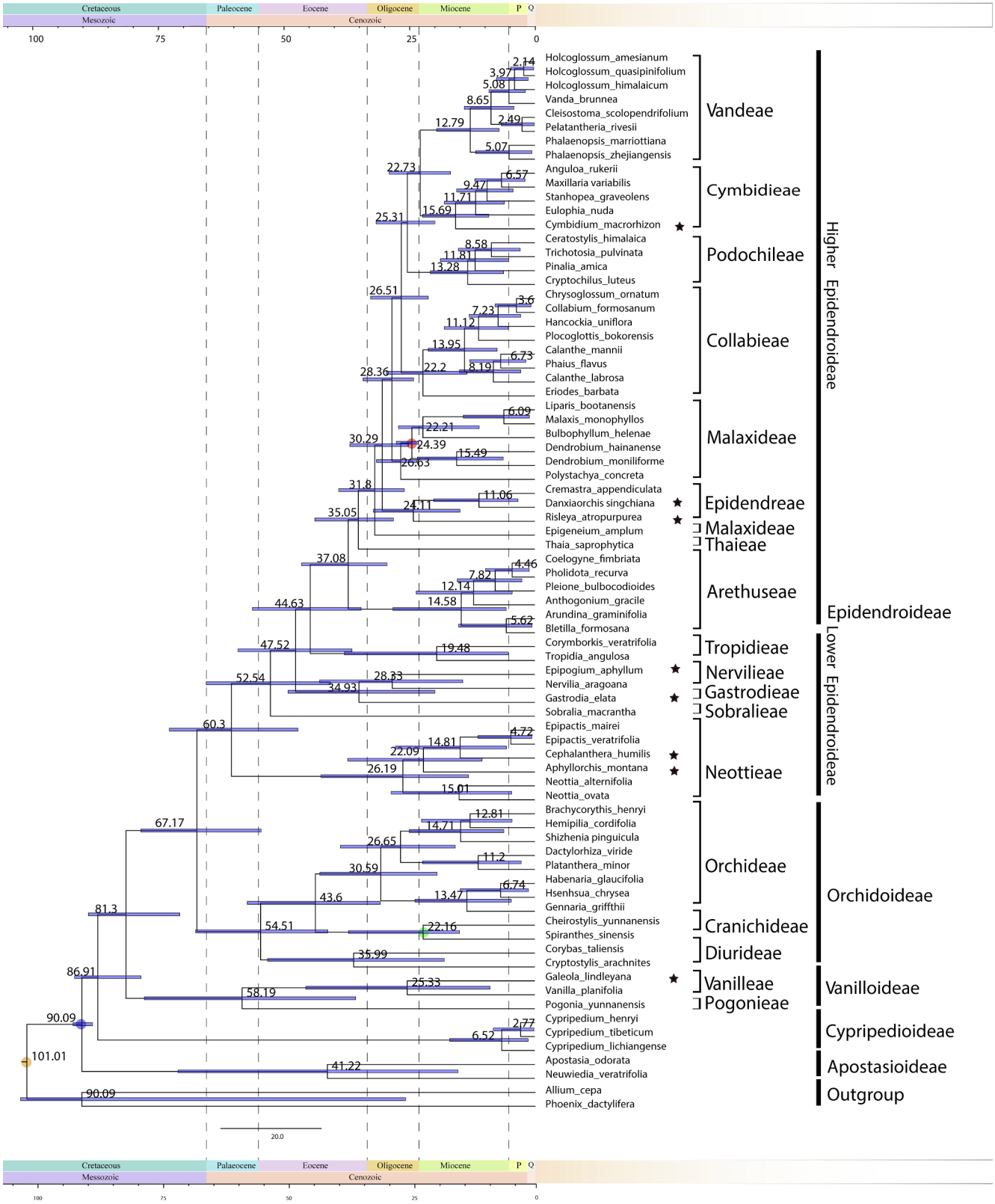

Fig. 4. Time-calibrated tree of Orchidaceae based on three protein-coding genes (*atp1*, *matR,* and *nad5*) for 74 orchid species obtained with BEAST and constrained by the topology of the mitochondrial phylogenomic concatenated tree. Numbers at nodes are median ages in million years ago (My). The topology of tree from concatenated analyses of 38 protein-coding genes of mitochondrial genome 76 taxa. Asterisks indicate mycoheterotrophic taxa. Orchid tribe and subfamily are shown for each placeholder.

