The evolution of mycoheterotrophy in genus *Cymbidium* (Orchidaceae)

ラン科シュンラン属における菌従属栄養の進化

2023 年 2 月

千葉大学大学院 融合理工学府

先進理化学専攻 生物学コース

茶木慧太

The evolution of mycoheterotrophy in genus *Cymbidium* (Orchidaceae)

February 2023

Graduate School of Science and Engineering

Keita Chagi

The evolution of mycoheterotrophy in Cymbidium (Orchidacdeae)

General introduction

Most land plants are autotrophic (AT), performing photosynthesis to obtain carbon. However, for the other nutrients especially for nitrogen, they receive a considerable amount of them through symbiotic fungi by forming mycorrhizas in the underground tissues (Wang & Qiu, 2006). In these mycorrhizal symbiotic systems, plants provide carbon produced by their photosynthesis to mycorrhizal fungi, and vice versa, the fungi provide the plants with nutrients such as phosphorus and nitrogen acquired from the soil (Smith & Gianinazzi-Pearson, 1988). This allows plants to obtain nutrients which are difficult to obtain from the soil by their own, and fungi to obtain the carbohydrates (Smith & Gianinazzi-Pearson, 1988). This symbiosis system is considered as the key event for plants to be able to expand into various terrestrial environments (Heckman et al., 2001).

Despite the importance of photosynthesis for the plants, some plants lack this ability and obtain not only phosphorus and nitrogen but also carbon from mycorrhizal fungi (Leake, 1994). These plants are called mycorrhizal heterotrophic (MH) plants. Among these MH plants, some are completely parasitic on mycorrhizal fungi and have evolved to obtain all the carbon they need from the fungi (Leake, 1994). These plants are called fully mycoheterotrophic (FM) and do not perform photosynthesis at all. The FM species have evolved independently in many taxonomic groups of plants, including liverworts, ferns, and vascular plants, with about 880 species (Merckx et al., 2013). Since FM plants have no need to perform photosynthesis, they lack photosynthetic organs represented by leaves (Leake, 1994). This feature allows them to inhabit the darker places such as forest floor, and it has been suggested that the FM habit may be an adaptation to such conditions (Bidartondo et al., 2004). Their mycorrhizal components also show higher specialization than for the closely related non-FM species. Previous study on Ericaceae and Orchidaceae showed that most FM species have different components of mycorrhizal fungi (Jacquemyn & Merckx, 2019;

Wang et al., 2021). In addition, the pollination ecology and floral morphology of several FM species are differing from their closely related non-FM species suggesting the changes in the pollination systems linked to the shift in trophic level (Suetsugu, 2015). All these phenomena suggest that the lack of a photosynthetic system has an impact not only on physiology, but also on the ecology of the plants.

The degree of dependency against the mycorrhizal fungi in partial mycoheterotrophic (mixotrophic; MX) plants, which utilize both carbon from photosynthesis and from mycorrhizal fungi (Selosse & Roy, 2009), could vary among species. The leaves of these species varied ranging from those with normal-sized leaves similar to AT plants to species with tiny leaves (Sakamoto et al., 2016), which might reflect the variation in dependency against their mycorrhizal fungi. Phylogenetic analysis on Aneuraceae, Ericaceae and Orchidaceae suggested that many FM plants did not evolve directly from AT plants, but through intermediate MX species (Duckett & Ligrone, 2008; Jacquemyn & Merckx, 2019; Wang et al., 2021). However, debate is ongoing about the evolution of FM species from their MX ancestral species, with two conflicting theories (Tsukaya, 2018). One hypothesis claims that the leafless individuals originated from albinistic variants of MX species (Julou et al., 2005), while the other postulates that these species lost the ability to photosynthesize after losing their leaves (Sakamoto et al., 2016). In either case, MX species might demonstrate the midpoint of evolution of FM plants, thus studying MX plants will provide insights on the physiological, ecological, and evolutional backgrounds of FM plants.

FM plants have been studied for a long time because of their strange morphology and ecology (Bidartondo, 2005). However, the study of FM species is challenging, because they are difficult to cultivate, their above-ground parts appear only during the flowering and fruiting seasons, and their mycorrhizal fungi are difficult to observe because of underground lifestyle. Nevertheless, recent advancements in molecular biology and genomics are overcoming these

problems. The next-generation sequencing and the third-generation sequencing technologies enable the generation of vast amounts of data from genomes and transcriptomes from both plants and fungi, allowing for a deeper understanding of genetic backgrounds of the FM plants and their mycorrhizas. The evolution of plants to abandon photosynthesis and lose their leaves has taken place several times independently across different plant groups, despite being a rare event in the majority of plant species. Examining these rare events might provide us with novel insights into the evolution of morphology, physiology, and symbiosis systems in plants.

Orchidaceae is one of the largest plant families and comprises more than 25,000 species (Dressler, 2005). All species are known to be MH, at least when juvenile, and about 300 species are FM (Jacquemyn & Merckx, 2019). This is the largest number of FM species in a single plant family, thus Orchidaceae is relevant to investigate the evolution of MH plants (Wang et al, 2021). Among them, I focused on genus *Cymbidium* and investigated what happened during the evolution of FM plants from MX plants. The genus *Cymbidium* mainly distributes in Southeast Asia, consisting of about 80 MH plants. Among them, a few species are almost FM, about 40 are MX, and the others are assumed to be AT (MH when juvenile). By focusing on this genus which consisted of a mosaic of these different nutrient modes, we could study the genetic and physiologic backgrounds of MH plants, including FM plants. Furthermore, this genus exhibits notable diversification not only in its nutrient modes but also in its habitats and photosynthetic pathways, making it an optimal model for exploring the evolution of the traits related to the changes in trophic level from AT to FM.

In chapter 1, I reconstruct the phylogenetic relationships in genus *Cymbidium* to elucidate the evolutionary dynamics of MH plants in this genus. Although the basic infrageneric relationships were provided by several authors (Yukawa et al., 2002; van den Berg et al., 2002), the evolution processes within the genus were little known due to the incomplete taxon sampling in

the previous studies. To solve this problem, I sampled a sufficient number of taxa from *Cymbidium* and reconstructed the phylogenetic trees and ancestral states of habitats and trophic levels for this genus. This would enable the proper interpretations of the evolution of trophic modes and related features in this genus.

In chapter 2, I focused on MX species, *C. goeringii*, to find out how the MH plants control the relationship with their mycorrhizal fungi. Some MH plants can adjust their dependency levels against their mycorrhizal fungi in response to the surrounding environmental conditions (Preiss et al., 2010; Motomura et al., 2010; Matsuda et al., 2012). However, the genetic mechanisms of this adjustment are still unclear. I inhibited their photosynthesis by shading them to increase the dependency against the mycorrhizal fungi then compared the gene expressions between shaded and unshaded individuals. This approach would give us novel clues to illuminate the systems underlying the control of fungal dependences of MH plants.

In the last chapter, chapter 3, I reconstructed the genome of *C. goeringii,* the MX species, and compared with the genome of the other MX species, *C. nagifolium*, and almost FM species, *C. macrorhizon*, to disclose the genomic backgrounds of the FM plants and its related features. A recent study conducted a comparative analysis of the genomes of closely related MX and FM orchids revealed the potential molecular basis underlying mycoheterotrophy (Li et al., 2022). They revealed the FM species showed increased substitution rates and gene loss, especially in photoreceptor and auxin transporter genes. However, it remains uncertain whether there are shared mechanisms underlying parallel evolutions of FM within the Orchidaceae. By comparing the genome of closely related MX and FM species, we could discover the further genetic backgrounds in mycoheterotrophy in plant evolution.

References

Bidartondo, M. I., Burghardt, B., Gebauer, G., Bruns, T. D., & Read, D. J. (2004). Changing partners in the dark: isotopic and molecular evidence of ectomycorrhizal liaisons between forest orchids and trees. Proceedings of the Royal Society of London. Series B: Biological Sciences, 271(1550), 1799-1806.

Bidartondo, M. I. (2005). The evolutionary ecology of myco - heterotrophy. New Phytologist, 167(2), 335-352.

Dressler, R. L. (2005). How many orchid species?. Selbyana, 155-158.

Duckett, J. G., & Ligrone, R. (2008). Basidiomycetous endophytes in New Zealand Aneuraceae (simple thalloid liverworts, Metzgeriidae) and the derived status of the monotypic genus Verdoornia. Botany, 86(4), 346-358.

Heckman, D. S., Geiser, D. M., Eidell, B. R., Stauffer, R. L., Kardos, N. L., & Hedges, S. B. (2001). Molecular evidence for the early colonization of land by fungi and plants. science, 293(5532), 1129-1133.

Jacquemyn, H., & Merckx, V. S. (2019). Mycorrhizal symbioses and the evolution of trophic modes in plants. Journal of Ecology, 107(4), 1567-1581.

Julou, T., Burghardt, B., Gebauer, G., Berveiller, D., Damesin, C., & Selosse, M. A. (2005). Mixotrophy in orchids: insights from a comparative study of green individuals and nonphotosynthetic individuals of Cephalanthera damasonium. New Phytologist, 166(2), 639-653.

Leake, J. R. (1994). The biology of myco - heterotrophic ('saprophytic') plants. New Phytologist, 127(2), 171-216.

Matsuda, Y., Shimizu, S., Mori, M., Ito, S. I., & Selosse, M. A. (2012). Seasonal and environmental changes of mycorrhizal associations and heterotrophy levels in mixotrophic Pyrola

japonica (Ericaceae) growing under different light environments. American Journal of Botany, 99(7), 1177-1188.

Merckx, V. S., Freudenstein, J. V., Kissling, J., Christenhusz, M. J., Stotler, R. E., Crandall-Stotler, B., ... & Maas, P. J. (2013). Taxonomy and classification. Mycoheterotrophy: the biology of plants living on fungi, 19-101.

Motomura, H., Selosse, M.A., Martos, F., Kagawa, A., & Yukawa, T. (2010). Mycoheterotrophy evolved from mixotrophic ancestors: Evidence in Cymbidium (Orchidaceae). Annals of Botany, 106(4), 573–581.

Preiss, K., Adam, I. K., & Gebauer, G. (2010). Irradiance governs exploitation of fungi: fine-tuning of carbon gain by two partially myco-heterotrophic orchids. Proceedings of the Royal Society B: Biological Sciences, 277(1686), 1333-1336.

Sakamoto, Y., Ogura-Tsujita, Y., Ito, K., Suetsugu, K., Yokoyama, J., Yamazaki, J., ... & Maki, M. (2016). The tiny-leaved orchid Cephalanthera subaphylla obtains most of its carbon via mycoheterotrophy. Journal of plant research, 129, 1013-1020.

Selosse MA, Roy M. Green plants that feed on fungi: facts and questions about mixotrophy. Trends in plant science, 2009;14(2):64-70.

Smith, S. E., & Gianinazzi-Pearson, V. (1988). Physiological interactions between symbionts in vesicular-arbuscular mycorrhizal plants. Annual review of plant physiology and plant molecular biology, 39(1), 221-244.

Suetsugu, K. (2015). Autonomous self-pollination and insect visitors in partially and fully mycoheterotrophic species of Cymbidium (Orchidaceae). Journal of Plant Research, 128, 115-125.

Tsukaya, H. (2018). How leaves of mycoheterotrophic plants evolved–from the view point of a developmental biologist. New Phytologist, 217(4), 1401-1406.

van den Berg, C., Ryan, A., Cribb, P. J. & Chase, M. W. (2002). Molecular phylogenetics of Cymbidium (Orchidaceae: Maxillarieae): sequence data from internal transcribed spacers (ITS) of Nuclear Ribosomal DNA and Plastid matK. Lindleyana, 17(2), 102–111.

Wang, B., & Qiu, Y. L. (2006). Phylogenetic distribution and evolution of mycorrhizas in land plants. Mycorrhiza, 16(5), 299-363.

Wang, D., Jacquemyn, H., Gomes, S. I., Vos, R. A., & Merckx, V. S. (2021). Symbiont switching and trophic mode shifts in Orchidaceae. New Phytologist, 231(2), 791-800.

Yukawa, T., Miyoshi, K., & Yokoyama, J. (2002). Molecular Phylogeny and Character Evolution of Cymbidium (Orchidaceae). Bull. Natn. Sci. Mus., Tokyo, Ser. B, 28(4), 129–139.

Chapter 1: Molecular phylogeny of *Cymbidium* (Orchidaceae) and its character evolution

Abstract

The genus *Cymbidium* comprises about 75 species mainly distribute in east to south Asia. Because this genus shows remarkable variations in habitats, nutrient modes, and photosynthetic pathways, it will be a good model to study evolutionary trends of the characters. Despite these interesting features, little was known about evolution processes within this genus, because of the incomplete taxon sampling in the previous studies. In this study, I sampled 61 taxa of *Cymbidium* and reconstructed the phylogenetic trees and ancestral states of this genus. I generated nuclear and plastid trees based on two nuclear (*Xdh* and nrITS) and five plastid (*rbcL*, *psaB*, *matK*, *ycf1*, and *rpl16* intron) DNA markers, respectively. My results suggested the monophyly of most sections previously recognized, except for section *Floribundum*, section *Cyperorchis*, and section *Jensoa*. Polyphyly of section *Cyperorchis* and section *Jensoa* might be caused by incomplete lineage sorting, or suggesting hybrid origin of some species within these sections. Section *Floribundum* comprises two different lineages, thus warrants the separation of this section. My trees also suggest the needs of reconsideration for delimitations of some species, e.g., *C. goeringii*, *C. ensifolium* and *C. lancifolium*. Ancestral state reconstructions based on the stochastic mapping approach suggests that epiphytic and autotrophic are the ancestral states of *Cymbidium*. Although my study successfully resolved the overall relationships within the genus, the delimitations of the

"variable" species, and the positions of the morphologically unique species require further work.

Introduction

The genus *Cymbidium* Sw. (Orchidaceae) are epiphytic, lithophytic, or terrestrial herbs comprising about 75 species mainly occurring in Indochina (Du Puy & Cribb, 2007; Liu *et al.*, 2009; WCSP, 2020). *Cymbidium* species show divergence not only in life forms but also in nutritional modes with being composed of autotrophic, mixotrophic and mycoheterotrophic species (Motomura *et al.*, 2010). Moreover, two different types of photosynthetic pathway, C_3 and CAM, are developed within this genus (Motomura *et al.*, 2008). Because the evolution of these characters has been only estimated using a limited number of species or unresolved phylogenetic tree in previous studies (Yukawa *et al.*, 2002; Motomura *et al.*, 2008; Motomura *et al.*, 2010), more detailed estimations using more concrete phylogeny were needed to firmly evaluate the evolutionary dynamics in this genus.

Genus *Cymbidium* established by Swartz (1799) includes several taxa with distinctive morphologies, which led some authors to split this genus to several genera, namely *Jensoa* Raf., *Cyperorchis* Blume, *Iridorchis* Blume, *Arethusantha* Finet, *Pachyrhizanthe* (Schltr.) Nakai, and *Cymbidiopsis* H.J.Chowdhery. These genera are now all recognized as synonyms of *Cymbidium* based on mainly molecular evidence (Du Puy & Cribb, 2007; WCSP, 2020). Although basic infrageneric relationship by molecular phylogeny were provided by van den Berg *et al.* (2002), Yukawa *et al.* (2002) and Zhang *et al.* (2002), these studies utilise only one or two DNA markers, and the support of each branch, especially in the basal nodes, are very low. Zhang *et al.* (2021) utilised eight molecular markers for phylogenetic reconstruction, but the evolutionary rates of the markers were too high and failed to reconstruct the secure phylogenetic tree. Although these previous studies provided the data to delimitate the

sectional ranks of the genus, the taxonomic placements of a few morphologically distinct species, namely *C. elongatum* J.J.Wood, Du Puy & Shim and *C. devonianum* Paxton, are undetermined. Furthermore, although several taxonomic studies provided phylogenetic trees of *Cymbidium*, these exclusively focused on describing species but not on intrageneric classification (e.g. Jiang *et al.*, 2020; Zhou *et al.*, 2020). On the other hand, although some recent studies use plastid genomes of species for phylogenetic reconstruction (e.g. Yang *et al.*, 2013; Jiang *et al.* 2019), the taxon samplings are sparse without any data from the nucleic regions.

The species delimitation among lineages are also confusing especially in several species in section *Jensoa* (Raf.) Schltr. and section *Pachyrhizanthe* Schltr., namely *C. goeringii* (Rchb.f.) Rchb.f., *C. ensifolium* (L.) Sw. and *C. lancifolium* Hook., which are treated as "variable" species. Within them, several similar entities had been described under different names, and most of them are now synonymized although there are noticeable morphological differences among them (Du Puy & Cribb, 2007; WCSP, 2020). However, confirmation or split of each species, especially using macromolecular markers, have never been attempted yet.

In this study, I established the most concrete molecular phylogeny of genus *Cymbidium* using five plastid and two nuclear DNA sequences. To summarise and examine the phylogenetic and infrataxonomic problems of the genus, I primarily focus on the following points: 1, Sectional rank relationship within the genus; 2, Unitary of morphologically variable "species"; 3, Effects of hybridizations against the speciation in the genus; 4, Evolution of character, especially, life forms and nutrient modes.

Materials and Methods

Taxon sampling

71 samples representing 60 taxa (58 species, one subspecies and one variety) out of about 75 extant *Cymbidium* species and three outgroups were included in this analysis (Table 1-1). The samples included the type species of all the 11 sections proposed in Du Puy & Cribb (2007). The outgroups were selected on the bases of Li *et al.* (2016). In the present study, I refer to the taxonomic system proposed in Du Puy & Cribb (2007). Furthermore, I tentatively treat *C. lancifolium*, *C. nagifolium* Masam., and *C. aspidistrifolium* Fukuy. as *C. lancifolium* complex, *C. goeringii s.s.* (Japanese population of *C. goeringii*), *C. goeringii s.l.*, *C. serratum* Schltr., *C. tortisepalum* Fukuy., and *C. formosanum* Hayata as *C. goeringii* complex, and *C. ensifolium*, *C. haematodes* Lindl., *C. nanulum* Y.S.Wu & S.C.Chen, *C. koran* Makino, and *C. munronianum* King & Pantl. as *C. ensifolium* complex. I also preliminary define section *Floribundum s.s.* as section *Floribundum* sensu Du Puy & Cribb (2007) but excluding *C. elongatum*.