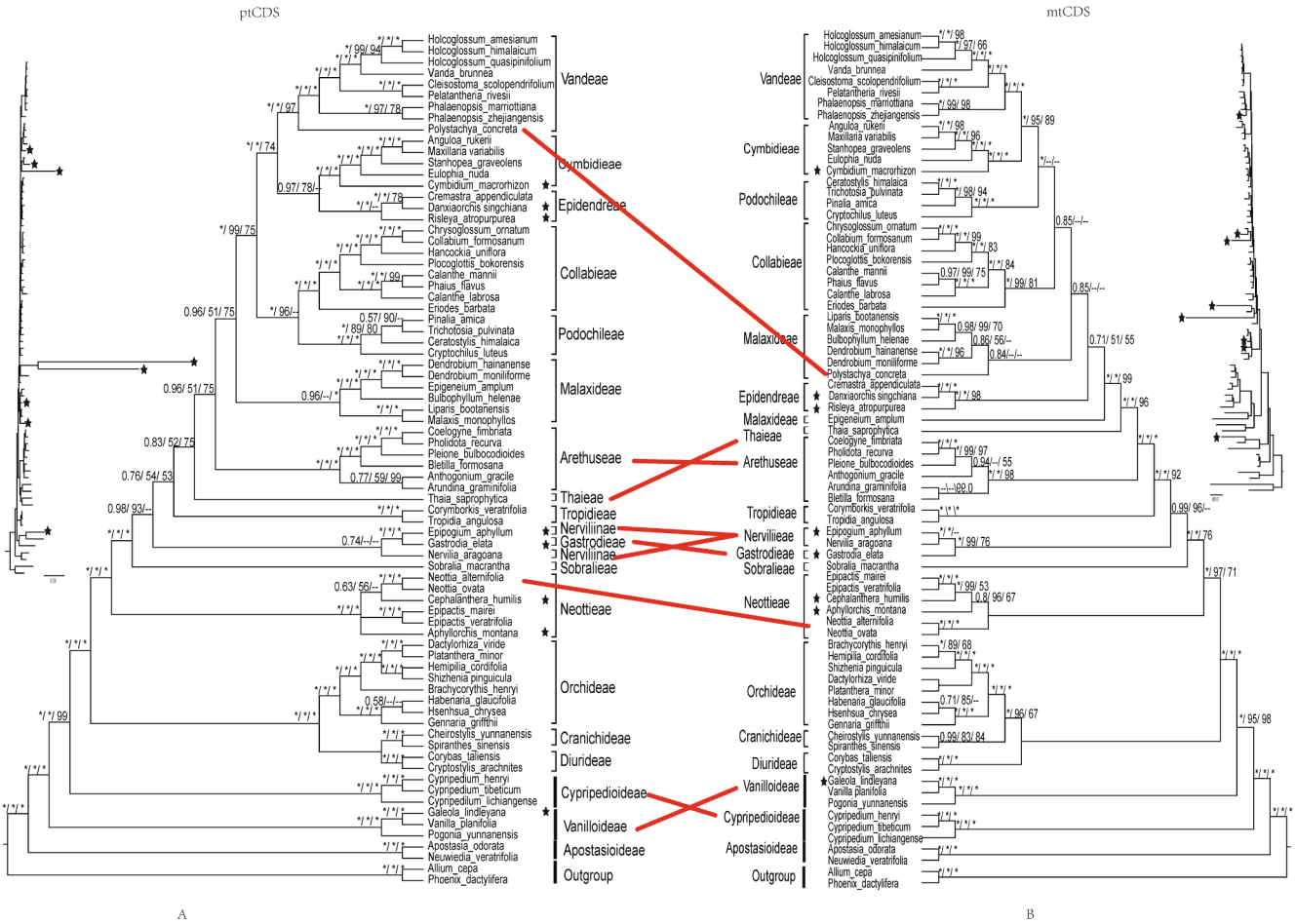

Fig. 5. Phylograms from the 76 plastid genome protein-coding genes (ptCDS) and 38 mitochondrial genome (mtCDS) for 76 taxa, respectively. A, phylogram of ptCDS; B, phylogram of mtCDS. Insets are outline of Bayesian inference tree. Asterisks indicate mycoheterotrophic taxa. Red lines indicate the relative position of some taxa in ptCDS and mtCDS phylograms. Numbers above branch are Bayesian posterior probabilities/ML bootstrap percentages/maximum parsimony bootstrap percentages. "*" indicates BS = 100% or PP = 1.00. "–" indicates BS < 50% or PP < 0.50. Orchid tribe and subfamily are shown for each placeholder.

likely that the greater number of phylogenetically informative plastid sites mask the incongruence with the mitochondrial results or that there is an obscure phylogenetic signal in the mtCDS that agrees with and enhances that in the ptCDS matrix.

4.2. Orchidaceae phylogeny based on mtCDS

Authors of recent molecular analyses of Orchidaceae have consistently recommended recognition of five subfamilies, but in the mtCDS analysis the relative placements of Cypripedioideae and Vanilloideae are reversed to that previous observed (Cameron, 2004; Cameron et al., 1999; Chase et al., 2003; Gorniak et al., 2010). This reversal of positions has important alternative implications for the evolution of some orchid characters, for example the origin of orchid velamen, and the number of fertile anther. These changed positions are more compatible with pre-molecular ideas about orchid evolution, which envisages the shift from three anthers to go to an intermediate of two before falling to the single anther present in Vanilloideae, Orchidoideae and Epidendroideae. This scenario is complicated because the anther present in orchids with a single anther is not among the two present in Cypripedioideae, so it is clear that these transitions to one or two anthers happened independently (Dressler, 1993), but the independent loss in Vanilloideae from that in Orchidoideae/Epidendroideae is not a parsimonious explanation.

Our studies have produced new insights into the orchid tree of life. Due to the loss and the elevated substitution rate of plastid genes, the phylogenetic position of several mycoheterotrophic tribes, subtribes and genera, such as Gastrodieae, Epipogiinae and *Risleya*, have not resolved or not included in previous studies (Cameron, 2004; Cameron et al., 1999; Chase et al., 2003; Gorniak et al., 2010), but the phylogenetic positions of these mycoheterotrophic groups here have high support. Gastrodieae were sister to Nervilieae with high support. The phylogenetic placement of monotypic *Risleya* was highly problematic based on morphological grounds, and it was previously considered as a member of Malaxideae or Collabieae (Chase et al., 2015; Dressler, 1993; Xiang et al., 2014); here, it falls with Calypsoinae (Epidendroideae) with high support. Monotypic Thaieae have been considered as sister to higher Epidendroideae (Xiang et al., 2012), but our results indicate that it is a member of higher Epidendroideae, which is supported by its the general morphological characters, such as corms, pseudostem formed by petioles, inflorescences arising from corms, which are similar to those of *Anthogonium* (Arethuseae) and *Cremastra* (Epidendreae). *Polystachya* has previously been included in Vandeae, but our results indicated that it may instead belong to Malaxideae, which is not in conflict with its general morphological characters, such as their shared distinctive pseudobulbs, terminal inflorescences, floral mentum and waxy pollinia. These alternative positions should be important foci in future phylogenetic and morphological studies.

Most tribes of Epidendroideae received support, and tribal interrelationships were resolved with exception of Arethuseae and

Malaxideae. The genera of Neottieae, *Aphyllorchis*, *Cephalanthera* and *Epipactis*, are highly similar in morphological terms, whereas *Neottia* is morphologically similar to *Palmorchis*, which has usually been found to be sister to the rest of Neottieae (see (Xiang et al., 2012; Zhou and Jin, 2018)). For Collabieae, relationships of each clade agree with previous results (Xiang et al., 2014).

4.3. Temporal evolution of Orchidaceae

Although one fossil of Orchidaceae (*Earina*, see (Conran et al., 2009)) was not used in our analyses, our molecular dating of Orchidaceae was based on previous results of crown age and two fossils of Orchidaceae, which seems to us to be a reliable approach. The dating indicate that there might have been a period of stasis after the early rapid divergence of the first three small subfamilies between 90 and 81 Ma. Another 14 Ma passed before divergence of two largest subfamilies (Fig. 3). The two tribes of Vanilloideae, Pogonieae and Vanilleae diverged approximate 59 Ma, pushing forward their crown ages by 18 million years (My) relative to the estimate of Givnish et al. (2015). Major clades of higher Epidendroideae, including Arethuseae, Cymbidieae, Collabieae, Epidendreae, Malaxideae, Podochileae, Vandeae and constituting about 70% of orchids species, diverged from each other between 37 and 22 My, which was younger and covered a longer period of time relative to other studies (Givnish et al., 2016a; Givnish et al., 2015; Gustafsson et al., 2010). This finding suggests that these clades, i.e., Cymbidieae and Vandeae, might have higher net species diversification than previous thought (Givnish et al., 2015).

Our results indicate that most mycoheterotrophic orchid taxa are relatively young and evolved in last 30 My, with the exception of Gastrodieae, which are slightly older. There are two possible explanations for the last result. One is that mycoheterotrophs evolved throughout the history of the orchids but have been evolutionary dead ends that did not survive for a long, resulting in no old mycoheterotrophic clades. The other is that orchids have only recently evolved the capacity to be fully mycoheterotrophic, a phenomenon that has been triggered by some unknown evolutionary innovation and/or altered environmental factors. However, the systematic position of mycoheterotrophic orchid species shows that mycoheterotrophy must have evolved several times independently, as it occurs in three subfamilies and in genera that are sometimes only distantly related within the same subfamily. Gastrodieae are the largest (approximately 100 species) and probably also the oldest mycoheterotrophic clade of extant Orchidaceae. They are widespread in Asia, Africa, Oceania, and South America (Pridgeon et al., 2005). Recently, many new species of Gastrodieae have been discovered (such as (Aung and Jin, 2018; Huang et al., 2015; Martos et al., 2015; Suetsugu, 2017)), suggesting that mycoheterotrophs could be evolutionarily active and successful, at least in the short term.