DNA extraction, amplification and sequencing

Total DNA was extracted from fresh plant material with DNeasy Plant Mini Kit (Qiagen) according to manufacturer's protocols. Two nuclear markers (*Xdh* and nrITS) and five plastid markers (*rbcL*, *psaB*, *matK, ycf1*, and *rpl16* intron) were amplified. The *Xdh* gene was amplified using two pairs of primers: X502F/Xdh-1513R_2 or X551F/Xdh_849R_2 (Górniak *et al.*, 2010; Li *et al.*, 2016). The nrITS region was amplified using the primer pair 17SE_2/845_R (Sun *et al.*, 1994). The *rbcL* gene was amplified using the primer pair 1F_2/rbcL1345R (Goldman *et al.*,

2001). The *psaB* gene was amplified using the primer pair NY159F_2/NY160R_2 (Cameron, 2004). The *matK* gene was amplified using the primer pair OMAT1F/trnK-2R (Hidayat *et al.*, 2005). The *ycf1* gene was amplified using either primer pairs 3720F/5500R (Neubig *et al.*, 2009) or ycf1N_F/ycf1N_R (designed for this study). The *rpl16* intron was amplified using the primer pair rpL16F1N_2/rpL16R2N (Inoue & Yukawa, 2002). Some primers were newly constructed or modified to match the sequences of *Cymbidium* based on the complete plastid genomes of *C. aloifolium* (L.) Sw. (GenBank accession: KC876122), *C. goeringii* (KT722982), *C. macrorhizon* Lindl. (KU179437) and *C. tracyanum* L.Castle (KC876127). The sequence of the primers used for the amplification were listed in Table 1-2.

The PCR mixture contained 10 μL of 2×MightyAmp Buffer Ver.3 (4 mM $Mg2^{+}$, 600 μM dNTP) (Takara, Japan), 0.3 μL of each primer (20 μM) (Eurofins Genomics, Japan), 0.3 μL MightyAmp DNA Polymerase Ver.3 (1.25 U/μL) (Takara, Japan), 1.0 μL of unquantified template DNA, and H_2O to a Final volume of 20 μL. The PCR condition was as follows: 98° C for 2 minutes; 30 cycles of 20 seconds denaturation at 98° C, 30 seconds annealing at 46-53° C and extension at 68° C for 1-2 minutes. The detailed PCR parameters for all regions are summarised in Table 1-2. PCR products were isolated and purified using GEL/PCR Purification Mini Kit (FAVORGEN, Taiwan) following the manufacturer's protocols. The sequence was performed by Eurofins Genomics (Japan). Sequence reactions were performed in both directions using the same primer sets as for PCR for nrITS and *Xdh*. For the plastid regions, internal sequencing primers were designed in addition to the primer sets for PCR (Table 1-2).

Phylogenetic analysis

Sequence chromatograms were imported into GeneStudio (GeneStudio, Inc.) and assembled

into contig sequences. The previously generated nrITS and *matK* data by Yukawa *et al.* (2002) were also utilised. The sequences were aligned with MUSCLE implemented in MEGA v7.0.21 (Edgar, 2004; Kumar *et al.*, 2016), and alignment errors were manually corrected. Ambiguously aligned characters were excluded prior to the tree building.

Partitioned Maximum Likelihood (ML) analysis was performed using RAxML v8.2.10 (Stamatakis, 2014) under GTR+ γ +I substitution model with four gamma categories. Bootstrap percentages (BP) were calculated from 1,000 bootstrap replicates. Partitioned Bayesian Inference (BI) analysis was performed using MrBayes v3.2.7 (Ronquist *et al.*, 2012) under GTR+ γ +I substitution model. Markov Chain Monte Carlo (MCMC) chains were run for 50,000,000 generations, starting with a random tree. Tree sampling was done once in each 1,000 generations. Posterior probabilities (PP) in a majority-rule consensus tree were assessed with a burn-in of the first 25% of the resulting trees. For several species, the nrITS data in GenBank were utilised if these were available. The phylogenetic position of these species were estimated by adding these sequences to my nrITS matrix. For this matrix, I only performed ML analysis in the same parameters as other trees. I defined BP of 50-74 as weak, 75-84 as moderate, and 85-100 as strong support as in Chase *et al.* (2000). Similarly, PP ≥ 95 was defined as strong support as in Martínez-Azorín *et al.* (2011).

To detect potential incongruences between nuclear and plastid trees, I used the Procrustes Approach to Cophylogeny (PACo; Balbuena *et al.*, 2013) by utilising the R script introduced by Pérez-Escobar *et al.* (2016). The species detected as incongruent were excluded from the combined dataset analysis. I could not perform PACo analysis between two nuclear regions because of the lack of the sequences from several species in Xdh.

Ancestral state reconstructions

I estimated the ancestral states for life forms and existence of subterranean rhizomes after the germination using the stochastic mapping approach implemented in phytools v0.7.47 (Revell, 2012). I estimated ancestral states using the 'make.simmap' function under the ER model, and simulated for 1,000 times. These simulated histories were summarised using the 'summary' function. I utilised an ultrametric plastid and nuclear tree delivered from BEAST v2.6.2 (Bouckaert *et al.*, 2014) as an input tree, respectively. Since there is no fossil record of *Cymbidium* or the related taxon, I set a secondary calibration point at the most recent common ancestor (MRCA) of *Cymbidium* + outgroup (*Grammatophyllum* Blume, *Acriopsis* Reinw. ex Blume and *Thecostele* Rchb.f.), which was dated to 21 ± 8 (95% credible interval) Ma based on Pérez-Escobar *et al.* (2017). I used the GTR+ γ +I substitution model with four rate categories, a Yule tree prior, and an MCMC chain of 100,000,000 generations, with parameters and trees sampled every 1,000 generations. I used Tracer v1.7 (Rambaut *et al.*, 2018) to check that the runs had converged. Trees were summarised in TreeAnnotator v1.8 with 25% of the trees as burn-in.

Life forms were listed based on Du Puy & Cribb (2007). Some *Cymbidium* species form subterranean rhizomes after they germinate, and this trait is considered as a feature of mixo- or mycoheterotrophic species (Du Puy & Cribb, 2007; Motomura *et al.*, 2010) (Table 1-3).

Results

Sequences and alignment

I generated a total of 444 sequences. Few accessions failed to be amplified for certain regions therefore I treated them as missing data. No large deletions were found except for the *rpl16* intron region, thus I excluded poorly aligned regions from the sequences manually, and divided the sequence into three partitions. Properties for each dataset are provided in Table 1-5. As a summary, for the nuclear regions, the nrITS region, and for the plastid regions, *rpl16* intron yielded the highest parsimony-informative characters. The combined plastid data represented a greater number of the parsimony-informative characters than the combined nuclear data, while the nuclear data showed a higher percentage of it.

Phylogenetic reconstruction and incongruence between nuclear and plastid loci

ML trees and BI trees showed similar topology (Figs. 1-1, 1-2). All the sections except for section *Floribundum* were monophyletic, at least in either plastid or nuclear tree as follows. (1) Monophyletic in both nuclear and plastid tree: section *Austrocymbidium*, section *Borneense*, section *Cymbidium*, and section *Pachyrhizanthe*. (2) Monophyletic in the nuclear tree but not in the plastid tree: section *Jensoa.* (3) Monophyletic in the plastid tree but not in the nuclear tree: section *Cyperorchis.* (4) Polyphyletic in both nuclear and plastid tree: section *Floribundum.* (5) Monotypic section: section *Annamaea*, section *Bigibbarium*, section *Himantophyllum*, and section *Parishiella*. The nuclear tree supports the monophyly of each section presented in Du Puy & Cribb (2007) with moderate to high support (BP ≥ 82, PP ≥ 99)

except for section *Cyperorchis* (Blume) P.F.Hunt and section *Floribundum* (Fig. 1-2). However, species especially in section *Jensoa* and section *Cyperorchis,* were left unsolved exclusively by the nuclear tree. *Cymbidium nagifolium*, the species often treated as a heterotypic synonym of *C. lancifolium*, is placed as a sister of *C. macrorhizon* and *C. nipponicum* (Franch. & Sav.) Rolfe, not as other *C. lancifolium* complex species. The *C. ensifolium* complex shows paraphyly, but the *C. goeringii* complex is monophyletic.

The tree reconstructed by the plastid dataset is shown in Figure 1-1. Monophyly of all the sections presented in Du Puy & Cribb (2007) other than section *Floribundum* and section *Jensoa,* were confirmed. Placement of *C. cyperifolium* Wall. subspp. which are often treated as section *Jensoa*, were unsolved in BI tree, or placed as sister to section *Pachyrhizanthe*, not in section *Jensoa* in ML tree although without the support (BP < 50). In species rank, the *C. goeringii* complex and the *C. ensifolium* complex showed paraphyly, but the *C. lancifolium* complex showed monophyly unlike the nuclear tree.

PACo detected 24 taxa of the genus as incongruent between nuclear and plastid trees, thus I excluded these species from the combined analysis. The combined tree shows similar topology with the plastid tree, although species from three sections are totally removed (Figs. 1-3, 1-S4). Three out of four species in section *Floribundum* s.s. were excluded from the analysis, although they seem to be congruent. In other cases, some taxa, i.e. *C. nagifolium,* seem to be incongruent despite PACo has detected them as congruent.

Ancestral state reconstructions

The ancestral states of the genus were reconstructed as epiphyte and autotroph in both nuclear and plastid trees (Fig. 1-4). Based on the plastid tree, terrestrial species were evolved

five times in different clades of the genus, namely (1), in the common ancestor of section *Borneense* Du Puy & P.J.Cribb; (2), in the common ancestor of section *Jensoa* and section *Pachyrhizanthe*; (3), ancestor of *C. hartinahianum* J.B.Comber & Nasution; (4), *C. insigne* Rolfe; and (5) *C. elongatum,* respectively. The nuclear tree showed a similar trend but estimated the evolution of terrestrial species as four times, namely (1), in the common ancestor of section *Borneense* Du Puy & P.J.Cribba and *C. elongatum*; (2), in the common ancestor of section *Jensoa* and section *Pachyrhizanthe*; (3), ancestor of *C. hartinahianum* J.B.Comber & Nasution; and (4), *C. insigne* Rolfe. Species with subterranean rhizomes had evolved from the non-rhizomatous species just once in the common ancestor of section *Jensoa* and section *Pachyrhizanthe*. Mycoheterotrophic species evolved once from the non-mycoheterotrophic species with rhizomes in section *Pachyrhizanthe*.

Discussion

Phylogenetic reconstruction and conflicts between nuclear and plastid trees

My phylogenetic tree includes more comprehensive species compared to the previous studies. The phylogenetic tree based on more inclusive genetic markers successfully resolved the relationship between most of the sections. The topology of my combined and plastome trees generally agree with the plastome trees formerly published (e.g. Yang *et al.*, 2013; Jiang *et al.* 2019). Although the phylogenetic trees based on nuclear and plastid proposed by Zhang *et al.* (2021) show polyphyly in most sections suggesting the hybrid origin of them, my

phylogenetic tree did not show such a tendency. This might be because the genetic regions utilised in Zhang *et al.* (2021) have too high evolutionary rates which might be caused by artificial errors, such as sequencing error. My nuclear tree showed relatively low BP support especially in terminal branches. This is probably because my nuclear data set is not enough for the concrete phylogenetic reconstruction. I failed to sequence the *Xdh* region of section *Borneense* and *C. elongatum*, which might decrease the support of the basal nodes. On the other hand, the plastid trees showed comparatively higher support in each basal node, but had lower support in the nodes containing the diverse groups, namely *C. ensifolium* complex and its relatives, and some nodes in section *Cyperorchis*. PACo detected 24 incongruent species between nuclear and plastid data. Among them, 17 species are conflicting within the focal section, two species are from the monotypic section (*C. dayanum* Rchb.f. and *C. erythrostylum* Rolfe), three species seem to be misdetection of incongruence (section *Floribundum s.s.* except for *C. floribundum* Lindl.), and two species seem to be detected as incongruent due to the lack of data (*C. aliciae* Quisumb. and *C. borneense* J.J.Wood). The conflicts mainly occurred within each section, which should reflect the hybrid origin of these species, or the incomplete lineage sorting (Degnan & Rosenberg, 2009).

Infrageneric relationships of *Cymbidium*

Overview of the sectional relationships within the genus

Because the sectional phylogenetic relationships have been poorly studied for *Cymbidium*, the subgeneric delimitations are still unclear (Du Puy & Cribb, 2007). Here, I discuss the phylogenetic position and the relationships of sections referring to the scenario proposed in Du Puy & Cribb (2007). The overview of my infrataxonomic recognition is provided in Figure

1-5. All the phylogenetic trees constructed in this study have four basal clades in general. Here, I tentatively define four clades based on combined tree and morphology for the following discussions: *Cymbidium* clade (section *Cymbidium* P.F.Hunt, section *Borneense*, section *Austrocymbidium* Schltr., and *C. elongatum*); *Jensoa* clade (section Jensoa, section *Pachyrhizanthe*, and section *Floribundum s.s.*); *Cyperorchis* clade (section *Cyperorchis*, section *Annamaea* (Schltr.) P.F.Hunt, section *Himantophyllum* Schltr., and section *Parishiella* (Schltr.) P.F.Hunt); and *Bigibbarium* clade (section *Bigibbarium* Schltr. only) as shown in Figure 1-5. The former three clades roughly correspond to three subgenera, subgenus *Cymbidium*, subgenus *Jensoa* (Raf.) Seth. & P.J.Cribb, and subgenus *Cyperorchis* (Blume) Seth & Cribb proposed in Du Puy & Cribb (1988). Exact branching order of each clade is unclear due to < 50 BS/BI values of the basal node, especially for the position of *C. devonianum*. By eliminating *C. devonianum* from the plastid tree, the relationship between *Cymbidium* clade and *Jensoa* clade becomes more secure with high BS support (Fig. 1-S1). This topology is also found in trees previously constructed from the plastid genomes, although only a few species are recruited in the trees (Yang *et al.*, 2013). Some of the species, especially those in clade *Jensoa* and clade *Cyperorchis* are doubtful, which might be described based on hybrid origin or abnormal individuals (Du Puy & Cribb 2007). Careful reexamination of the specimens along with reconsideration of "morphological variable" species are needed. Although I implied the genetic regions with low evolutionary rate, namely *psaB*, *rbcL* and *Xdh*, for the sequencing, I failed to reconstruct the branching order of these basal nodes, suggesting the rapid speciation of these species on ancestral lineages. Further data, for example, plastome, MIG-seq or RAD-seq data may resolve this problem.

Cymbidium *clade*

The *Cymbidium* clade contains three sections and one species, namely section *Cymbidium*, section *Borneense*, section *Austrocymbidium* and *C. elongatum*. The plastid tree supports the monophyly of this clade, and nuclear trees also support this monophyletic clade if excluding *C. elongatum* and section *Borneense*. I failed to sequence the *Xdh* gene of *C. elongatum* and any species in section *Borneense,* which might affect the unstable topology of the tree. The common character among these sections in this clade are unknown. According to my phylogenetic trees, all the species in this clade show relatively long branches compared to the species in the other clades (Figs. 1-3). This suggests the lower diversification rates in this clade, although its causes are unknown.

Morphologically unique species, *C. elongatum,* had been tentatively placed in section *Floribundum* by Du Puy & Cribb (2007), but all the trees provided in the present study rejected this treatment. Both the combined and plastid trees placed this species sister to section *Cymbidium*, section *Borneense*, and section *Austrocymbidium*, not to section *Floribundum*. Although I failed to sequence the *Xdh* region of this species, the nuclear tree also denied putting this species in section *Floribundum*, but on sister with section *Borneense*. However, since I failed to sequence the *Xdh* region of section *Borneense* also, this topology might be changed. Because *C. elongatum* has quite a unique morphology, especially the elongating stem, and for the maintenance of the monophyly of section *Floribundum* and other sections in *Cymbidium* clade, it is worth to place this species in a new section, although more examination on morphology, especially in vegetative anatomy, are needed.

Jensoa *clade*

This clade contains three sections, which are section *Jensoa*, section *Pachyrhizanthe*, and section *Floribundum s.s.* The common character among the species in this clade is unknown. Although there is no morphological support for the relatedness among these three sections in this clade (Yukawa & Stern 2002; Du Puy & Cribb, 2007), highly supported nodes of combined and plastid trees suggest the monophyly of this clade. Section *Nanula* which was described in Liu *et al.* (2006) with non-articulate leaved species, namely *C. nanulum*, *C. faberi*, *C. serratum* and *C. tortisepalum*, are not monophyletic without any support of my trees. Thus, separation of section *Nanula* from section *Jensoa* is not supported. Section *Axillaria*, the monotypic section described in Liu *et al.* (2006), includes only *C. cyperifolium* and its subspecies and varieties. *Cymbidium cyperifolium* have unique vegetative morphology, in particular, numerous leaves which are distichous-equitant at base, and axillary scape. My plastid tree supports the specificity of this section, but the nuclear tree prevents it. There is little merit for the set of this section, thus I do not support the concept of section *Axillaria*.

Although some previous authors moved *C. macrorhizon* and its relative species to other genus, e.g. *Pachyrhizanthe* (Schltr.) Nakai or *Cymbidiopsis* H.J.Chowdhery, my results showed that these treatments are not reasonable by the point of maintaining the monophyly of the genus *Cymbidium*.

Cyperorchis *clade*

This clade contains four sections, namely section *Cyperorchis*, section *Parishiella*, section *Annamaea*, and section *Himantophyllum*. The combined and plastid trees provided by the present study supports this clade with moderate and high bootstrap values, respectively. The

synapomorphy of this clade is the fusion of lip and column at the base with the exception of section *Himantophyllum* and *C. daweishanense* G.Q.Zhang & Z.J.Liu, the species resembles with *C. erythraeum* Lindl. recently described from China (Du Puy & Cribb, 2007; Zhang *et al.*, 2018). I could not resolve the phylogenetic relationships among section *Cyperorchis*, section *Annamaea*, and section *Himantophyllum* by my nuclear data, which suggests the hybrid origin of each section, or the incomplete lineage sorting. Further data delivered from nuclear by MIG-seq or RAD-seq might be needed to resolve this problem. Although the four sections in this clade have distinctive morphological traits, nuclear data suggest the needs of reconsidering the maintenance of these four sections. This clade might have undergone a complex evolutionary history with interspecific hybridizations, thus further investigations with more numerical data are needed to fully resolve its evolutionary history.

Bigibbarium *clade*

The section *Bigibbarium*, consisting of only *C. devonianum*, is tentatively placed as a distinct clade. *Cymbidium devonianum* shows the equivocal morphology, such as pollinia morphology similar to section *Cymbidium* and leaf margin morphology similar to section *Cyperorchis*, prevent the reliable classification of the species. However, the floral morphology of this species is unique, which strongly suggests independence of the species (Du Puy and Cribb, 2007). Any of the previous studies also failed to decide its infrageneric position (Yukawa and Stern, 2002; van den Berg *et al.*, 2002; Yukawa *et al.*, 2002; Zhang *et al.*, 2021). The additional molecular data with lower evolutionary rates might resolve this problem.

Unplaced taxa

Notably, two species that have recently been described, namely *C. repens* Aver. & Q.T.Phan and *C. lii* M.Z.Huang, J.M.Yin & G.S.Yang, have stemlike pseudobulbs which are connected with each other by long plagiotropic rhizome (Averynov *et al.*, 2016; Huang *et al.*, 2017). This morphological character is unique among the species in this genus, which supports the need for the placement of these species in a new section. Although I could not include these species into my study, further investigation on these species will provide important insights to the infrataxonomy of this genus.