5. Conclusions

The results of our analyses of ptCDS parallel those of previous studies of plastid markers, but make important contributions in terms of the larger taxon sampling included here relative to e.g. Givnish et al. (2015, 2016a). These plastid CDS results also agree with higher-level relationships as evidenced in analyses of low-copy nuclear genes (e.g. Gorniak et al., 2010; Deng et al., 2015). To a large extent, our analyses of mtCDS agree with these previous studies, but for some taxa, it provides an alternative positions that seem generally to be better aligned with at least some morphological characters. It seems obvious to us that it is not a matter of one set of genes being “correct” and the other “wrong”. These differences suggest that the evolution of these taxa has likely been more complex than previously hypothesized. These taxa should be important foci in future orchid phylogenetic studies. Our temporal results based on the mtCDS matrix also generally support the hypotheses developed in previous such studies on orchids, but as with the phylogenetic placements they indicate slightly different scenarios that could be more accurate reflections of events. These results require further evaluation with respect to their implication for intepretation of molecular evolution. Finally, inclusion of less rate-heterogeneous mtCDS in future studies of holomycotrophic orchids holds great promise for overcoming the loss of some common phylogenetic markers in the plastid genome and widely varying substitution rates in both plastid and low-copy nuclear genes, which have made estimating phylogenetic positions of these taxa problematic in previous studies.

Authors’ contributions

X.H.J. conceived the study; Y.X.L and Z.H.L. obtained the molecular data; X.H.J., Z.H.L., and Y.X.L. participated in data analysis; X.H.J., J.W.L, W.C.H., A.H. provided plant material; X.H.J. and Y.X.L. drafted the manuscript; M.W.C, A.S. & S.S. W. revised the manuscript; all authors provided comments and final approval.

Funding

This study was supported by grants from the National Natural Science Foundation of China, China (31870195 and 31670194 to X.H.J), Southeast Asia Biodiversity Research Institute, China, Chinese Academy of Sciences, China (Y4ZK111B01 to X.H.J) and Shanghai Administration Department of Afforestation and City Appearance, China (F122432 to W.C.H.).

Data availability

These phylogenomic matrices have been deposited in the Treebase http://purl.org/phylo/treebase/phylows/study/TB2:S24472.

Appendix A. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ympev.2019.106540.

References

Aung, Y.L., Jin, X.H., 2018. *Gastrodia kachinensis* (Orchidaceae), a new species from Myanmar. PhytoKeys 94, 23–29.
Barrett, C.F., Davis, J.I., 2012. The plastid genome of the mycoheterotrophic *Corallorhiza striata* (Orchidaceae) is in the relatively early stages of degradation. Am. J. Bot. 99, 1513–1523.
Bouckaert, R., Heled, J., Kühnert, D., Vaughan, T., Wu, C.-H., Xie, D., Suchard, M.A., Rambaut, A., Drummond, A.J., 2014. BEAST 2: a software platform for Bayesian evolutionary analysis. PLoS Comput. Biol. 10, e1003537.
Bowe, L.M., DePamphilis, C.W., 1996. Effects of RNA editing and gene processing on phylogenetic reconstruction. Mol. Biol. Evol. 13, 1159–1166.
Braun, E.L., Kimball, R.T., 2002. Examining basal avian divergences with mitochondrial sequences: model complexity, taxon sampling, and sequence length. Syst. Biol. 51, 614–625.
Burns-Balogh, P., Funk, V.A., 1986. A phylogenetic analysis of the Orchidaceae. Smithsonian Contribut Botany 61, 1–79.
Cameron, K.M., 2004. Utility of plastid psaB gene sequences for investigating intrafamilial relationships within Orchidaceae. Mol. Phylogenet. Evol. 31, 1157–1180.
Cameron, K.M., 2009. On the value of nuclear and mitochondrial gene sequences for reconstructing the phylogeny of vanilloid orchids (Vanilloideae, Orchidaceae). Ann. Bot. 104, 377–385.
Cameron, K.M., Chase, M.W., Whitten, W.M., Kores, P.J., Jarrell, D.C., Albert, V.A., Yukawa, T., Hills, H.G., Goldman, D.H., 1999. A phylogenetic analysis of the Orchidaceae: evidence from rbcL nucleotide sequences. Am. J. Bot. 86, 208–224.
Chase, M.W., Cameron, K.M., Barrett, R.L., Freudenstein, J.V., 2003. DNA data and Orchidaceae systematics: a new phylogenetic classification. Orchid Conserv. 69, 89.
Chase, M.W., Cameron, K.M., Freudenstein, J.V., Pridgeon, A.M., Salazar, G., Van den Berg, C., Schuiteman, A., 2015. An updated classification of Orchidaceae. Bot. J. Linn. Soc. 177, 151–174.
Chernomor, O., von Haeseler, A., Minh, B.Q., 2016. Terrace aware data structure for phylogenomic inference from supermatrices. Syst. Biol. 65, 997–1008.
Conran, J.G., Bannister, J.M., Lee, D.E., 2009. Earliest orchid macrofossils: early Miocene *Dendrobium* and *Earina* (Orchidaceae: Epidendroideae) from New Zealand. Am. J. Bot. 96, 466–474.
Davis, J.I., Stevenson, D.W., Petersen, G., Seberg, O., Campbell, L.M., Freudenstein, J.V., Goldman, D.H., Hardy, C.R., Michelangeli, F.A., Simmons, M.P., Specht, C.D.,

Vergara-Silva, F., Gandolfo, M., 2004. A Phylogeny of the monocots, as inferred from rbcL and atpA sequence variation, and a comparison of methods for calculating jackknife and bootstrap values. Syst. Bot. 29, 467–510.
Delannoy, E., Fujii, S., Colas des Francs-Small, C., Brundrett, M., Small, I., 2011. Rampant gene loss in the underground orchid *Rhizanthella gardneri* highlights evolutionary constraints on plastid genomes. Mol. Biol. Evol. 28, 2077–2086.
Delsuc, F., Brinkmann, H., Philippe, H., 2005. Phylogenomics and the reconstruction of the tree of life. Nat. Rev. Genet. 6, 361.
Deng, H., Zhang, G.Q., Lin, M., Wang, Y., Liu, Z.J., 2015. Mining from transcriptomes: 315 single-copy orthologous genes concatenated for the phylogenetic analyses of Orchidaceae. Ecol. Evol. 5, 3800–3807.
Douglas, S.E., 1998. Plastid evolution: origins, diversity, trends. Curr. Opin. Genet. Dev. 8, 655–661.
Doyle, J., Doyle, J., 1987. CTAB DNA extraction in plants. Phytochem. Bull. 19, 11–15.
Dressler, R.L., 1981. The Orchids: Natural History and Classification. Harvard University Press, Cambridge, Massachusetts.
Dressler, R.L., 1993. Phylogeny and Classification of the Orchid Family. Cambridge University Press.
Drouin, G., Daoud, H., Xia, J., 2008. Relative rates of synonymous substitutions in the mitochondrial, chloroplast and nuclear genomes of seed plants. Mol. Phylogenet. Evol. 49, 827–831.
Drummond, A.J., Rambaut, A., 2007. BEAST: Bayesian evolutionary analysis by sampling trees. BMC Evol. Biol. 7, 214.
Drummond, A.J., Suchard, M.A., Xie, D., Rambaut, A., 2012. Bayesian phylogenetics with BEAUti and the BEAST 1.7. Mol. Biol. Evol. 29, 1969–1973.
Dunn, C.W., Hejnol, A., Matus, D.Q., Pang, K., Browne, W.E., Smith, S.A., Seaver, E., Rouse, G.W., Obst, M., Edgecombe, G.D., 2008. Broad phylogenomic sampling improves resolution of the animal tree of life. Nature 452, 745–749.
Egan, A.N., Schlueter, J., Spooner, D.M., 2012. Applications of next-generation sequencing in plant biology. Am. J. Bot. 99, 175–185.
Farris, J.S., Kallersjo, M., Kluge, A.G., Bult, C., 1994. Permuation. Cladistics- Int. J. Willi Hennig Soc. 10, 65–76.
Feng, Y.L., Wicke, S., Li, J.W., Han, Y., Lin, C.S., Li, D.Z., Zhou, T.T., Huang, W.C., Huang, L.Q., Jin, X.H., 2016. Lineage-specific reductions of plastid genomes in an orchid tribe with partially and fully Mycoheterotrophic species. Genome Biol. Evol. 8, 2164–2175.
Freudenstein, J.V., Chase, M.W., 2015. Phylogenetic relationships in Epidendroideae (Orchidaceae), one of the great flowering plant radiations: progressive specialization and diversification. Ann. Bot. 115, 665–681.
Freudenstein, J.V., Rasmussen, F.N., 1999. What does morphology tell us about orchid relationships? - A cladistic analysis. Am. J. Bot. 86, 225–248.
Givnish, T.J., Ames, M., McNeal, J.R., McKain, M.R., Steele, P.R., dePamphilis, C.W., Graham, S.W., Pires, J.C., Stevenson, D.W., Zomlefer, W.B., Briggs, B.G., Duvall, M.R., Moore, M.J., Heaney, J.M., Soltis, D.E., Soltis, P.S., Thiele, K., Leebens-Mack, J.H., 2010. Assembling the tree of the monocotyledons: plstome sequence phylogeny and evolution of Poales. Ann. Mo. Bot. Gard. 97, 584–616.
Givnish, T.J., Spalink, D., Ames, M., Lyon, S.P., Hunter, S.J., Zuluaga, A., Doucette, A., Caro, G.G., McDaniel, J., Clements, M.A., Arroyo, M.T.K., Endara, L., Kriebel, R., Williams, N.H., Cameron, K.M., 2016a. Orchid historical biogeography, diversification, Antarctica and the paradox of orchid dispersal. J. Biogeogr. 43, 1905–1916.
Givnish, T.J., Spalink, D., Ames, M., Lyon, S.P., Hunter, S.J., Zuluaga, A., Iles, W.J.D., Clements, M.A., Arroyo, M.T.K., Leebens-Mack, J., Endara, L., Kriebel, R., Neubig, K.M., Whitten, W.M., Williams, N.H., Cameron, K.M., 2015. Orchid phylogenomics and multiple drivers of their extraordinary diversification. Proc. R. Soc. B-Biol. Sci. 282, 171–180.