Evolutionary trajectories in species-rich sections and unitary of morphologically variable "species"

Phylogenetic relationships among species in species-rich sections, especially section *Jensoa* and section *Cyperorchis*, were still unclear. Incongruences of nuclear and plastid trees and shallow branching patterns in those sections suggest the occurrence of interspecific hybridizations and recent rapid speciation within these sections.

Morphologically variable "species", particularly the *C. ensifolium* complex, the *C. goeringii* complex and the *C. lancifolium* complex, consist of several morphologically distinguishable entities. Although species in these complexes were often regarded as conspecific, my trees suggest the independence of some entities. For example, I treat all the species in the *C. goeringii* complex, namely *C. goeringii s.l.*, *C. serratum*, *C. tortisepalum* and *C. formosanum*, as distinct species according to their unique morphologies, though Du Puy & Cribb (2007) treated *C. serratum* and *C. formosanum* conspecific with *C. goeringii*, and Liu *et al.* (2009) treated *C. formosanum* conspecific with *C. goeringii.* Although my nuclear tree unified these species complex into one clade, the plastid tree placed them into two different

clades, which suggest the hybrid origins of some lineage in this species complex. Species in the *C. ensifolium* complex are also regarded as conspecific in some studies. Du Puy & Cribb (2007) placed *C. koran* and *C. haematodes* as conspecific with *C. ensifolium* (latter as subsp. *haematodes* (Lindl.) Du Puy & Cribb). Other authors often place *C. koran* as conspecific with *C. ensifolium*. Although I could not warrant the correspondence of my samples of *C. ensifolium* and *C. haematodes* to the type specimen of each taxon, my phylogenetic trees suggest that there are several species in *C. ensifolium s.l.*

Cymbidium lancifolium, *C. nagifolium* and *C. aspidistrifolium* are often treated as conspecific. Although my plastid data showed monophyly of this species complex, the nuclear data suggested the independence of *C. nagifolium*, placing it sister to mycoheterotrophic species, results in the paraphyly of the *C. lancifolium* complex. The observation of the auto self-pollination in *C. nagifolium* and *C. macrorhizon* in wild populations also supports the close connection between these species (Suetsugu, 2015). Several species, namely *C. rhizomatosum* Z.J.Liu & S.C.Chen, *C. biflorens* D.Y.Zhang, S.R.Lan & Z.J.Liu and *C. multiradicatum* Z.J.Liu & S.C.Chen, show intermediate morphology between *C. lancifolium* and *C. macrorhizon*, which suggest the phylogenetic continuity of *C. lancifolium* and *C. macrorhizon*. The nrITS tree placed former two species to the sister of *C. macrorhizon* also suggests the existence of the third group between *C. lancifolium*, the mixotrophic species with large leaves, and *C. macrorhizon*, mycoheterotrophic species with no leaves (Zhang *et al.*, 2020).

Although my molecular data suggests the heterospecificity of each species complex, examination of type materials, and vast molecular data from sufficient numbers of samples are needed to resolve this complex taxonomic status.

Evolution of traits in genus *Cymbidium*

Typically, most plant taxa evolve the epiphytic lifestyle from the terrestrial one, which could be applied to some Orchidaceae groups (Givnish *et al.,* 2015). However, my ancestral trait reconstruction suggests the opposite evolutional direction, which means the ancestor of the genus is epiphytic, and the terrestrial species evolved at least five times independently. Throughout the tribe Cymbidieae, terrestrial lineages are the minority (e.g. *Eulophia* R.Br., *Cyanaeorchis* Barb.Rodr. and *Guanchezia* G.A.Romeo & Carnevali) suggesting the MRCA of this tribe to be epiphytic (Pérez-Escobar *et al.*, 2017). In Polypodiaceae ferns, terrestrial species have evolved from the epiphytic lineage and are more radiated, because they explored novel habitat types (Sundue *et al.*, 2015). This scenario might also explain the higher number of species in the *Jensoa* clade.

The evolution of terrestrial lifestyle might also involve the evolution of mixo- and mycoheterotrophic species. Ogura-Tsujita *et al.* (2012) showed the shifts in mycorrhizal fungi species composition in mixo- and mycoheterotrophic species in *Cymbidium*. These species showed symbiotic relationship with ectomycorrhizal Sebacinales, whereas epiphytic species had no association with these fungi. Ectomycorrhizal fungi should be poor in an epiphytic environment, thus terrestrial habitat is required for the evolution of mixo- and mycoheterotrophy. Although the adaptation to terrestrial habitats has occurred at least four times in the genus *Cymbidium*, only one clade (section *Jensoa* + section *Pachyrhizanthe*) evolved the specific mycorrhiza. This suggests that there might be additional conditions other than terrestrial habitat for constructing the novel mutual relationships. Investigating and comparing the mycorrhiza of each terrestrial lineage and its sister lineages, especially section

Floribundum, the sister clade of mixotrophic lineages, might provide new insight for this concept. Most species in section *Jensoa* and section *Pachyrhizanthe* distribute in temperate areas or tropical areas with dry season (Du Puy & Cribb, 2007), suggesting the needs of dormancy. Currently, some species, namely *C. defoliatum*, *C. tamphianum* and *C. rhizomatosum* show vegetative dormancy (Liu *et al.*, 2006; Liu *et al.*, 2009; Averyanov *et al.*, 2018). Some of these species have no leaves when anthesis, suggesting the initial state of mycoheterotrophic species. Ectomycorrhizal fungi might increase the survival rate during the dormancy, hence suggesting the mycoheterotrophic lifestyle are extent from the dormancy (Suetsugu pers. com.).

Conclusions

I reconstructed the most species rich phylogeny of the genus *Cymbidium* organising the phylogenetic relationship within the genus as well as providing the insights for character evolution. The genus *Cymbidium* shows diverse lifestyles and morphology within the genus, thus the thorough investigation of the genus could undoubtedly bring new insights to the studies of plant evolution. My study suggests that the hybridizations are involved in speciation of some lineages. More detailed data of the nuclear region, i.e. from RAD-seq or MIG-seq might solve this complex evolutionary history of the genus. The character evolution was closely connected with the lifestyle of the species. Especially, the occurrence of mycoheterotrophic species in the genus were suggested to be evolved from the extension of the vegetative dormancy in the mixotrophic species. Further examination in genus *Cymbidium* may provide

crucial insights into character evolution caused by the shifts in the life forms, including the evolution of mycoheterotrophic plants.

References

Averyanov, L. V., Nong, V. D., Nguyen, K. S., Maisak, T. V., Nguyen, V. C., Phan, Q. T., Nguyen, P. T., Nguyen, T. T. & Truong, B. V. (2016) New species of Orchids (Orchidaceae) in Flora of Vietnam. Taiwania, 61(4), 319-354. https://doi.org/10.6165/tai.2016.61.319

Averyanov, L.V., Nguyen, V.C., Truong, B.V., Maisak, T.V., Luu, H.T., Nguyen, K.S., Dinh, Q.D., Nguyen, H.T., Chu, X.C., Tran, G., Nguyen, V.K., & Le, .HS. (2018). New Orchids (Orchidaceae: Cymbidieae and Vandeae) in the Flora of Vietnam. Taiwania, 63(2), 119-138. 10.6165/tai.2018.63.119.

Balbuena JA, Míguez-Lozano R, Blasco-Costa I. (2013) PACo: a novel procrustes application to cophylogenetic analysis. PLoSONE8: e61048. Barkman

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 Computational Biology, 10(4), e1003537. doi:10.1371/journal.pcbi.1003537

Cameron, K. M. (2004) Utility of plastid psaB gene sequences for investigating intrafamilial relationships within Orchidaceae. Molecular Phylogenetics and Evolution, 31(3), 1157–1180. https://doi.org/10.1016/j.ympev.2003.10.010

Chase MW, de Bruijn AY, Cox AV, Reeves G, Rudall PJ, Johnson MAT, Eguiarte LE. (2000) Phylogenetics of Asphodelaceae (Asparagales): an analysis of plastid rbcL and trnL-F DNA sequences. Annals of Botany 86:935-951.

Chen, G.Z., Zhang, G.Q., Huang, J., Wang, M., Rao, W.H., & Chen, L.J. (2019) *Cymbidium shidianense* (Orchidaceae; Epidendroideae), a new species from China: evidence from morphology and molecular data. Phytotaxa, 399(1): 100-108. https://doi.org/10.11646/phytotaxa.399.1.10

Du Puy, D. & Cribb, P.J. (2007) The genus *Cymbidium*. Royal Botanic Gardens, Kew, Richmond.

Degnan, J. H., and N. A. Rosenberg. (2009) Gene tree discordance, phylogenetic inference and the multispecies coalescent. Trends Ecol. Evol. 24:332–340.

Edgar, R.C. (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput, Nucleic Acids Research. 32(5), 1792–1797.

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. Proceedings of the Royal Society B: Biological Sciences, 282(1814), 20151553. https://doi.org/10.1098/rspb.2015.1553

Goldman, D. H., Freudenstein, J. V, Kores, P. J., Molvray, M., Jarrell, D. C., Whitten, W. M., Cameron, K. M., Jansen, R. K., Chase, M. W. (2001) Phylogenetics of Arethuseae (Orchidaceae) Based on Plastid matK and rbcL Sequences. Systematic Botany, 26(3), 670–695. https://doi.org/10.1043/0363-6445-26.3.670

Górniak M, Paun O, Chase MW. Phylogenetic relationships within Orchidaceae based on a low-copy nuclear coding gene, Xdh: Congruence with organellar and nuclear ribosomal

DNA results. (2010) Molec Phylogen Evol. 56(2), 784–795. http://dx.doi.org/10.1016/j.ympev.2010.03.003.

Huang, M.Z., Liu, Z.L., Yang, G.S. & Yin, J.M. (2017) An unusual new epiphytic species of *Cymbidium* (Orchidaceae: Epidedroideae) from Hainan, China. Phytotaxa. 314(2): 289-293. https://doi.org/10.11646/phytotaxa.314.2.12

Inoue, K. & Yukawa, T. (2002) A New Species of *Yoania*(Orchidaceae) from Southern Nagano, Central Japan. Acta Phytotax, Geobot. 53 (2), 107–114

Jiang, Y. T., Lin, R. Q., Liu, B., Zeng, Q. M., Liu, Z. J. & Chen, S. P. (2019) Complete chloroplast genome of *Cymbidium ensifolium* (Orchidaceae). Mitochondrial DNA Part B, 4:2, 2236-2237, DOI: 10.1080/23802359.2019.1624637

Jiang, Y. T., Ma, L., Lin, R. Q., Xiong, Y. F., Jiang, S. Z., Liu, B., Liu, Z. J., & Chen, S. P. (2020) *Cymbidium codonanthum* (Orchidaceae; Epidendroideae; Cymbidiinae), a new species from China: Evidence from morphological and molecular analyses. Phytotaxa, 453(3), 275–283. https://doi.org/10.11646/phytotaxa.453.3.9

Kumar S, Stecher G, and Tamura K (2016) MEGA7: Molecular Evolutionary Genetics Analysis version 7.0 for bigger datasets. Molecular Biology and Evolution 33:1870–1874.

Lan, S., Chen, L., Chen, G., Wu, X., Rao, W., Zhang, P. and Liu, Z. (2018) *Cymbidium densiflorum* (Orchidaceae; Epidendroideae; Cymbidieae): a new orchid species from China based on morphological and molecular evidence. Phytotaxa 345(1), 51-58. https://doi.org/10.11646/phytotaxa.345.1.6

Li, M. H., Zhang, G. Q., Liu, Z. J., & Lan, S. R. (2016) Subtribal relationships in Cymbidieae (Epidendroideae, Orchidaceae) reveal a new subtribe, Dipodiinae, based on plastid

and nuclear coding DNA. Phytotaxa, 246(1), 037–048. https://doi.org/10.11646/phytotaxa.246.1.3

Liao, X.Y., Liu, X.D., Jiang, Y.T., Zhang, D.Y., Lan, S.R. and Liu Z.J. (2019) *Cymbidium atrolabium* (Orchidaceae; Epidendroideae), a new species from China: evidence from morphological and molecular data. Phytotaxa, 423(3): 87-92. https://doi.org/10.11646/phytotaxa.423.3.4

Liu, Z.J., Chen, S.C., & Ru, Z.Z. (2006) The genus *Cymbidium* in China. Science Press, Beijing.

Liu, Z.J., Chen, X.Q., & Cribb, P.J. (2009) *Cymbidium*. In: wu, Z.Y., Raven, P.H. & Hong, D. (Eds.) Flora of China, vol. 25. Science Press, Beijing & Missouri Botanical Garden Press, St. Louis, pp. 260–280.

Martínez-Azorín M., Crespo M.B., Juan A., & Fay M.F. (2011) Molecular phylogenetics of subfamily Ornithogaloideae (Hyacinthaceae) based on nuclear and plastid DNA regions, including a new taxonomic arrangement. Annals of Botany 107:1 ?37.

Motomura, H., Selosse, M.A., Martos, F., Kagawa, A., & Yukawa, T. (2010) Mycoheterotrophy evolved from mixotrophic ancestors: Evidence in *Cymbidium* (Orchidaceae). Annals of Botany, 106(4), 573–581. https://doi.org/10.1093/aob/mcq156

Motomura, H., Yukawa, T., Ueno, O., & Kagawa, A. (2008) The occurrence of crassulacean acid metabolism in *Cymbidium* (Orchidaceae) and its ecological and evolutionary implications. Journal of Plant Research, 121(2), 163–177. https://doi.org/10.1007/s10265-007-0144-6

Neubig, K.M., Whitten, W.M., Carlsward, B.S., Blanco, M.A., Endara, L., Williams, N.H., & Moore, M. (2009) Phylogenetic utility of ycf1 in orchids: A plastid gene more variable than matK.

Plant Systematics and Evolution, 277(1–2), 75–84. https://doi.org/10.1007/s00606-008-0105-0

Ogura - Tsujita, Y., Yokoyama, J., Miyoshi, K., & Yukawa, T. (2012) Shifts in mycorrhizal fungi during the evolution of autotrophy to mycoheterotrophy in Cymbidium (Orchidaceae). American Journal of Botany, 99(7):1158-1176.

Peng, Y.L., Zhou, Z., Lan, S.R., & Liu Z.J. (2019) *Cymbidium jiangchengense* (Orchidaceae; Epidendroideae; Cymbidiinae), a new species from China: evidence from morphology andDNA sequences. Phytotaxa, 408(1):77-84. https://doi.org/10.11646/phytotaxa.408.1.6

Pérez-Escobar, O.A., Balbuena, J.A., & Gottschling, M. (2016) Rumbling orchids: how to assess divergent evolution between chloroplast endosymbionts and the nuclear host. Systematic Biology 65:51–65

Pérez-Escobar, O.A, Chomicki, G., Condamine, F.L., Karremans, A.P., Bogarín, D., Matzke, N.J., Silvestro, D., & Antonelli, A. (2017) Recent origin and rapid speciation of Neotropical orchids in the world's richest plant biodiversity hotspot. New Phytologist. 215:891–905.

Rambaut, A., Drummond, A.J., Xie, D., Baele, G. & Suchard, M.A. (2018) Posterior summarisation in Bayesian phylogenetics using Tracer 1.7. Systematic Biology. syy032. doi:10.1093/sysbio/syy032

Revell, L. J. (2012) phytools: An R package for phylogenetic comparative biology (and other things). Methods Ecol. Evol., 3, 217–223.

Ronquist, F., M. Teslenko, P. van der Mark, D.L. Ayres, A. Darling, S. Höhna, B. Larget, L. Liu, M.A. Suchard, and J.P. Huelsenbeck. (2012) MRBAYES 3.2: Efficient Bayesian phylogenetic inference and model selection across a large model space. Syst. Biol. 61:539-542.

Stamatakis, A. (2014) RAxML Version 8: A tool for Phylogenetic Analysis and Post-Analysis of Large Phylogenies. Bioinformatics, 30(9), 1312–1313.

Suetsugu, K. (2015) Autonomous self-pollination and insect visitors in partially and fully mycoheterotrophic species of Cymbidium (Orchidaceae). J. Plant Res., 128(1):115-125. doi: 10.1007/s10265-014-0669-4.

Sun Y, Skinner DZ, Liang GH, Hulbert SH (1994) Phylogenetic analysis of sorghum and related taxa using internal transcribed spacers of nuclear ribosomal DNA. Theor Appl Genet 89, 26–32

Sundue, M. A., Testo, W. L., & Ranker, T. A. (2015). Morphological innovation, ecological opportunity, and the radiation of a major vascular epiphyte lineage. Evolution, 69(9), 2482–2495. https://doi.org/10.1111/evo.12749

Swartz, O. (1799). Dianome Epidendri generis Linn. Nova Acta Regiae Soc. Sci. Upsal. 6, 61–88.

van den Berg, C., Ryan, A., Cribb, P. J. & Chase, M. W. (2002). Molecular phylogenetics of *Cymbidium* (Orchidaceae: Maxillarieae): sequence data from internal transcribed spacers (ITS) of Nuclear Ribosomal DNA and Plastid matK. Lindleyana, 17(2), 102–111.

WCSP (2020). 'World Checklist of Selected Plant Families. Facilitated by the Royal Botanic Gardens, Kew. Published on the Internet; http://wcsp.science.kew.org/ Retrieved Dec. 1, 2020.

Xu, X., Ding, C.C., Hu, W., Yu, X., Zheng, Y., Zhang, D., Liao, X., Liu, X., Liu, Z.J. and Lan, S. (2021). *Cymbidium xichouense* (Orchidaceae; Epidendroideae), a new species from China:

evidence from morphological and molecular data. Phytotaxa, 484(3): 291-297. https://doi.org/10.11646/phytotaxa.484.3.4

Yang, J.-B., Tang, M., Li, H.-T., Zhang, Z.-R., & Li, D.-Z. (2013). Complete chloroplast genome of the genus *Cymbidium*: lights into the species identification, phylogenetic implications and population genetic analyses. BMC Evolutionary Biology, 13(1), 84. https://doi.org/10.1186/1471-2148-13-84

Yukawa, T., Miyoshi, K., & Yokoyama, J. (2002). Molecular Phylogeny and Character Evolution of *Cymbidium* (Orchidaceae). Bull. Natn. Sci. Mus., Tokyo, Ser. B, 28(4), 129–139.