Givnish, T.J., Zuluaga, A., Marques, I., Lam, V.K., Gomez, M.S., Iles, W.J., Ames, M., Spalink, D., Moeller, J.R., Briggs, B.G., 2016b. Phylogenomics and historical biogeography of the monocot order Liliales: out of Australia and through Antarctica. Cladistics 32, 581–605.
Givnish, T.J., Zuluaga, A., Spalink, D., Gomez, M.S., Lam, V.K.Y., Saarela, J.M., Sass, C., Iles, W.J.D., Lima de Sousa, D.J., Leebens-Mack, J., Pires, J.C., Zomlefer, W.B., Gandolfo, M.A., Davis, J.I., Stevenson, D.W., dePamphilis, C., Specht, C.D., Graham, S.W., Barrett, C.F., Ane, C., 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. Am. J. Bot. 105, 1888–1910.
Gorniak, M., Paun, O., Chase, M.W., 2010. Phylogenetic relationships within Orchidaceae based on a low-copy nuclear coding gene, Xdh: Congruence with organellar and nuclear ribosomal DNA results. Mol. Phylogenet. Evol. 56, 784–795.
Gray, M.W., Burger, G., Lang, B.F., 1999. Mitochondrial evolution. Science 283, 1476–1481.
Gualberto, J.M., Newton, K.J., 2017. Plant mitochondrial genomes: dynamics and mechanisms of mutation. In: In: Merchant, S.S. (Ed.), Annual Review of Plant Biology, vol. 68. pp. 225–252.
Gustafsson, A.L.S., Verola, C.F., Antonelli, A., 2010. Reassessing the temporal evolution of orchids with new fossils and a Bayesian relaxed clock, with implications for the diversification of the rare South American genus *Hoffmannseggella* (Orchidaceae: Epidendroideae). BMC Evol. Biol. 10, 177.
Hall, T.A., 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. In: Nucleic acids symposium series. [London]: Information Retrieval Ltd. pp. c197995–c200098.
Hipp, A.L., Hall, J.C., Sytsma, K.J., 2004. Congruence versus phylogenetic accuracy: revisiting the incongruence length difference test. Syst. Biol. 53, 81–89.
Hosner, P.A., Faircloth, B.C., Glenn, T.C., Braun, E.L., Kimball, R.T., 2015. Avoiding missing data biases in phylogenomic inference: an empirical study in the landfowl (Aves: Galliformes). Mol. Biol. Evol. 33, 1110–1125.
Huang, X.Y., Hu, A.Q., Hsu, T.C., Liu, Y., 2015. *Gastrodia huapingensis* (Orchidaceae: Epidendroideae: Gastrodieae): a remarkable new mycoheterotrophic orchid with dimorphic columns from China. Phytotaxa 222, 290–294.
Jarvis, E.D., Mirarab, S., Aberer, A.J., Li, B., Houde, P., Li, C., Ho, S.Y., Faircloth, B.C., Nabholz, B., Howard, J.T., 2014. Whole-genome analyses resolve early branches in the tree of life of modern birds. Science 346, 1320–1331.
Kalyaanamoorthy, S., Minh, B.Q., Wong, T.K., von Haeseler, A., Jermiin, L.S., 2017. ModelFinder: fast model selection for accurate phylogenetic estimates. Nat. Methods 14, 587.
Katoh, K., Standley, D.M., 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol. Biol. Evol. 30, 772–780.
Kim, J., 1998. Large-scale phylogenies and measuring the performance of phylogenetic estimators. Syst. Biol. 47, 43–60.
Kuhner, M.K., Felsenstein, J., 1994. A simulation comparison of phylogeny algorithms under equal and unequal evolutionary rates. Mol. Biol. Evol. 11, 459–468.
Kumar, S., Filipski, A.J., Battistuzzi, F.U., Kosakovsky Pond, S.L., Tamura, K., 2011. Statistics and truth in phylogenomics. Mol. Biol. Evol. 29, 457–472.
Lam, V.K.Y., Darby, H., Merckx, V.S.F.T., Lim, G., Yukawa, T., Neubig, K.M., Abbott, J.R., Beatty, G.E., Provan, J., Gomez, M.S., Graham, S.W., 2018. Phylogenomic inference in extremis: A case study with mycoheterotroph plastomes. Am. J. Bot. 105, 480–494.
Lanfear, R., Frandsen, P.B., Wright, A.M., Senfeld, T., Calcott, B., 2016. PartitionFinder 2: new methods for selecting partitioned models of evolution for molecular and morphological phylogenetic analyses. Mol. Biol. Evol. 34, 772–773.
Li, Z.H., Ma, X., Wang, D.Y., Li, Y.X., Wang, C.W., Jin, X.H., 2019. Evolution of plastid genomes of *Holcoglossum* (Orchidaceae) with recent radiation. BMC Evol. Biol. 19, 63.
Lin, C.S., Chen, J.J., Chiu, C.C., Hsiao, H.C., Yang, C.J., Jin, X.H., Leebens-Mack, J., de Pamphilis, C.W., Huang, Y.T., Yang, L.H., 2017. Concomitant loss of NDH complex-related genes within chloroplast and nuclear genomes in some orchids. Plant J. 90, 994–1006.