Yukawa, T. & Stern, W. L. (2002). Comparative vegetative anatomy and systematics of Cymbidium (Orchidaceae). Botanical Journal of the Linnean Society, 138(1), 21-41. https://doi.org/10.1111/j.1095-8339.2009.00818.x

Zhang, D.Y., Tu, X.D., Liu, B., Liao, X.Y., Lan, S.R., & Liu, Z.J. (2020) *Cymbidium biflorens* (Orchidaceae; Epidendroideae), a new species from China: evidence from morphological and molecular data. Phytotaxa, 428(3): 271-278. https://doi.org/10.11646/phytotaxa.428.3.8

Zhang, G.Q., Chen, G.Z., Liu, Z.J. and Lan, S.R. (2018) *Cymbidium daweishanense*, a new species from China: evidence from morphological and molecular analysis. Phytotaxa. 374(3): 249-256. https://doi.org/10.11646/phytotaxa.374.3.6

Zhang, G.Q., Chen, G.Z., Chen, L.J. and Lan, S.R. (2019) *Cymbidium yunnanensis*: a new orchid species (Orchidaceae; Epidendroideae) from China based on morphological and molecular evidence. Phytotaxa. 387(2): 149-157. https://doi.org/10.11646/phytotaxa.387.2.7

Zhang, G.Q., Chen, G.Z., Chen, L.J., Zhai, J.W., Huang J., Wu, X.Y., Li M.H., Peng, D.H., Rao, W.H., Liu, Z.J. and Lan, S.R. (2021). Phylogenetic incongruence in *Cymbidium* orchids. Plant Diversity. 43(5): https://doi.org/10.1016/j.pld.2021.08.002

Zhang, M. Y., Sun, C. Y., Hao, G., Ye, X. L., Liang, C. Y. & Zhu, G. H. (2002). A Preliminary analysis of phylogenetic relationship in *Cymbidium* (Orchidaceae) based on nrITS sequence data. Acta Botanica Sinica. 44(5), 588–592

Zhou, Z., Zhang, D., Chen, G. Z., Yu, X., Li, C. R., Lan, S. R., & Liu, Z. J. (2020). *Cymbidium brevifolium* (Orchidaceae; Epidendroideae), a new species from China: evidence from morphological and molecular data. Phytotaxa, 464(3), 236–242. https://doi.org/10.11646/phytotaxa.464.3.5

Figures and Tables

Table 1-1. A list of taxa and their voucher information used in this study. ○: sequence obtained in this study. ◎: sequence obtained in Yukawa *et al.*, 2002. HT indicates Holotype.

Taxa	Voucher	*Xdh*	nrITS	*rbcL*	*psaB*	*matK*	*ycf1*	*rpl16* intron
Cymbidium aliciae Quisumb.	TBG119071	-	◎	○	○	◎	○	◎
C. aloifolium (L.) Sw.	TBG118253	○	◎	○	○	◎	○	◎
C. aspidistrifolium Fukuy.	TBG122348	○	◎	○	○	◎	○	◎
C. atropurpureum (Lindl.) Rolfe	KC48	○	○	○	○	○	○	○
C. banaense Gagnep.	KC50	○	○	○	○	○	○	○
C. bicolor subsp. *pubescens* (Lindl.) Du Puy & P.J.Cribb	TBG118509	○	◎	○	○	◎	○	◎
C. borneense J.J.Wood	KCR2	-	-	-	-	◎	-	◎
C. borneense	KC24	-	○	○	○	-	○	-
C. canaliculatum R.Br.	TBG135991	-	-	-	-	◎	-	◎
C. canaliculatum	KC47	-	○	○	○	-	○	-
C. chloranthum Lindl.	TBG127438	○	◎	○	○	◎	○	◎

C. cochleare Lindl.	KCT12	○	◎	○	○	◎	○	◎
C. crassifolium Herb.	TBG124328	○	◎	○	○	◎	○	◎
C. cyperifolium Wall. subsp. *cyperifolium*	KC46	○	○	○	○	○	○	○
C. cyperifolium subsp. *indochinense* Du Puy & P.J.Cribb	KC42	○	○	○	○	○	○	○
C. dayanum Rchb.f.	TBG127899	○	◎	○	○	◎	○	◎
C. devonianum Paxton	TBG119073	○	◎	○	○	◎	○	◎
C. eburneum Lindl.	KC18	○	◎	○	○	◎	○	◎
C. elegans Lindl.	KCT07	○	◎	○	○	◎	○	◎
C. elongatum J.J.Wood, Du Puy & Shim	KCR5	-	-	-	-	◎	-	◎
C. elongatum J.J.Wood	KC22	-	○	○	○	-	○	-
C. ensifolium (L.) Sw.	TBG1182555	○	◎	○	○	◎	○	◎
C. erythraeum Lindl.	TBG135399	○	◎	○	○	◎	○	◎
C. erythrostylum Rolfe	TBG126650	-	-	-	-	◎	-	◎
C. erythrostylum	TBG144174	○	○	○	○	-	○	-
C. faberi Rolfe	TBG133593	-	-	-	-	◎	-	◎
C. faberi	KC6	○	○	○	○	-	○	-
C. finlaysonianum Lindl.	TBG122826	○	◎	○	○	◎	○	◎
C. floribundum Lindl.	TBG127357	○	◎	○	○	◎	○	◎
C. formosanum Hayata	KC2	○	○	○	○	○	○	○
C. goeringii s.s. (Rchb.f.) Rchb.f.	TBG115919	○	◎	○	○	◎	○	◎

C. goeringii s.l.	KC26	○	○	○	○	○	○	○
C. haematodes Lindl.	TBG133251	○	◎	○	○	◎	○	○
C. hartinahianum J.B.Comber & Nasution	KCT08	○	◎	○	○	◎	○	◎
C. hookerianum Rchb.f.	TBG133817	○	◎	○	○	◎	○	◎
C. insigne Rolfe	TBG135526	○	◎	○	○	◎	○	◎
C. iridioides D.Don	KCT21	○	◎	○	○	◎	○	◎
C. kanran Makino	TBG133782	○	◎	○	○	◎	○	◎
C. koran Makino	KC34	○	○	○	○	○	○	○
C. lancifolium Hook.	TBG56111	○	◎	○	○	◎	○	◎
C. lowianum (Rchb.f.) Rchb.f.	TBG54943	○	◎	○	○	◎	○	◎
C. lowianum var. *iansonii* (Rolfe) P.J.Cribb & Du Puy	KCT17	○	◎	○	○	◎	○	◎
C. macrorhizon Lindl.	KCT06	○	◎	○	○	◎	○	◎
C. madidum Lindl.	TBG119076	○	◎	○	○	◎	○	◎
C. maguanense F.Y.Liu	KC28	○	○	○	○	○	○	○
C. mastersii Griff. ex Lindl.	KC16	○	○	○	○	○	○	○
C. munronianum King & Pantl.	KCT16	○	◎	○	○	◎	○	◎
C. nagifolium Masam.	TBG122388	○	◎	○	○	◎	○	◎
C. nanulum Y.S.Wu & S.C.Chen	KC43	○	◎	○	○	◎	○	○
C. nipponicum (Franch. & Sav.) Rolfe	KCT15	○	◎	○	○	◎	○	◎
C. qiubeiense K.M.Feng & H.Li	TBG133821	○	◎	○	○	○	○	○

C. rectum Lindl.	KCT10	○	◎	○	○	◎	○	◎
C. roseum J.J.Sm.	TBG133724	○	◎	○	○	◎	○	◎
C. sanderae (Rolfe) P.J.Cribb & Du Puy	TBG133251	○	◎	○	○	◎	○	◎
C. schroederi Rolfe	KC49	○	○	○	○	○	○	○
C. serratum Schltr.	KC25	○	○	○	○	○	○	○
C. sigmoideum J.J.Sm.	KCT14	○	◎	○	○	◎	○	◎
C. sinense (Andrews) Willd.	KCR9	-	-	-	-	◎	-	◎
C. sinense	KC21	○	○	○	○	-	○	-
C. suave R.Br.	KCT18	○	◎	○	○	◎	○	◎
C. suavissimum Sander ex C.H.Curtis	TBG128817	-	-	-	-	◎	-	◎
C. suavissimum	KCT19	○	○	○	○	-	○	-
C. tigrinum C.S.P.Parish ex Hook.	KCR11	-	-	-	-	◎	-	◎
C. tigrinum	KC23	○	○	○	○	-	○	-
C. tortisepalum Fukuy.	KC7	○	○	○	○	○	○	○
C. tracyanum L.Castle	TBG144175	○	◎	○	○	◎	○	◎
C. wadae T.Yukawa	KCR13 (HT)	-	-	-	-	◎	-	◎
C. wadae	TBG127439	○	○	○	○	-	○	-
C. wenshanense Y.S.Wu & F.Y.Liu	KC5	○	◎	○	○	◎	○	○
C. whiteae King & Pantl.	KC15	-	-	-	-	◎	-	◎
C. whiteae	TBG140251	○	○	○	○	-	○	-

Outgroups								
Acriopsis liliifolia (J.Koenig) Ormerod	KCN01	○	◎	○	○	◎	○	◎
Grammatophyllum speciosum Blume	KCT05	○	○	○	○	◎	○	◎
Thecostele alata (Roxb.) C.S.P.Parish & Rchb.f.	TBG118540	○	◎	○	○	○	○	◎

Table 1-2. Primers utilised in this study. Primers with "*" after their references were those modified from the original study to match the sequence of the genus *Cymbidium*. "-" in the Annealing and Extension section were the primers used only for the sequencing.

Target region	Primers		Sequences	Annealing, Extension	References
Xdh	X502F	Forward	TGTGATGTCGATGTATGC	53° C, 1 min	Gorniak *et al.*,2010
	Xdh-1513R_2	Reverse	GAGCAATATCATCTTCTCTCCG		Li *et al.*, 2016 *
	X551F	Forward	GAAGAGCAGATTGAAGAATGCC	53° C, 1 min	Gorniak *et al.*,2010
	Xdh-849R_2	Reverse	CAAACTGCCTGCTCCATGGTAA		Tsutsumi pers. com. *
nrITS	17SE_2	Forward	TCATGGTCCGGTGAAGTG	50° C, 1 min	Sun *et al.*, 1994 *
	845R	Reverse	TTGATATGCTTAAACTCGGCG		This study
rbcL	1F_2	Forward	ATGTCACCACAAACAGAA	46° C, 1 min	Goldman *et al.*, 2001 *
	rbcL1345R	Reverse	GCTCCATTTGCTAGCTTC		This study
	505F	Internal	CTATTGGGATGTACTATTAAACCA	-	This study
	891R	Internal	TCTGTCTATCAATAACTGCATG	-	This study
psaB	NY159F_2	Forward	ACTCGTCGTATTTGGTTTG		Cameron, 2004 *

	NY160R_2	Reverse	AAAGTAACCCATCCAATG	52° C, 1 min	Cameron, 2004 *
	IntF	Internal	TAGTTCATGTCGCTATTCC	-	This study
	IntR	Internal	GGAAAAATTGGTTAAAAGTTT	-	This study
matK	OMAT1F	Forward	CAATATGGTCAGAACGGCGT	49° C, 2 min	Hidayat *et al.*, 2005
	trnK-2R	Reverse	AACTAGTCGGATGGAGTAG		Hidayat *et al.*, 2005
	577F	Internal	CCATCCTATCCATTTGGA	-	This study
	1418R	Internal	TAGCACACGAAAGTCGAAG	-	This study
ycf1	3720F_2	Forward	CGTATGTAATGAACGAATGGAA	53° C, 1 min	Neubig et al., 2009 *
	5500R_2	Reverse	CTGTTATTGGCATCAAACCA		Neubig *et al.*, 2009 *
	ycf1N_F	Forward	CAATATCAAGGGAGCTATCC	49° C, 1 min	This study
	ycf1N_R	Reverse	CCATTATTGGTATCAAACCA		This study
	IntF	Internal	GAGATCTGGACCAATGCACATA	-	This study
	IntR	Internal	CAAGGAATCAGCCCATC	-	This study
	1113R	Internal	GCAAGAACTAGGTCCTTTTGAA	-	This study
rpl16 intron	rpL16F1N_2	Forward	CTCAGTGTGTGACTCGTTAG	49° C, 1 min	Inoue & Yukawa, 2002 *
	rpL16R2N	Reverse	TGCTTCTATTTGTCTAGCTG		Inoue & Yukawa, 2002
	F466	Internal	AAAGGCAGTGTTATAAAGCAT	-	This study

R1021	Internal	TCCTTCATTTGTAGGGTCA	-	This study

Table 1-3. A list of characters of each species examined in this study. "T" indicates a terrestrial habitat and "E" indicates an epiphytic habitat. "0" indicates that the species have no subterranean rhizome but have normal leaves, "1" indicates that the species have subterranean rhizome and leaves, and "2" indicates the species with subterranean rhizome but have no normal leaves.

Taxa	Habitat	rhizome
Cymbidium aliciae.	T	0
C. aloifolium.	E	0
C. aspidistrifolium	T	1
C. atropurpureum	E	0
C. banaense	E	0
C. bicolor subsp. *pubescens*	E	0
C. borneense	T	0
C. canaliculatum	E	0
C. chloranthum	E	0
C. cochlear	E	0
C. crassifolium	E	0
C. cyperifolium subsp. *cyperifolium*	T	1
C. cyperifolium subsp. *indochinense*	T	1
C. dayanum	E	0
C. devonianum	E	0
C. eburneum	E	0

C. elegans	E	0
C. elongatum	T	0
C. ensifolium	T	1
C. erythraeum	E	0
C. erythrostylum	E	0
C. faberi	T	1
C. finlaysonianum	E	0
C. floribundum	E	0
C. formosanum	T	1
C. goeringii s.s.	T	1
C. goeringii s.l.	T	1
C. haematodes	T	1
C. hartinahianum	T	0
C. hookerianum	E	0
C. insigne	T	0
C. iridioides	E	0
C. kanran	T	1
C. koran	T	1
C. lancifolium	T	1
C. lowianum	E	0
C. lowianum var. *iansonii*	E	0
C. macrorhizon	T	2
C. madidum	E	0
C. maguanense	E	0
C. mastersii	E	0
C. munronianum	T	1
C. nagifolium	T	1
C. nanulum	T	1
C. nipponicum	T	2
C. qiubeiense	T	1
C. rectum	E	0
C. roseum	E	0

C. sanderae	E	0
C. schroederi	E	0
C. serratum	T	1
C. sigmoideum	E	0
C. sinense	T	1
C. suave	E	0
C. suavissimum	E	0
C. tigrinum	E	0
C. tortisepalum	T	1
C. tracyanum	E	0
C. wadae	E	0
C. wenshanense	E	0
C. whiteae	E	0
Outgroups		
Acriopsis liliifolia	E	0
Grammatophyllum speciosum	E	0
Thecostele alata	E	0

Table 1-4. Properties of DNA datasets used in this study.

locus	No. of taxa	Length range (bp)	Alignment length (bp)	Variable characters (bp)	Parsimony informative characters (bp)	Excluded ambiguously aligned characters (bp)
Xdh	60	752	752	123 (16%)	44 (5.9%)	0
nrITS	64	703–732	761	208 (27%)	121 (16%)	0
combined nrDNA	64	-	1513	331 (22%)	165 (11%)	-
rbcL	64	1141	1141	71 (6.2%)	29 (2.5%)	0
psaB	64	1582	1582	108 (6.8%)	46 (2.9%)	0
matK	64	1541–1547	1550	291 (19%)	117 (7.5%)	0
ycf1	64	1416–1485	1566	406 (26%)	163 (10%)	0
rpl16 intron	64	1004–1399	1734	538 (31%)	271 (16%)	891

combined cpDNA	64	-	6682	1040 (16%)	431 (6.5%)	-
combined cp & nrDNA	64	-	8195	1147 (14%)	461 (5.6%)	-

Figure legends

Figure 1-1. ML best-score tree inferred from five plastid loci (*rbcL*, *psaB*, *matK*, *ycf1*, and *rpl16* intron). Bootstrap percentages (BP) and Bayesian posterior probabilities (PP) are displayed on the branches in this order ("*" indicates a support value of 100 and "-" indicates a support value less than 50). The sections with "*" are non-monophyletic sections. Tips of the non-monophyletic sections are coloured red (section *Jensoa*) and blue (section *Floribundum*), respectively.

Figure 1-2. ML best-score tree inferred from two nuclear loci (*Xdh* and nrITS). Bootstrap percentages (BP) and Bayesian posterior probabilities (PP) are displayed on the branches in this order ("*" indicates a support value of 100 and "-" indicates a support value less than 50). The sections/clades with "*" are non-monophyletic sections/clades. Tips of the non-monophyletic sections are coloured green (section *Cyperorchis*) and blue (section *Floribundum*), respectively.

Figure 1-3. ML best-score tree inferred from combined plastid and nuclear data with five plastid and two nuclear loci (*rbcL*, *psaB*, *matK*, *ycf1*, *rpl16* intron, *Xdh*, and nrITS). Bootstrap percentages (BP) and Bayesian posterior probabilities (PP) are displayed on the branches in this order ("*" indicates a support value of 100 and "-" indicates a support value less than 50). The sections with "*" are non-monophyletic sections.

Figure 1-4. Summaries of the ancestral state reconstruction of two states from 1,000 simulations using stochastic mapping. Ancestral state reconstructions of habitats (terrestrial or epiphyte) (plastid genes: A; nuclear genes: C) and rhizome formation and nutrient modes (rhizome present and leaf absent, rhizome present and leaf present, or rhizome absent and leaf present) (plastid genes: B; nuclear genes: D).

Figure 1-5. Infrataxonomic recognition of genus *Cymbidium* in previous studies and the present study. The lines indicate merges and splits of the sections among the authors. Species classified as "?" in the present study possibly need their own section. *Cymbidium repens* was described in 2016, thus former studies do not treat the taxonomic position of it.