Liu, Y., Cox, C.J., Wang, W., Goffinet, B., 2014. Mitochondrial phylogenomics of early land plants: mitigating the effects of saturation, compositional heterogeneity, and codon-usage bias. Syst. Biol. 63, 862–878.
Logacheva, M.D., Schelkunov, M.I., Penin, A.A., 2011. Sequencing and analysis of plastid genome in mycoheterotrophic orchid Neottia nidus-avis. Genome Biol. Evol. 3, 1296–1303.
Martos, F., Johnson, S.D., Bytebier, B., 2015. *Gastrodia madagascariensis* (Gastrodieae, Orchidaceae): from an historical designation to a description of a new species from Madagascar. Phytotaxa 221, 48–56.
Merckx, V.S., 2013. Mycoheterotrophy: an introduction. In: Mycoheterotrophy. Springer, pp. 1–17.
Miller, F.P., 2010. Molecular Genetics. Alphascript Publishing.
Minh, B.Q., Nguyen, M.A.T., von Haeseler, A., 2013. Ultrafast approximation for phylogenetic bootstrap. Mol. Biol. Evol. 30, 1188–1195.
Mirarab, S., Reaz, R., Bayzid, M.S., Zimmermann, T., Swenson, M.S., Warnow, T., 2014. ASTRAL: genome-scale coalescent-based species tree estimation. Bioinformatics 30, i541–i548.
Moore, M.J., Soltis, P.S., Bell, C.D., Burleigh, J.G., Soltis, D.E., 2010. Phylogenetic analysis of 83 plastid genes further resolves the early diversification of eudicots. PNAS 107, 4623–4628.
Mower, J.P., Touzet, P., Gummow, J.S., Delph, L.F., Palmer, J.D., 2007. Extensive variation in synonymous substitution rates in mitochondrial genes of seed plants. BMC Evol. Biol. 7, 135.
Nguyen, L.-T., Schmidt, H.A., von Haeseler, A., Minh, B.Q., 2014. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. Mol. Biol. Evol. 32, 268–274.
Nosenko, T., Schreiber, F., Adamska, M., Adamski, M., Eitel, M., Hammel, J., Maldonado, M., Müller, W.E., Nickel, M., Schierwater, B., 2013. Deep metazoan phylogeny: when different genes tell different stories. Mol. Phylogenet. Evol. 67, 223–233.
Patel, R.K., Jain, M., 2012. NGS QC Toolkit: a toolkit for quality control of next generation sequencing data. PLoS One 7, e30619.
Petersen, G., Seberg, O., Davis, J.I., Goldman, D.H., Stevenson, D.W., Campbell, L.M., Michelangeli, F.A., Specht, C.D., Chase, M.W., Fay, M.F., 2006. Mitochondrial data in monocot phylogenetics. Aliso 22, 52–62.
Philippe, H., Brinkmann, H., Lavrov, D.V., Littlewood, D.T.J., Manuel, M., Wörheide, G., Baurain, D., 2011. Resolving difficult phylogenetic questions: why more sequences are not enough. PLoS Biol. 9, e1000602.
Philippe, H., Delsuc, F., Brinkmann, H., Lartillot, N., 2005. Phylogenomics. Annu. Rev. Ecol. Evol. Syst. 36, 541–562.
Pridgeon, A.M., Cribb, P.J., Chase, M.W., Rasmussen, F.N., 2005. Genera Orchidacearum Vol. 4. Epidendroideae (Part 1). Oxford University Press, Oxford.
Prum, R.O., Berv, J.S., Dornburg, A., Field, D.J., Townsend, J.P., Lemmon, E.M., Lemmon, A.R., 2015. A comprehensive phylogeny of birds (Aves) using targeted next-generation DNA sequencing. Nature 526, 569.
Qiu, Y.L., Li, L., Wang, B., Xue, J.Y., Hendry, T.A., Li, R.Q., Brown, J.W., Liu, Y., Hudson, G.T., Chen, Z.D., 2010. Angiosperm phylogeny inferred from sequences of four mitochondrial genes. J. Systemat. Evol. 48, 391–425.
Ramírez, S.R., Gravendeel, B., Singer, R.B., Marshall, C.R., Pierce, N.E., 2007. Dating the origin of the Orchidaceae from a fossil orchid with its pollinator. Nature 448, 1042–1045.
Rambaut, A., Drummond, A., Suchard, M., 2014. Tracer v1. 6 http://beast. bio. ed. ac. uk. Tracer (visited on 2017-06-12).
Richardson, A.O., Rice, D.W., Young, G.J., Alverson, A.J., Palmer, J.D., 2013. The "fossilized" mitochondrial genome of Liriodendron tulipifera: ancestral gene content and order, ancestral editing sites, and extraordinarily low mutation rate. BMC Biol. 11, 29. https://doi.org/10.1186/1741-7007-1111-1129.
Rokas, A., Williams, B.L., King, N., Carroll, S.B., 2003. Genome-scale approaches to resolving incongruence in molecular phylogenies. Nature 425, 798–804.