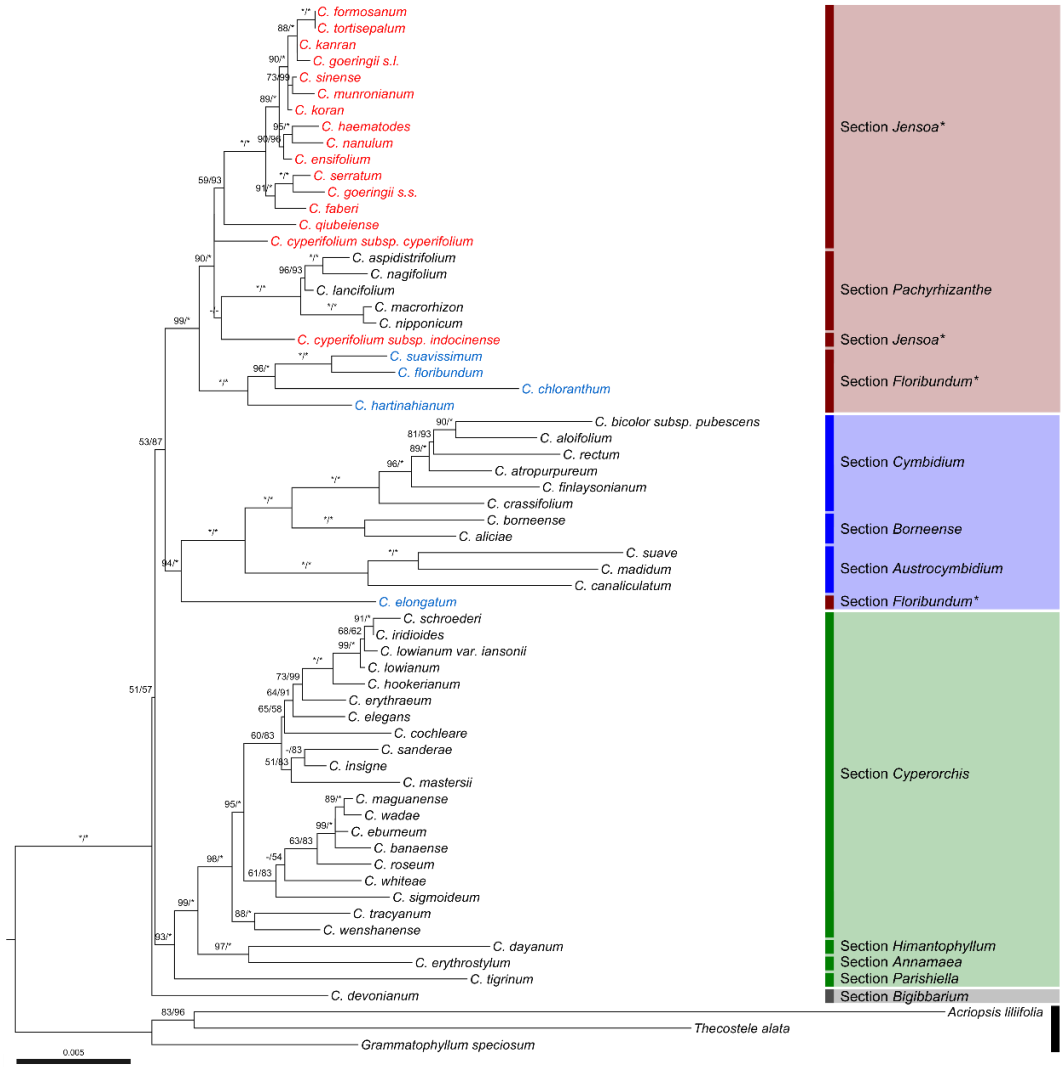

C. formosanum
C. tortisepalum
C. kanran
C. goeringii s.l.
C. sinense
C. munronianum
C. koran
C. haematodes
C. nanulum
C. ensifolium
C. serratum
C. goeringii s.s.
C. faberi
C. qiubeiense
C. cyperifolium subsp. cyperifolium
C. aspidistrifolium
C. nagifolium
C. lancifolium
C. macrorhizon
C. nipponicum
C. cyperifolium subsp. indocinense
C. suavissimum
C. floribundum
C. chloranthum
C. hartinahianum
C. bicolor subsp. pubescens
C. aloifolium
C. rectum
C. atropurpureum
C. finlaysonianum
C. crassifolium
C. borneense
C. aliciae
C. suave
C. madidum
C. canaliculatum
C. elongatum
C. schroederi
C. iridioides
C. lowianum var. iansonii
C. lowianum
C. hookerianum
C. erythraeum
C. elegans
C. cochleare
C. sanderae
C. insigne
C. mastersii
C. maguanense
C. wadae
C. eburneum
C. banaense
C. roseum
C. whiteae
C. sigmoideum
C. tracyanum
C. wenshanense
C. dayanum
C. erythrostylum
C. tigrinum
C. devonianum
Acriopsis liliifolia
Thecostele alata
Grammatophyllum speciosum
0.005
Section Jensoa*
Section Pachyrhizanthe
Section Jensoa*
Section Floribundum*
Section Cymbidium
Section Borneense
Section Austrocymbidium
Section Floribundum*
Section Cyperorchis
Section Himantophyllum
Section Annamaea
Section Parishiella
Section Bigibbarium
Jensoa Clade
Cymbidium Clade
Cyperorchis Clade
Bigibbarium Clade
Outgroups

Figure 1-1.

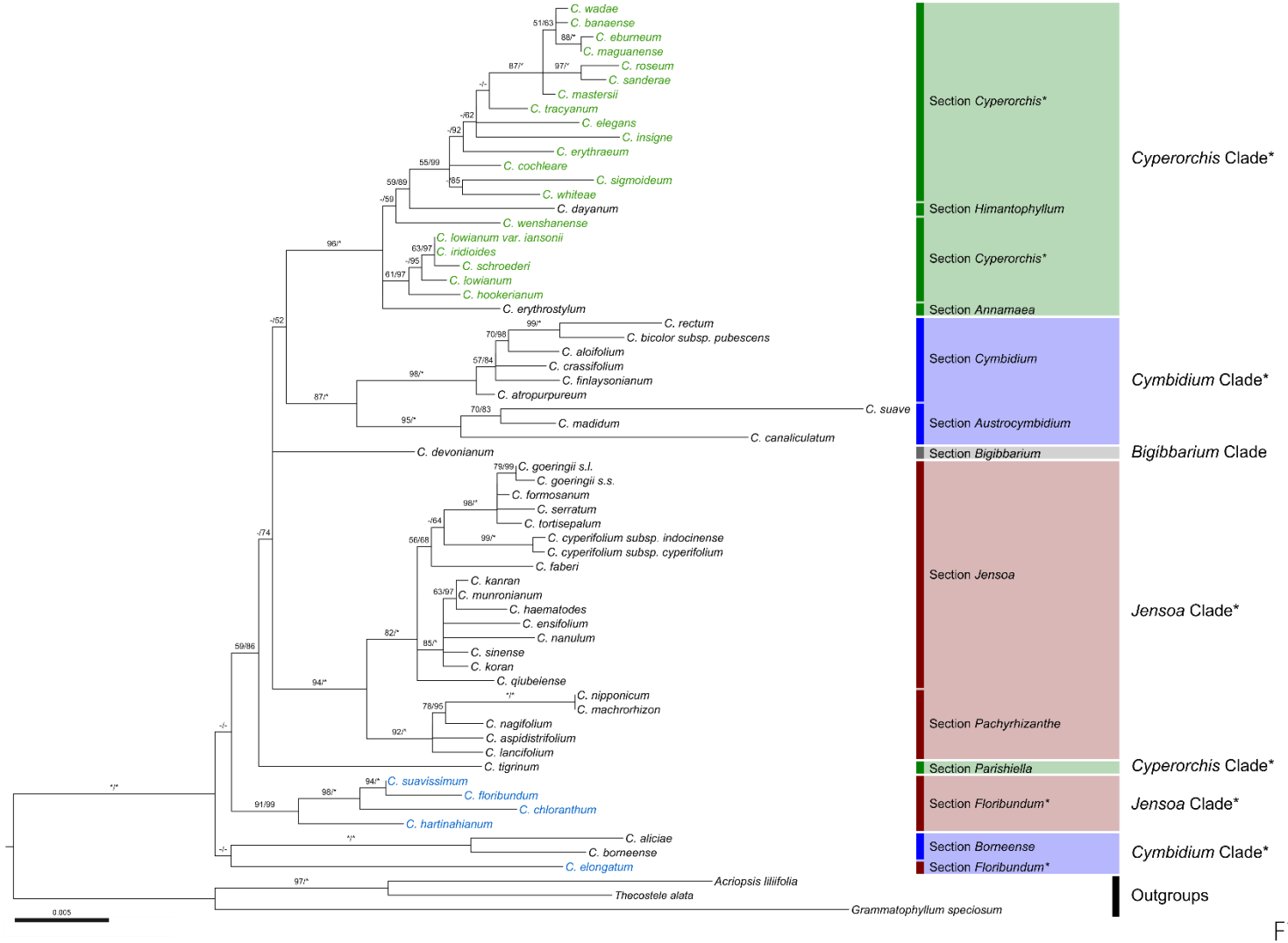

C. wadae
C. banaense
C. eburneum
C. maguanense
C. roseum
C. sanderae
C. mastersii
C. tracyanum
C. elegans
C. insigne
C. erythraeum
C. cochleare
C. sigmoideum
C. whiteae
C. dayanum
C. wenshanense
C. lowianum var. iansonii
C. iridioides
C. schroederi
C. lowianum
C. hookerianum
C. erythrostylum
C. rectum
C. bicolor subsp. pubescens
C. aloifolium
C. crassifolium
C. finlaysonianum
C. atropurpureum
C. suave
C. madidum
C. canaliculatum
C. devonianum
C. goeringii s.l.
C. goeringii s.s.
C. formosanum
C. serratum
C. tortisepalum
C. cyperifolium subsp. indocinense
C. cyperifolium subsp. cyperifolium
C. faberi
C. kanran
C. munronianum
C. haematodes
C. ensifolium
C. nanulum
C. sinense
C. koran
C. qiubeiense
C. nipponicum
C. machrorhizon
C. nagifolium
C. aspidistrifolium
C. lancifolium
C. tigrinum
C. suavissimum
C. floribundum
C. chloranthum
C. hartinahianum
C. aliciae
C. borneense
C. elongatum
Acriopsis liliifolia
Thecostele alata
Grammatophyllum speciosum
0.005
Section Cyperorchis*
Section Himantophyllum
Section Cyperorchis*
Section Annamaea
Cyperorchis Clade*
Section Cymbidium
Section Austrocymbidium
Cymbidium Clade*
Section Bigibbarium
Bigibbarium Clade
Section Jensoa
Jensoa Clade*
Section Pachyrhizanthe
Section Parishiella
Cyperorchis Clade*
Section Floribundum*
Jensoa Clade*
Section Borneense
Section Floribundum*
Cymbidium Clade*
Outgroups

Figure 1-2.

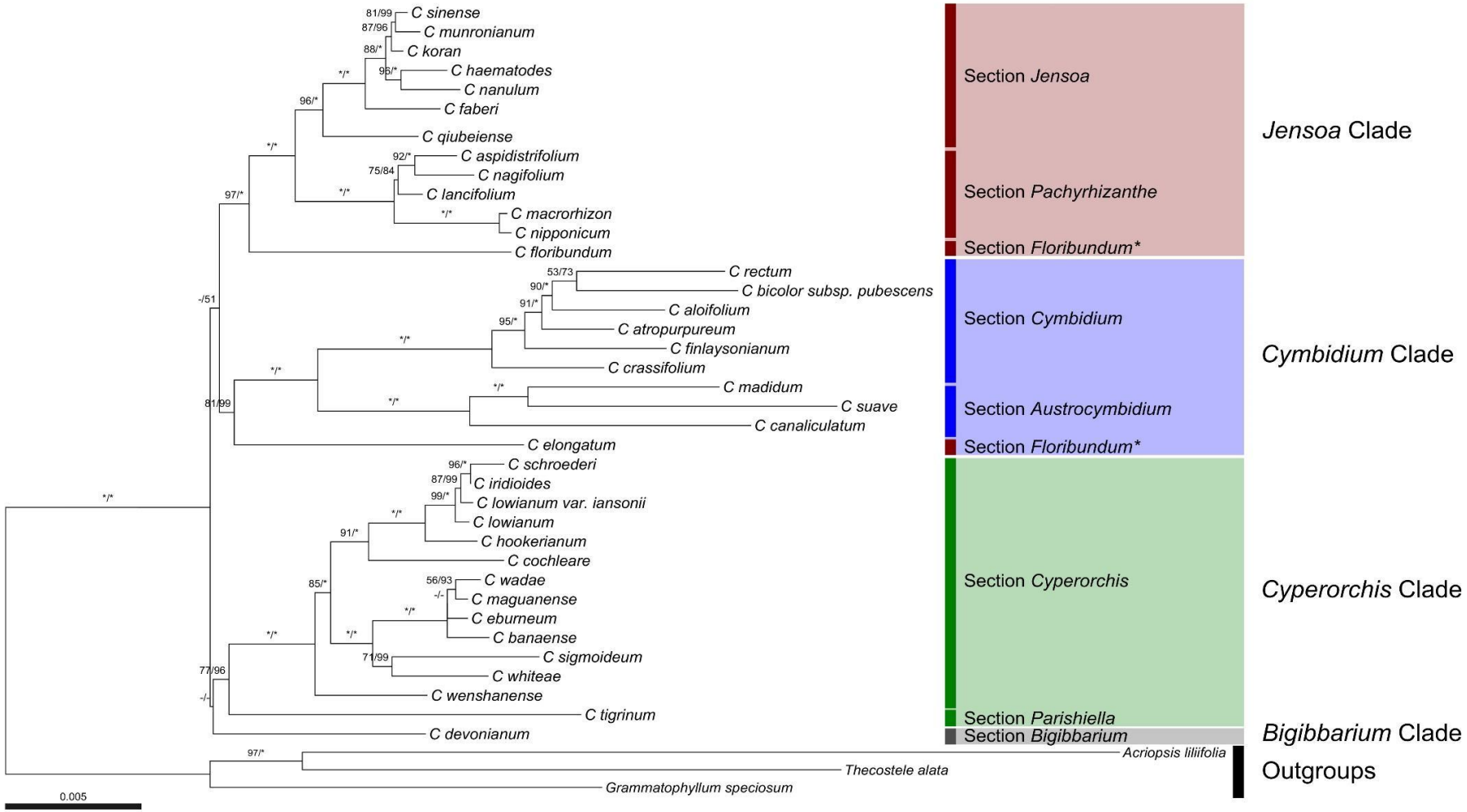
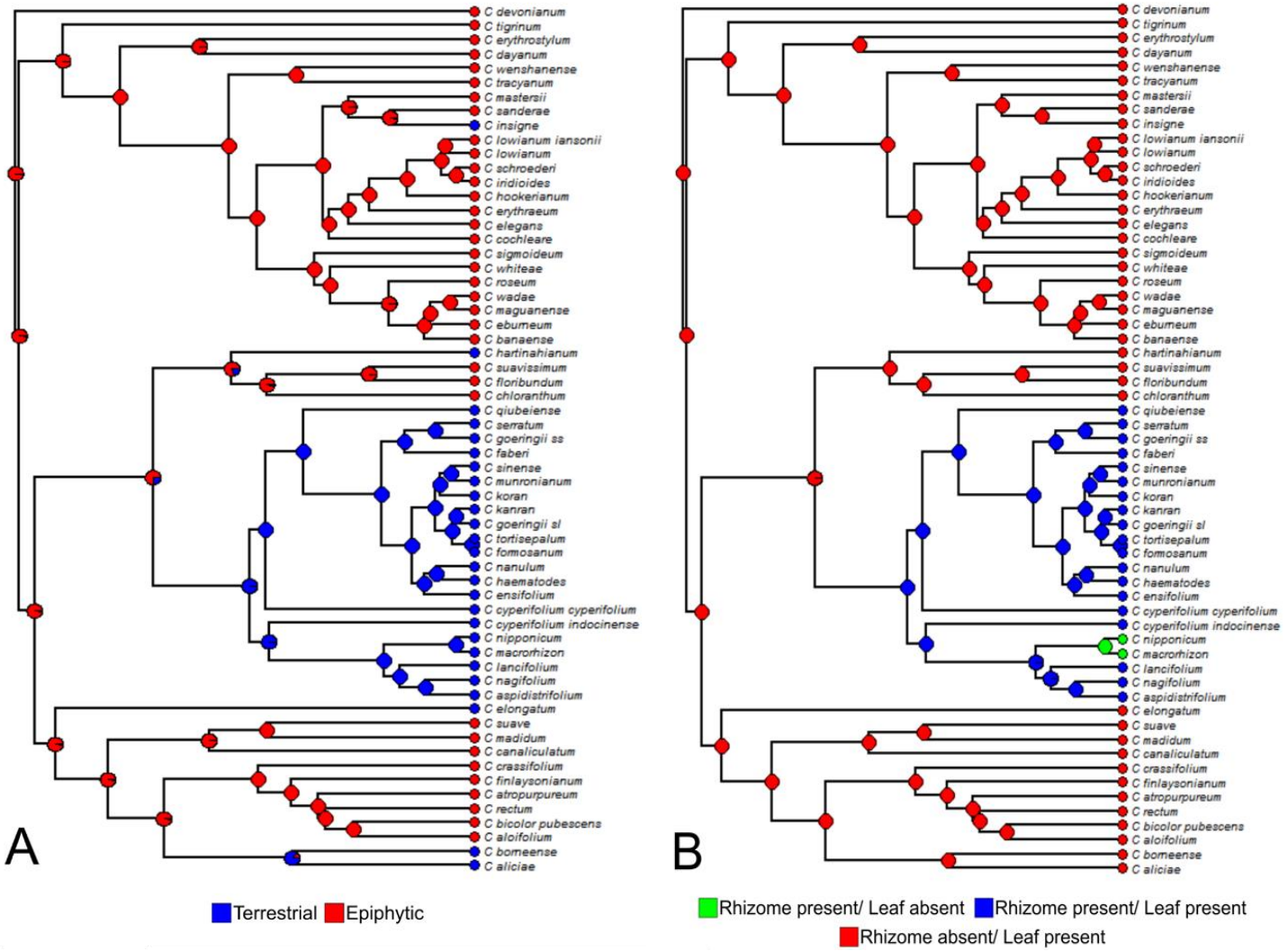

C sinense
C munronianum
C koran
C haematodes
C nanulum
C faberi
C qiubeiense
C aspidistrifolium
C nagifolium
C lancifolium
C macrorhizon
C nipponicum
C floribundum
C rectum
C bicolor subsp. pubescens
C aloifolium
C atropurpureum
C finlaysonianum
C crassifolium
C madidum
C suave
C canaliculatum
C elongatum
C schroederi
C iridioides
C lowianum var. iansonii
C lowianum
C hookerianum
C cochleare
C wadae
C maguanense
C eburneum
C banaense
C sigmoideum
C whiteae
C wenshanense
C tigrinum
C devonianum
Acriopsis liliifolia
Thecostele alata
Grammatophyllum speciosum
0.005
Section Jensoa
Section Pachyrhizanthe
Section Floribundum*
Section Cymbidium
Section Austrocymbidium
Section Floribundum*
Section Cyperorchis
Section Parishiella
Section Bigibbarium
Jensoa Clade
Cymbidium Clade
Cyperorchis Clade
Bigibbarium Clade
Outgroups

Figure 1-3.