Rothacker, E.P., 2007. The primitive Epidendroideae (Orchidaceae): phylogeny, character evolution and the systematics of *Psilochilus* (Triphoreae). (Doctoral dissertation). The Ohio State University.
Salamin, N., Hodkinson, T.R., Savolainen, V., 2005. Towards building the tree of life: a simulation study for all angiosperm genera. Syst. Biol. 54, 183–196.
Sanderson, M., Wojciechowski, M., Hu, J.-M., Khan, T.S., Brady, S., 2000. Error, bias, and long-branch attraction in data for two chloroplast photosystem genes in seed plants. Mol. Biol. Evol. 17, 782–797.
Schierwater, B., Eitel, M., Jakob, W., Osigus, H.-J., Hadrys, H., Dellaporta, S.L., Kolokotronis, S.-O., DeSalle, R., 2009. Concatenated analysis sheds light on early metazoan evolution and fuels a modern "urmetazoon" hypothesis. PLoS Biol. 7, e1000020.
Sloan, D.B., Alverson, A.J., Chuckalovcak, J.P., Wu, M., McCauley, D.E., Palmer, J.D., Taylor, D.R., 2012. Rapid evolution of enormous, multichromosomal genomes in flowering plant mitochondria with exceptionally high mutation rates. PLoS Biol. 10, e1001241.
Suetsugu, K., 2017. Two new species of *Gastrodia* (Gastrodieae, Epidendroideae, Orchidaceae) from Okinawa Island, Ryukyu Islands, Japan. Phytotaxa 302, 251–258.
Swofford, D.L., 2002. PAUP*: Phylogenetic Analysis Using Parsimony (and Other Methods), version 4.0b10. Sinauer, Sunderland, MA.
Vaidya, G., Lohman, D.J., Meier, R., 2011. SequenceMatrix: concatenation software for the fast assembly of multi-gene datasets with character set and codon information. Cladistics 27, 171–180.
Warnow, T., 2017. Computational challenges in constructing the tree of life. Parallel and Distributed Processing Symposium (IPDPS), 2017 IEEE International. IEEE pp. 1-1.
Wicke, S., Müller, K.F., Quandt, D., Bellot, S., Schneeweiss, G.M., 2016. Mechanistic model of evolutionary rate variation en route to a nonphotosynthetic lifestyle in plants. Proc. Natl. Acad. Sci. 113, 9045–9050.
Wicke, S., Schneeweiss, G.M., dePamphilis, C.W., Mueller, K.F., Quandt, D., 2011. The evolution of the plastid chromosome in land plants: gene content, gene order, gene function. Plant Mol. Biol. 76, 273–297.
Wolfe, K.H., Li, W.-H., Sharp, P.M., 1987. Rates of nucleotide substitution vary greatly among plant mitochondrial, chloroplast, and nuclear DNAs. Proc. Natl. Acad. Sci. 84, 9054–9058.
Xi, Z.X., Liu, L., Rest, J.S., Davis, C.C., 2014. Coalescent versus concatenation methods and the placement of *Amborella* as sister to water lilies. Syst. Biol. 63, 919–932.
Xia, X., 2013. DAMBE5: a comprehensive software package for data analysis in molecular biology and evolution. Mol. Biol. Evol. 30, 1720–1728.
Xiang, X.G., Jin, W.T., Li, D.Z., Schuiteman, A., Huang, W.C., Li, J.W., Jin, X.H., Li, Z.Y., 2014. Phylogenetics of tribe Collabieae (Orchidaceae, Epidendroideae) based on four chloroplast genes with morphological appraisal. PLoS One 9, e87625.
Xiang, X.G., Li, D.Z., Jin, W.T., Zhou, H.L., Li, J.W., Jin, X.H., 2012. Phylogenetic placement of the enigmatic orchid genera *Thaia* and *Tangtsinia*: evidence from molecular and morphological characters. Taxon 61, 45–54.
Xiang, X.G., Mi, X.C., Zhou, H.L., Li, J.W., Chung, S.W., Li, D.Z., Huang, W.C., Jin, W.T., Li, Z.Y., Huang, L.Q., Jin, X.H., 2016. Biogeographical diversification of mainland Asian *Dendrobium* (Orchidaceae) and its implications for the historical dynamics of evergreen broad-leaved forests. J. Biogeogr. 43, 1310–1323.