C devonianum
C tigrinum
C erythrostylum
C dayanum
C wenshanense
C tracyanum
C mastersii
C sanderae
C insigne
C lowianum iansonii
C lowianum
C schroederi
C iridioides
C hookerianum
C erythraeum
C elegans
C cochleare
C sigmoideum
C whiteae
C roseum
C wadae
C maguanense
C eburneum
C banaense
C hartinahianum
C suavissimum
C floribundum
C chloranthum
C qiubeiense
C serratum
C goeringii ss
C faberi
C sinense
C munronianum
C koran
C kanran
C goeringii sl
C tortisepalum
C formosanum
C nanulum
C haematodes
C ensifolium
C cyperifolium cyperifolium
C cyperifolium indocinense
C nipponicum
C macrorhizon
C lancifolium
C nagifolium
C aspidistrifolium
C elongatum
C suave
C madidum
C canaliculatum
C crassifolium
C finlaysonianum
C atropurpureum
C rectum
C bicolor pubescens
C aloifolium
C borneense
C aliciae
A
Terrestrial
Epiphytic
B
Rhizome present/ Leaf absent
Rhizome present/ Leaf present
Rhizome absent/ Leaf present

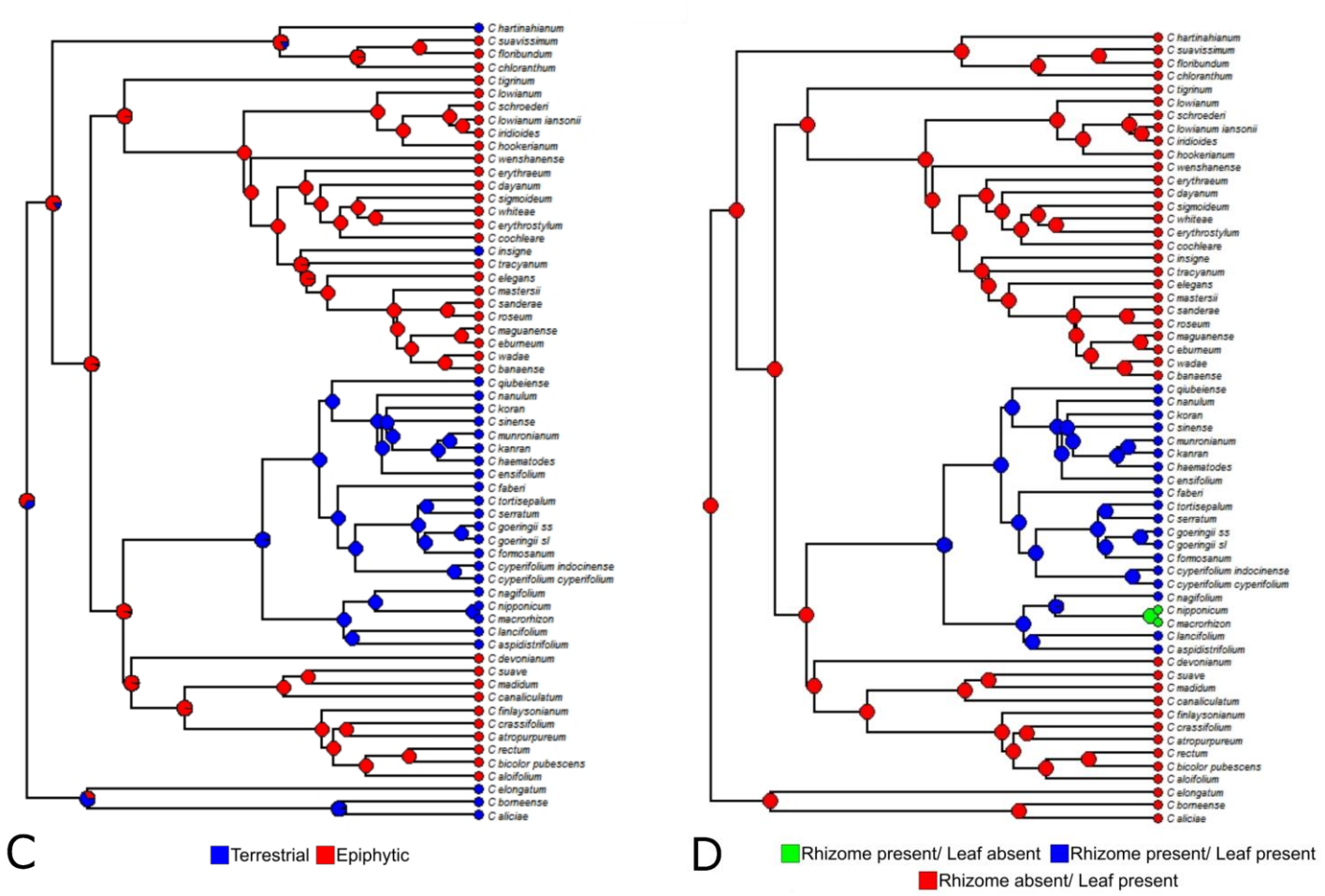

C hartinahianum
C suavissimum
C floribundum
C chloranthum
C tigrinum
C lowianum
C schroederi
C lowianum iansonii
C iridioides
C hookerianum
C wenshanense
C erythraeum
C dayanum
C sigmoideum
C whiteae
C erythrostylum
C cochleare
C insigne
C tracyanum
C elegans
C mastersii
C sanderae
C roseum
C maguanense
C eburneum
C wadae
C banaense
C qiubeiense
C nanulum
C koran
C sinense
C munronianum
C kanran
C haematodes
C ensifolium
C faberi
C tortisepalum
C serratum
C goeringii ss
C goeringii sl
C formosanum
C cyperifolium indocinense
C cyperifolium cyperifolium
C nagifolium
C nipponicum
C macrorhizon
C lancifolium
C aspidistrifolium
C devonianum
C suave
C madidum
C canaliculatum
C finlaysonianum
C crassifolium
C atropurpureum
C rectum
C bicolor pubescens
C aloifolium
C elongatum
C borneense
C aliciae
C
Terrestrial
Epiphytic
D
Rhizome present/ Leaf absent
Rhizome present/ Leaf present
Rhizome absent/ Leaf present

Figure 1-4.

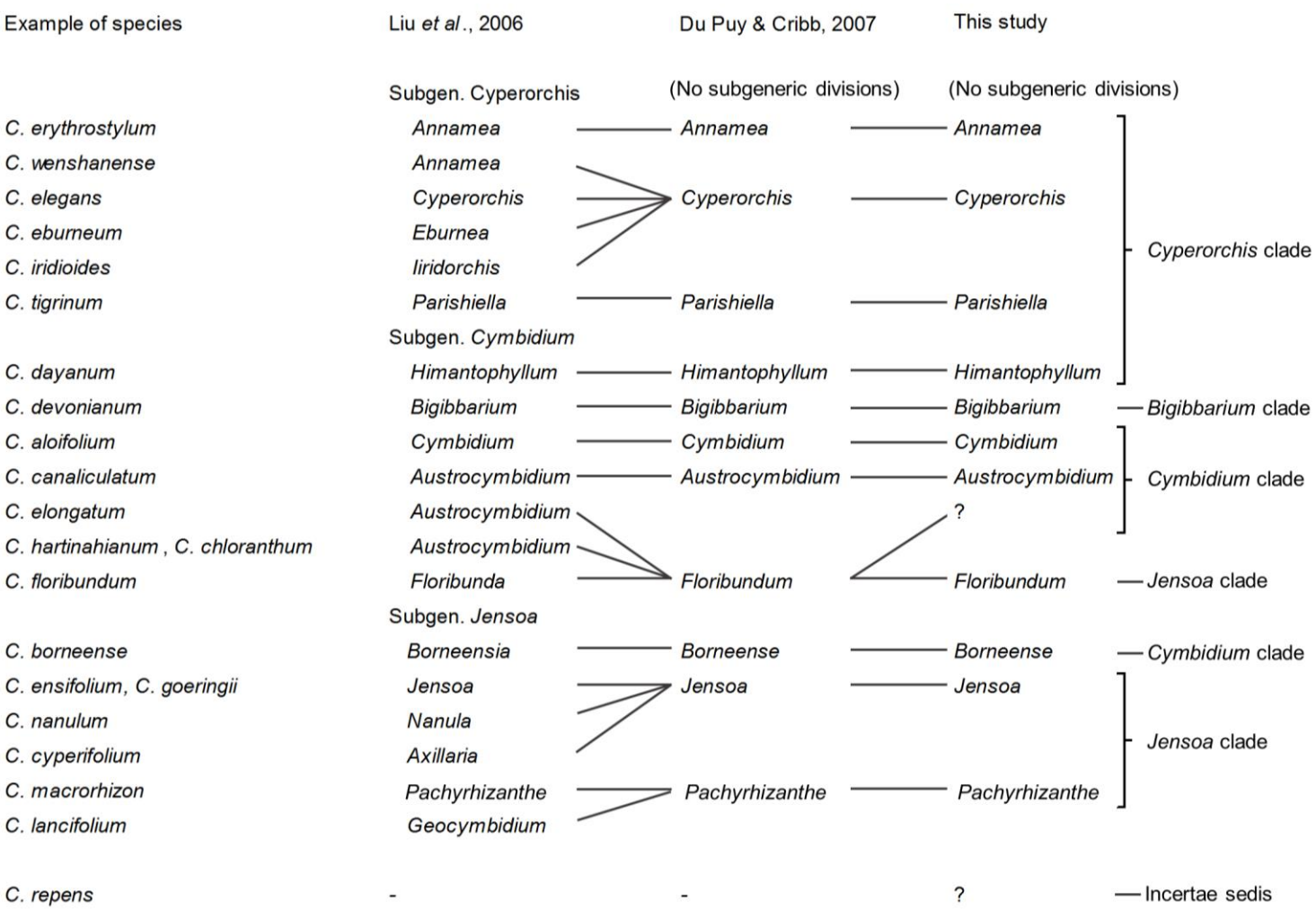

Example of species
Liu et al., 2006
Du Puy & Cribb, 2007
This study
Subgen. Cyperorchis
(No subgeneric divisions)
(No subgeneric divisions)
C. erythrostylum
C. wenshanense
C. elegans
C. eburneum
C. iridioides
C. tigrinum
Annamea
Annamea
Cyperorchis
Eburnea
Iiridorchis
Parishiella
Annamea
Cyperorchis
Parishiella
Annamea
Cyperorchis
Parishiella
Cyperorchis clade
Subgen. Cymbidium
C. dayanum
C. devonianum
C. aloifolium
C. canaliculatum
C. elongatum
C. hartinahianum , C. chloranthum
C. floribundum
Himantophyllum
Bigibbarium
Cymbidium
Austrocymbidium
Austrocymbidium
Austrocymbidium
Floribunda
Himantophyllum
Bigibbarium
Cymbidium
Austrocymbidium
Floribundum
Himantophyllum
Bigibbarium
Cymbidium
Austrocymbidium
?
Floribundum
Bigibbarium clade
Cymbidium clade
Jensoa clade
Subgen. Jensoa
C. borneense
C. ensifolium, C. goeringii
C. nanulum
C. cyperifolium
C. macrorhizon
C. lancifolium
Borneensia
Jensoa
Nanula
Axillaria
Pachyrhizanthe
Geocymbidium
Borneense
Jensoa
Pachyrhizanthe
Borneense
Jensoa
Pachyrhizanthe
Cymbidium clade
Jensoa clade
C. repens
-
-
?
Incertae sedis

Figure 1-5.

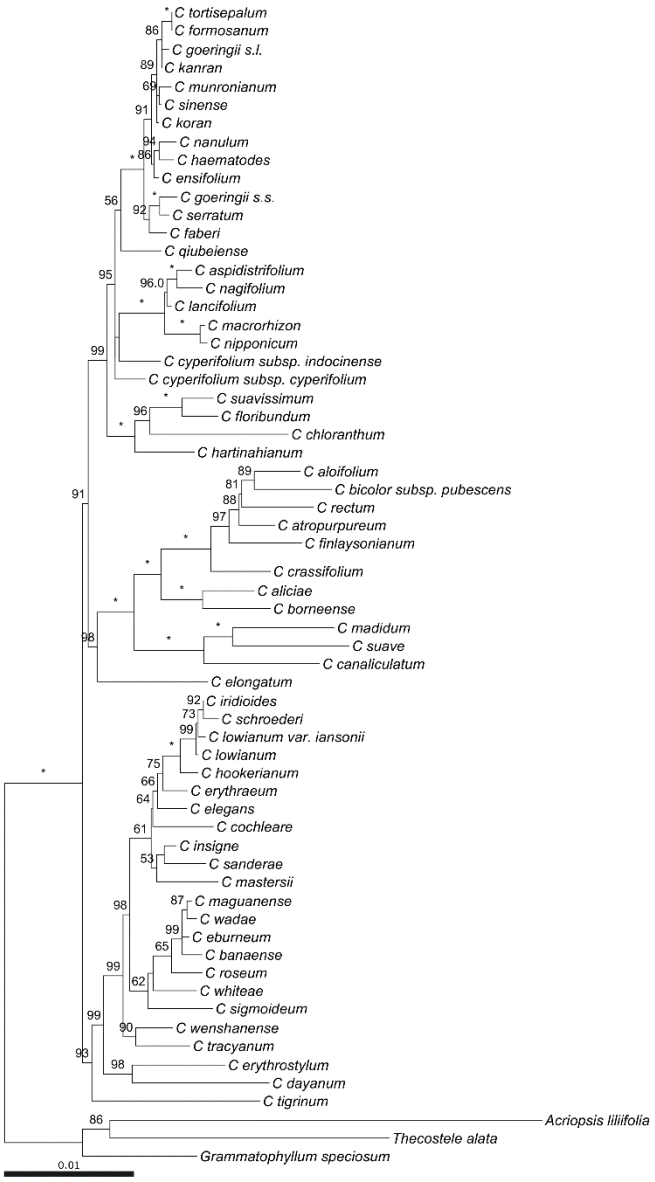
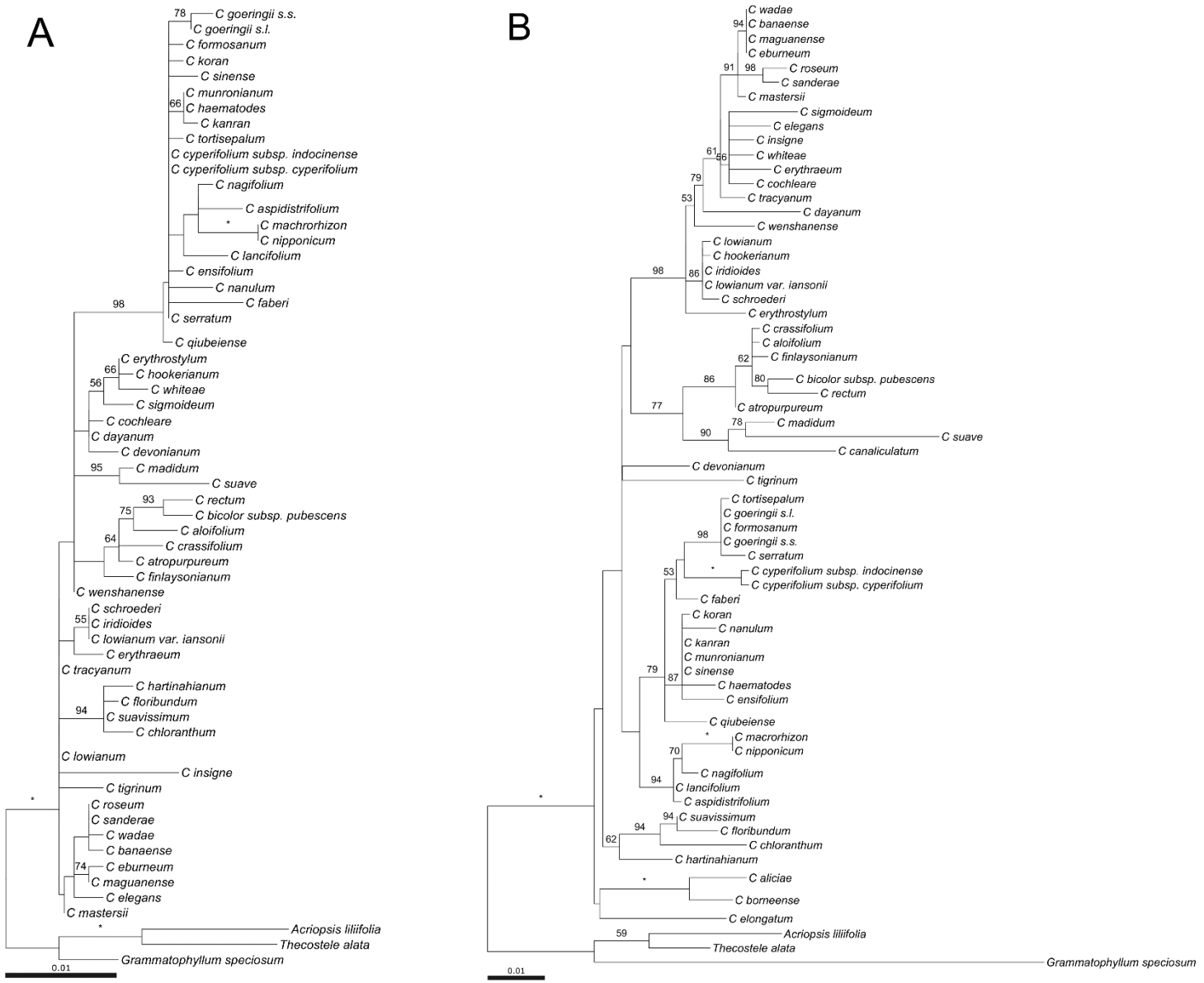
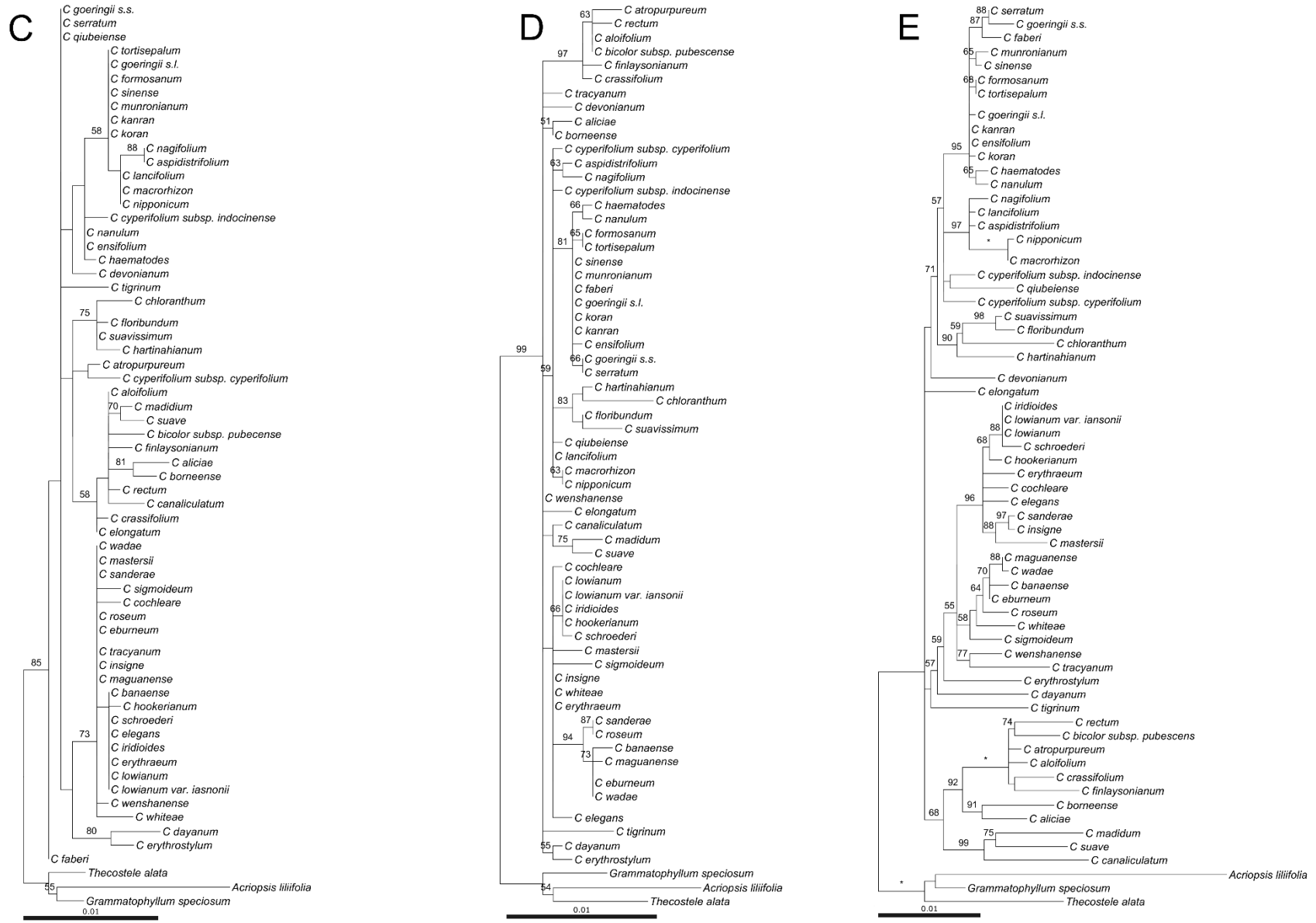


Figure 1-S1. ML best-score tree inferred from five plastid loci (*rbcL*, *psaB*, *matK*, *ycf1*, and *rpl16* intron) excluding *Cymbidium devonianum*. Bootstrap percentages (BP) are displayed on the branches ("*" indicates a support value of 100 and support values less than 50 are not shown). Note that support values of the basal nodes are higher compared with Figure 1-1.


A
C goeringii s.s.
C goeringii s.l.
C formosanum
C koran
C sinense
C munronianum
C haematodes
C kanran
C tortisepalum
C cyperifolium subsp. indocinense
C cyperifolium subsp. cyperifolium
C nagifolium
C aspidistrifolium
C machrorhizon
C nipponicum
C lancifolium
C ensifolium
C nanulum
C faberi
C serratum
C qiubeiense
C erythrostylum
C hookerianum
C whiteae
C sigmoideum
C cochleare
C dayanum
C devonianum
C madidum
C suave
C rectum
C bicolor subsp. pubescens
C aloifolium
C crassifolium
C atropurpureum
C finlaysonianum
C wenshanense
C schroederi
C iridioides
C lowianum var. iansonii
C erythraeum
C tracyanum
C hartinahianum
C floribundum
C suavissimum
C chloranthum
C lowianum
C insigne
C tigrinum
C roseum
C sanderae
C wadae
C banaense
C eburneum
C maguanense
C elegans
C mastersii
Acriopsis liliifolia
Thecostele alata
Grammatophyllum speciosum
0.01
B
C wadae
C banaense
C maguanense
C eburneum
C roseum
C sanderae
C mastersii
C sigmoideum
C elegans
C insigne
C whiteae
C erythraeum
C cochleare
C tracyanum
C dayanum
C wenshanense
C lowianum
C hookerianum
C iridioides
C lowianum var. iansonii
C schroederi
C erythrostylum
C crassifolium
C aloifolium
C finlaysonianum
C bicolor subsp. pubescens
C rectum
C atropurpureum
C madidum
C suave
C canaliculatum
C devonianum
C tigrinum
C tortisepalum
C goeringii s.l.
C formosanum
C goeringii s.s.
C serratum
C cyperifolium subsp. indocinense
C cyperifolium subsp. cyperifolium
C faberi
C koran
C nanulum
C kanran
C munronianum
C sinense
C haematodes
C ensifolium
C qiubeiense
C macrorhizon
C nipponicum
C nagifolium
C lancifolium
C aspidistrifolium
C suavissimum
C floribundum
C chloranthum
C hartinahianum
C aliciae
C borneense
C elongatum
Acriopsis liliifolia
Thecostele alata
Grammatophyllum speciosum
0.01


C
C goeringii s.s.
C serratum
C qiubeiense
C tortisepalum
C goeringii s.l.
C formosanum
C sinense
C munronianum
C kanran
C koran
C nagifolium
C aspidistrifolium
C lancifolium
C macrorhizon
C nipponicum
C cyperifolium subsp. indocinense
C nanulum
C ensifolium
C haematodes
C devonianum
C tigrinum
C chloranthum
C floribundum
C suavissimum
C hartinahianum
C atropurpureum
C cyperifolium subsp. cyperifolium
C aloifolium
C madidium
C suave
C bicolor subsp. pubecense
C finlaysonianum
C aliciae
C borneense
C rectum
C canaliculatum
C crassifolium
C elongatum
C wadae
C mastersii
C sanderae
C sigmoideum
C cochleare
C roseum
C eburneum
C tracyanum
C insigne
C maguanense
C banaense
C hookerianum
C schroederi
C elegans
C iridioides
C erythraeum
C lowianum
C lowianum var. iasnonii
C wenshanense
C whiteae
C dayanum
C erythrostylum
C faberi
Thecostele alata
Acriopsis liliifolia
Grammatophyllum speciosum
0.01
D
C atropurpureum
C rectum
C aloifolium
C bicolor subsp. pubescense
C finlaysonianum
C crassifolium
C tracyanum
C devonianum
C aliciae
C borneense
C cyperifolium subsp. cyperifolium
C aspidistrifolium
C nagifolium
C cyperifolium subsp. indocinense
C haematodes
C nanulum
C formosanum
C tortisepalum
C sinense
C munronianum
C faberi
C goeringii s.l.
C koran
C kanran
C ensifolium
C goeringii s.s.
C serratum
C hartinahianum
C chloranthum
C floribundum
C suavissimum
C qiubeiense
C lancifolium
C macrorhizon
C nipponicum
C wenshanense
C elongatum
C canaliculatum
C madidum
C suave
C cochleare
C lowianum
C lowianum var. iansonii
C iridioides
C hookerianum
C schroederi
C mastersii
C sigmoideum
C insigne
C whiteae
C erythraeum
C sanderae
C roseum
C banaense
C maguanense
C eburneum
C wadae
C elegans
C tigrinum
C dayanum
C erythrostylum
Grammatophyllum speciosum
Acriopsis liliifolia
Thecostele alata
E
C serratum
C goeringii s.s.
C faberi
C munronianum
C sinense
C formosanum
C tortisepalum
C goeringii s.l.
C kanran
C ensifolium
C koran
C haematodes
C nanulum
C nagifolium
C lancifolium
C aspidistrifolium
C nipponicum
C macrorhizon
C cyperifolium subsp. indocinense
C qiubeiense
C cyperifolium subsp. cyperifolium
C suavissimum
C floribundum
C chloranthum
C hartinahianum
C devonianum
C elongatum
C iridioides
C lowianum var. iansonii
C lowianum
C schroederi
C hookerianum
C erythraeum
C cochleare
C elegans
C sanderae
C insigne
C mastersii
C maguanense
C wadae
C banaense
C eburneum
C roseum
C whiteae
C sigmoideum
C wenshanense
C tracyanum
C erythrostylum
C dayanum
C tigrinum
C rectum
C bicolor subsp. pubescens
C atropurpureum
C aloifolium
C crassifolium
C finlaysonianum
C borneense
C aliciae
C madidum
C suave
C canaliculatum
Acriopsis liliifolia
Grammatophyllum speciosum
Thecostele alata
0.01

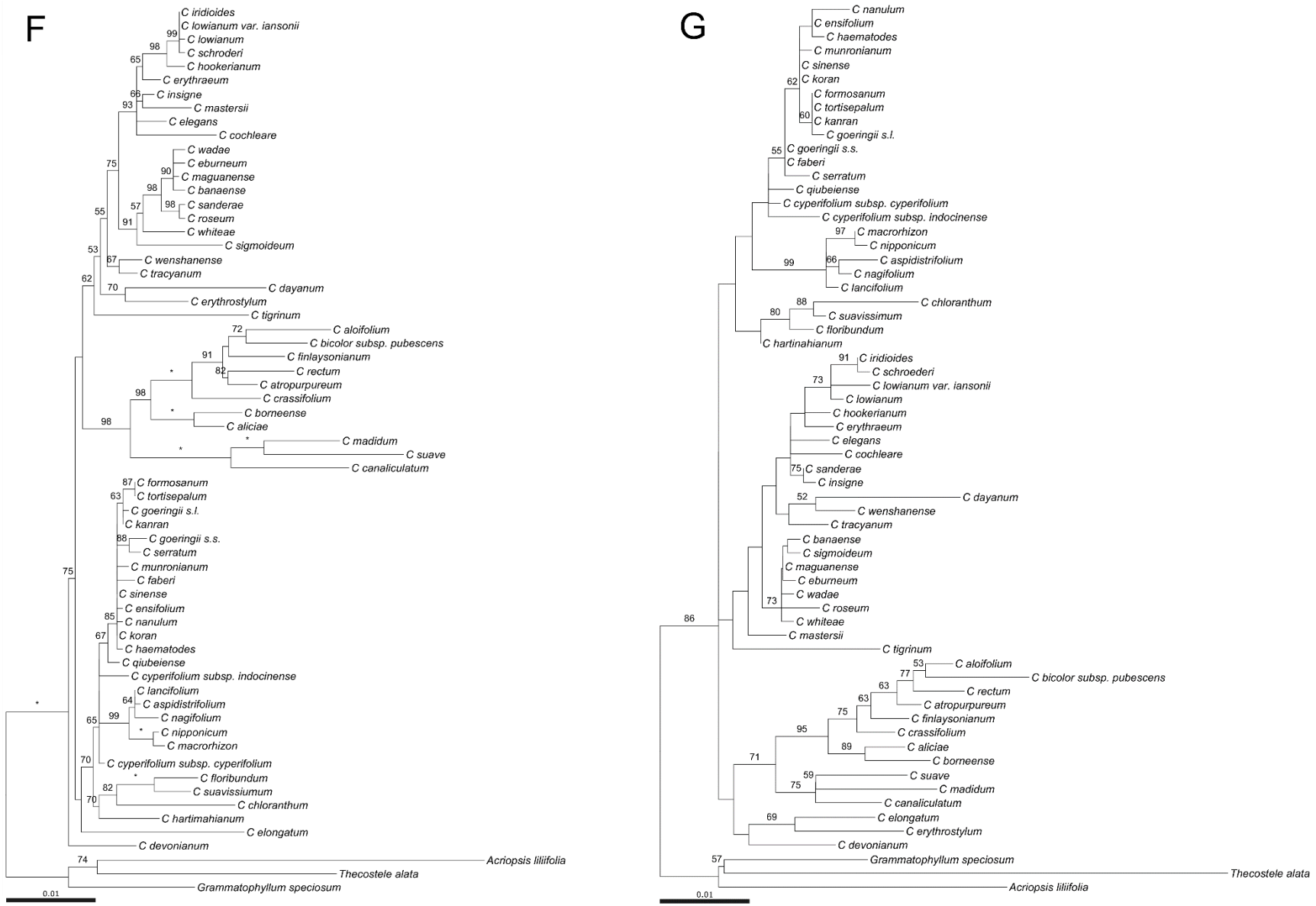

F
C iridioides
C lowianum var. iansonii
C lowianum
C schroderi
C hookerianum
C erythraeum
C insigne
C mastersii
C elegans
C cochleare
C wadae
C eburneum
C maguanense
C banaense
C sanderae
C roseum
C whiteae
C sigmoideum
C wenshanense
C tracyanum
C dayanum
C erythrostylum
C tigrinum
C aloifolium
C bicolor subsp. pubescens
C finlaysonianum
C rectum
C atropurpureum
C crassifolium
C borneense
C aliciae
C madidum
C suave
C canaliculatum
C formosanum
C tortisepalum
C goeringii s.l.
C kanran
C goeringii s.s.
C serratum
C munronianum
C faberi
C sinense
C ensifolium
C nanulum
C koran
C haematodes
C qiubeiense
C cyperifolium subsp. indocinense
C lancifolium
C aspidistrifolium
C nagifolium
C nipponicum
C macrorhizon
C cyperifolium subsp. cyperifolium
C floribundum
C suavissiumum
C chloranthum
C hartimahianum
C elongatum
C devonianum
Acriopsis liliifolia
Thecostele alata
Grammatophyllum speciosum
0.01
G
C nanulum
C ensifolium
C haematodes
C munronianum
C sinense
C koran
C formosanum
C tortisepalum
C kanran
C goeringii s.l.
C goeringii s.s.
C faberi
C serratum
C qiubeiense
C cyperifolium subsp. cyperifolium
C cyperifolium subsp. indocinense
C macrorhizon
C nipponicum
C aspidistrifolium
C nagifolium
C lancifolium
C chloranthum
C suavissimum
C floribundum
C hartinahianum
C iridioides
C schroederi
C lowianum var. iansonii
C lowianum
C hookerianum
C erythraeum
C elegans
C cochleare
C sanderae
C insigne
C dayanum
C wenshanense
C tracyanum
C banaense
C sigmoideum
C maguanense
C eburneum
C wadae
C roseum
C whiteae
C mastersii
C tigrinum
C aloifolium
C bicolor subsp. pubescens
C rectum
C atropurpureum
C finlaysonianum
C crassifolium
C aliciae
C borneense
C suave
C madidum
C canaliculatum
C elongatum
C erythrostylum
C devonianum
Grammatophyllum speciosum
Thecostele alata
Acriopsis liliifolia
0.01

Figure 1-S2. ML best-score tree inferred from each locus. A, *Xdh* B, nrITS C, *rbcL* D, *psaB* E, *matK* F, *ycf1* G, *rpl16* intron, respectively. Bootstrap percentages (BP) are displayed on the branches. ("*" indicates a support value of 100 and support values less than 50 are not shown.)

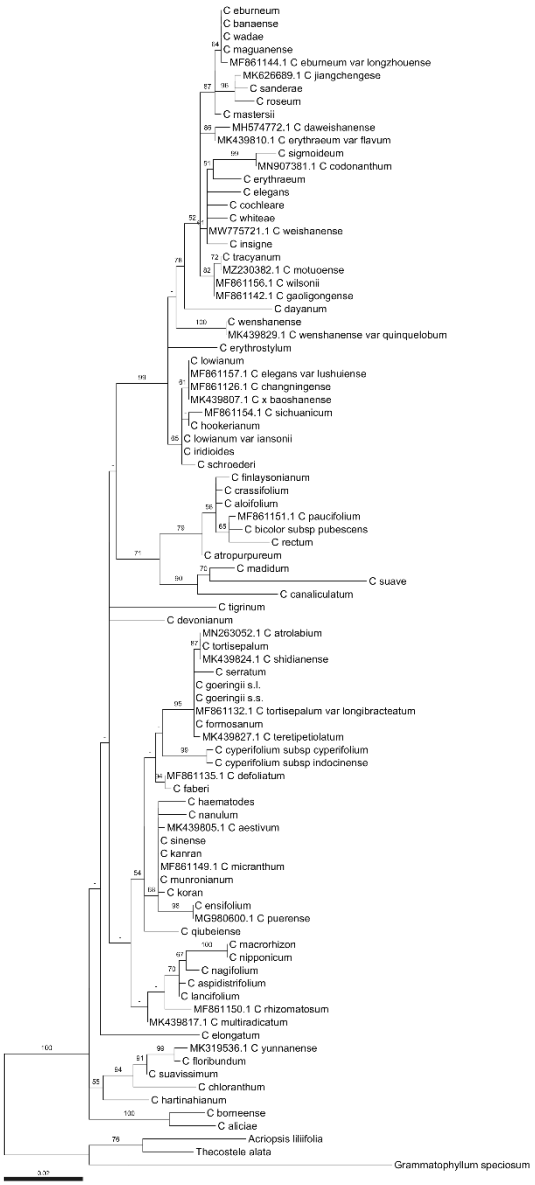


Figure 1-S3. ML best-score tree inferred from nrITS region (650 bp). Note that the data from GenBank are added. Bootstrap percentages (BP) are displayed on the branches. ("*" indicates a support value of 100 and support values less than 50 are not shown.)

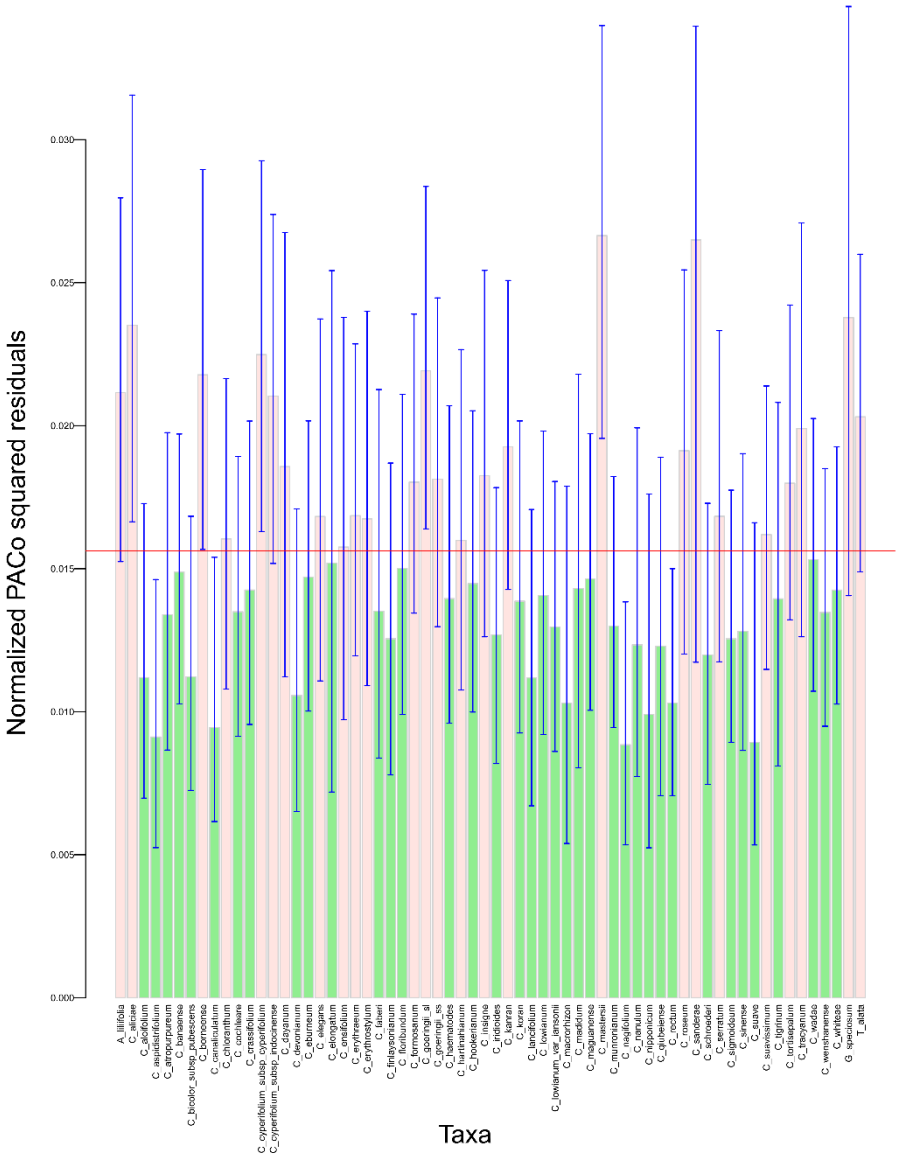


Figure 1-S4. Potential conflicting taxa between nuclear and plastid phylogenies detected by PACo. Terminals with normalised squared residual values higher than the cut-off threshold (red line) are potentially conflicting taxa (pink bar).

Chapter 2: Effect of light conditions on trophic level and gene expression of partially mycoheterotrophic orchid, *Cymbidium goeringii*

Abstract

Partial mycoheterotrophic i.e., mixotrophic, plants are the species which partially depend on mycorrhizal fungi for its nutrients. Although some of these plants are known to show plasticity in the degree of fungal dependence induced by the changes in light condition, the genetic background of this plasticity is largely unsolved. Here, I investigated the relationships between environmental conditions and nutrient sources based on ^{13}C and ^{15}N enrichment in mixotrophic orchid *Cymbidium goeringii.* I also shaded them for two months and evaluated the effect of light condition on the nutrient sources based on the abundance of ^{13}C and ^{15}N and the gene expressions by RNA-seq based de novo assembly. The shading had no effect on isotope enrichment, possibly because of the translocation of carbon and nitrogen from the storage organs. Gene expression analysis showed the upregulation of genes involved in jasmonic acid response in leaves of the shaded plants, which suggests that the jasmonic acid played an important role in regulation of degree of dependence against the mycorrhizal fungi. My results suggest that mixotrophic plants might be controlling their dependency against the mycorrhizal fungi by a common mechanism with the autotrophic plants.

Introduction

Carbon (C) acquisition is one of the fundamental functions in plants. While most plants fixed the atmospheric CO_2 through photosynthesis to fulfil their needs of C, several plants gained C through the mycorrhizal fungi mutualized with them. These plants are called mycoheterotrophic (MH) plants. The levels of dependency against mycorrhizas vary among plant species from fully mycoheterotrophic (FM) to partial mycoheterotrophic (mixotrophic; MX) (Selosse & Roy, 2009; Merckx *et al.*, 2013a). FM plants are observed in a variety of plant lineages from liverworts to Angiosperms for about 880 extant species (Merckx *et al.*, 2013a;b). These FM species frequently lack the chlorophyll and both morphologically and ecologically differ from AT plants in regards of degenerated leaves and roots, adaptation to darker environments, or changes in pollination systems (Leake, 1994; Bidartondo et al., 2004; Suetsugu, 2015).

Although many plants have been found to be MH, the physiological mechanism and the genetic backgrounds of mycoheterotrophy are little known. Several studies have provided insight into the genetic backgrounds of the evolution of FM plants by comparing transcriptomes of AT and FM plants. Transcriptomics of two FM orchids suggested that the shift from autotrophy to mycoheterotrophy is associated with function losses rather than metabolic innovations (Jąkalski et al., 2021). Another study focused on several monocots found convergent gene loss in FM plants across different lineages suggesting shared genetic backgrounds (Timilsena et al., 2022). The albinistic variants of MX plants, which lack chlorophyll, give perspective on

the evolution of FM plants as they gain their carbon needs solely through mycorrhizal fungi, unlike normal (green) MX plants (Julou et al., 2005; Roy et al., 2013; Suetsugu et al., 2017). Several studies comparing green and albinistic MX plants suggest that albinistic individuals show higher expression rate in gene related to nutrients transporting, suggesting those genes are playing important roles in nutrition exchange between plant and its mycorrhizal fungi (Suetsugu et al., 2017; Lallemand et al., 2019). Some of these genes are also involved in mycorrhizal symbiosis in ancestral AT plants suggesting the diversion of the genetic system of AT plants to the MH plants (Suetsugu *et al.*, 2017; Miura *et al.*, 2018).

On the other hand, some MX plants can also plastically change their dependency against mycorrhizal fungi according to their surrounding environments such as light availability (Preiss *et al.*, 2010; Matsuda *et al.*, 2012; Gonneau et al., 2014; Dearnaley et al., 2016). This plasticity may allow MX plants to live in various environments, from light open areas to dark habitats such as forest floor. Motomura *et al.* (2010) suggests that the dependency against mycorrhizal fungi of a MX plant *C. goeringii* varies among habitats, which might reflect the difference of light condition between the habitats. However, the gene expression between these habitats are not compared and the habitats were distant so that the environment apart from light conditions might affect this result. Hence, the genetic mechanisms of adjusting dependency levels against mycorrhizal fungi in response to changes in the environmental conditions are still unclear. The phytohormones such as gibberellins, auxins, strigolactones and jasmonic acids are known to be the key factors controlling the mycorrhizal symbiosis in AT plants (Takeda et al., 2015; Hanlon & Coenen, 2011;

Akiyama & Hayashi, 2006; Hause & Schaarschmidt, 2009). Gibberellins and auxins control mycorrhizal colonisation both positively and negatively through regulations of hyphal entry and arbuscule formation (Takeda et al., 2015; Hedden & Thomas, 2012; Hanlon & Coenen, 2011). Strigolactones promote the arbuscular mycorrhiza formation of *Lotus japonicus* and other plants (Akiyama & Hayashi, 2006) and also initiate the interaction of non-arbuscular mycorrhizal (AM) fungi with a FM orchid (Yuan et al., 2018). Jasmonic acids and its metabolite act throughout the communication between plants and fungi by inducing flavonoids that attract the mycorrhizas and directly affect the signalling cascade that controls the mycorrhizal symbiosis (Hause & Schaarschmidt, 2009). Jasmonic acids are also known to control AM colonisation according to the environmental light conditions (Nagata et al., 2015). Mycorrhizal development is also affected by the environmental factors such as availability of nutrients, water stresses, and light conditions, through the regulation of the phytohormones (Foo et al., 2013; Ruiz - Lozano et al., 2016; Nagata et al., 2015). Some of these phytohormones should also be involved in mycorrhizal symbiosis of MH plants by controlling the degree of mycoheterotrophy (Miura *et al.*, 2018).

Cymbidium is an orchid genus containing the MH species which varies in degree of mycoheterotrophy from AT (MH when juvenile) to almost FM (Motomura et al., 2010; Kobayashi et al., 2021). To name a few examples, *C. dayanum* is an AT plant with saprotrophic-endophytic rhizoctonias, while *C. goeringii* is a MX plant which has associations with both ectomycorrhizal fungi and saprotrophic-endophytic rhizoctonias (Ogura-Tsujita et al., 2012; Selosse et al., 2022). *Cymbidium macrorhizon* is an almost FM plant that has no leaves and photosynthesizes only on its scape and has

associations only with ectomycorrhizal fungi (Ogura-Tsujita et al., 2012; Kobayashi et al., 2021). Therefore, *Cymbidium* can be a good model for testing the evolutionary process of MH plants phylogenetically, morphologically, and physiologically by comparing closely related species with different nutrient modes.

Here, I performed a field sampling and shading experiment on *C. goeringii* to test if the MX species, *C. goeringii,* is able to change their dependency against mycorrhizal fungi in response to the variation in light condition. Generally, the relative concentrations of ^{13}C and ^{15}N are enriched in MH plants by the nutrients derived from their mycorrhizal fungi (Gebauer & Meyer, 2003). Although some studies pointed out that ^{15}N concentration are not directly affected by the light condition (Girlanda et al. 2011; Gonneau et al., 2014; Liebel et al., 2015), I provisionally hypothesised that the shading interrupts the photosynthesis and induces higher dependence on fungi in shaded plants, resulting in higher ^{13}C and ^{15}N concentrations. I further compared the gene expression of shaded and unshaded individuals to evaluate the genes involved in the adjustment of the mycorrhizal symbiosis. Here I address the hypothesis that the mycorrhizal symbiosis in MX plants is, at least in part, regulated by the similar phytohormone-induced system with AT plants.

Material and Methods

Sampling site and shading experiment

A population of *C. goeringii* was sampled at the temperate broadleaf deciduous forest located in the Ecological Garden of the Natural History Museum and Institute, Chiba (Chuo-ku, Chiba-Shi, Chiba-ken, Japan; E35.60°, N140.14°). For the first year of the observation (2018), I sampled 33 individuals of *C. goeringii* on 15th, August and 24th, October. For the second year of the observation (2019), I collected 32 individuals which are the same individuals as the last year (one individual was lost), and the additional 15 individuals on 15th, August and 17th, October. At the same time, I shaded 10 mature individuals of *C. goeringii* out of those 47 individuals by the frames covered with 60% black shade cloth (KLARK Co. Ltd., Japan). Each frame was covered with the cloth except for the bottom. The cloth was adjusted to be about 5 to 10 cm above the top of each plant. The shading was performed from 15th, August, 2019 until 17th, October, 2019 (64 days). The other 36 individuals were left unshaded and used as control.

Measurement of environmental condition

I measured precipitation, temperature, and canopy openness as environmental conditions that might affect the dependency of *C. goeringii* against mycorrhizas. Precipitation and temperature were measured as the macro-environmental factors which might affect all individuals in the site equally, whereas canopy openness was measured as the micro-environmental factors which might affect each individual plant. The precipitation and temperature were measured by the Meteorological equipment of the Ecological Garden. The total precipitation and the average temperature over ten days before the sampling date, respectively, were utilised (both values included the values on the sampling date). The photos of the canopy above each plant individual

were taken by a built-in camera of a smartphone (AQUOS sense lite SH-M05; SHARP, Japan) equipped with a fisheye lens (Daiso, Japan), which were imported to CanopOn 2 (http://takenaka-akio.org/etc/canopon2/) to major canopy openness. To index the maturity of each individual of *C. goeringii*, the number of pseudobulbs was recorded. Because *C. goeringii* has a sympodial growth and makes new shoots (pseudobulbs) once a year, the number of the pseudobulbs reflects the plant age after shoot formation.

Stable isotope analysis

About 1 cm of the leaf from the newest shoot of each individual of *C. goeringii* were collected seven times from 15th August to 17th October for the first year, and the first day (August, 15th) and the last day (October, 17th) of the shading for the second year. For the reference of $\delta^{13}C$ and $\delta^{15}N$ values, leaves (peduncles and rhizomes for the FM plant) of five AT (*Carpinus tschonoskii, Cleyera japonica*, *Machilus thunbergii, Pleioblastus chino, Quercus serrata*, and *Sasa nipponica*), one MX (*Ce. falcata*), and one FM (*C. macrorhizon*) plant species from the surrounding understorey were collected. The leaves were dried in a 60 ° C oven for five days and stored at room temperature until the weighing. Each 1.000 mg – 5.000 mg of the dried leaves were placed in a tin capsule, and $\delta^{13}C$ and $\delta^{15}N$ were measured at the UCDavis Stable Isotope Facility (U.S.A.). The relative abundance of C and N isotope ratios, i.e. $\delta^{13}C$ and $\delta^{15}N$ respectively, were calculated by the equation

$$\delta^{13}C \text{ or } \delta^{15}N = (R_{sample}/R_{standard} - 1) \times 1000[‰]$$

using Vienna Pee Dee Belemnite and atmospheric N_2 as standards for C and N respectively, where R is the molar ratio of the isotopes, hence $^{13}C/^{12}C$ or $^{15}N/^{14}N$. The standard deviations for replicate combustions of the internal standards were 0.05 ‰ for $\delta^{13}C$ and 0.07 ‰ for $\delta^{15}N$ respectively.

Differences in $\delta^{13}C$ or $\delta^{15}N$ among AT, MX, and FM plants were compared by Generalised linear model (GLM) adding nutrient modes as a fixed factor. I considered five models which contain all combinations of one to three groups of nutrient modes with null model. The best model was selected by Akaike's information criterion (AIC) values. Simple regression analyses were used to determine relationships between each environmental factor and isotope ratios. The isotope data of shaded and unshaded individuals were compared with *glmer* function in the R package *lme4* to fit a generalised linear mixed model (GLMM) adding shading and sampling date as fixed factor and plant ID as a random factor (Bates et al., 2015). All statistical analyses were performed in R v4.1.2 (R Core Team, 2021).

RNA extraction and RNA sequencing

About 3 cm of the leaf, 5 mm of the shoot tip (stem), and 1 cm of the matured root from the newest shoot, respectively, of each *C. goeringii* individual was collected and preserved in RNAlater (Invitrogen, U.S.A.) on the first day and the last day of the shading in the second year of the experiment (2019). The specimens were kept at -80 ° C until the total RNA extraction. Total RNA was extracted using Maxwell 16 LEV Plant RNA Kit with the Maxwell 16 Research Instrument (Promega, U.S.A.) according to

the manufacturer's instructions. RNA concentrations were measured using a Qubit 2.0 Fluorometer (Invitrogen, U.S.A.), and electrophoresis on agarose gel was performed to check the RNA degradation. RNA purity was estimated using a BioSpec-nano (Shimadzu, Japan). The cDNA library was constructed using TruSeq RNA Sample Prep Kits. Paired-end (150 bp) RNA sequencing (RNA-seq) was performed on the Illumina NovaSeq6000 platform.

De novo assembly and detection of differentially expressed genes (DEGs)

I utilised Trimmomatic v.0.39 for the removal of adaptor sequences and low-quality reads (Bolger et al., 2014), and used FastQC v.0.11.9 for quality control (http://www.bioinformatics.babraham.ac.uk/projects/fastqc/). The remaining reads were used for de novo assembly using Trinity v.2.13.2 (Grabherr et al., 2011). To estimate gene expression levels, trimmed sequences of each sample were mapped to the reference transcripts using RSEM v.1.3.3 (Li & Dewey, 2011). The read count data was used for gene expression analysis. The isoforms were concatenated by "Trinity_gene_splice_modeler.py" which is a part of the Trinity utilities, then I searched for homologues of every gene using BLAST searches for all protein sequences of *Arabidopsis thaliana* (Araport11). Genes with the best hit and with an e-value < 0.0001 were used for gene expression analysis. DEGs among the two light conditions were detected using the DESeq2 package. I considered genes with $q < 0.05$ as DEGs. Gene ontology (GO) enrichment analysis of the DEGs was performed using the statistical enrichment test of the PANTHER classification system (http://pantherdb.org/).

Results

Stable isotope analysis

The $\delta^{13}C$ and $\delta^{15}N$ values of the *C. goeringii* range widely, often higher than AT and lower than FM plants (Fig. 2-1). The model that distinguishes three nutrient modes was selected based on AIC for both $\delta^{13}C$ and $\delta^{15}N$ values ($AIC_{best} = 642$, $AIC_{null} = 716$ for $\delta^{13}C$, $AIC_{best} = 607$, $AIC_{null} = 856$ for $\delta^{15}N$). The MX orchid species, *Ce. falcata,* showed the similar values with *C. goeringii* in both $\delta^{13}C$ and $\delta^{15}N$. The $\delta^{13}C$ values of unshaded *C. goeringii* ranged from -34.6‰ to -28.3‰ and showed a positive correlation with canopy openness (Fig. 2-2A, $P < 0.001$; Tukey-Kramer test), while other environmental conditions i.e. temperature and precipitation and number of the bulbs have no effects on them (Fig. 2-2C, E, G; $P = 0.92$, $P = 0.57$, $P = 0.53$, respectively). The $\delta^{15}N$ values of unshaded *C. goeringii* ranged from -0.83‰ to -4.93‰. No effects of environmental conditions on the $\delta^{15}N$ were detected (Fig. 2-2B, D, F, H; canopy openness: $P = 0.92$; temperature: $P = 0.88$; precipitation: $P = 0.28$; number of bulbs: $P = 0.10$). The $\delta^{13}C$ values of the shaded individuals ranged from -33.7‰ to -28.3‰, while the $\delta^{15}N$ values ranged from -1.04‰ to -3.43‰(Fig. 2-3). When the $\delta^{13}C$ and $\delta^{15}N$ values of shaded and unshaded individuals were compared, none of the effects of shading were detected by the GLMM with the null model being selected based on AIC values.

Difference in gene expressions under the shaded and unshaded conditions

Eighty six genes were detected as DEGs including 73 upregulated and 13 downregulated ones expressed in the leaves of shaded individuals (Tables 2-1 and 2-S1). Fourteen GO terms were detected as enriched for the "Biological process" category and all of them were over-expressed in the shaded individuals. There was a single GO related to jasmonic acid, the phytohormone related to mycorrhizal symbiosis ("response to jasmonic acid", GO:0009753). Other GOs related to phytohormones, here gibberellins, auxins, and strigolactones, were not detected. One GO term from the "Cellular component" category was detected as enriched and was under-expressed in the shaded individuals (Table 2-2). No GO terms from the "Molecular function" category were enriched. Among the genes expressed in the stems, eight genes were identified as DEGs including three upregulated and five downregulated in shaded individuals (Tables 2-1 and 2-S1). There were no genes related to phytohormones. No GO terms were detected as enriched for the genes expressed in the stems. For the genes expressed at the roots, none of them were detected as DEGs.

Discussion

The effect of the environment against trophic levels of *C. goeringii*

The $\delta^{13}C$ and $\delta^{15}N$ values of *C. goeringii* showed a wider range than AT and FM plants (Fig. 2-1). Motomura et al. (2010) showed that there are considerable differences in

$\delta^{13}C$ and $\delta^{15}N$ values in *C. goeringii* inhabiting different localities, suggesting the ability to adjust the dependency against fungi in different macro-environments. My results showed a further wider range of $\delta^{13}C$ and $\delta^{15}N$ values than those shown in Motomura *et al.* (2010) even among the individuals inhabiting the same locality. This suggests that this species has an ability to adjust the dependency against fungi not only under the different macro-environment but also under the different micro-environments.

I hypothesised that the individuals in darker places had higher ^{13}C and ^{15}N concentrations because of the higher C and N uptake from the mycorrhizal fungi. Although my results showed that canopy openness had a positive effect on $\delta^{13}C$ value and no effect on $\delta^{15}N$ value, which is inconsistent with my hypothesis (Fig. 2A-F). Apart from the mycorrhizal dependency, there are several factors that might affect the ^{13}C and ^{15}N concentrations, which are translocation, water stress, and soil conditions (Farquhar & Richards, 1984; Damesin & Lelarge, 2003; Craine *et al.,* 2015). Generally, the concentrations of ^{13}C in leaves are increased when the water stress is high (Damesin & Lelarge, 2003). The place with higher canopy openness might have higher irradiance which results in higher ground temperature and lower soil water content, which causes higher water stress resulting in higher ^{13}C concentration. The other factor that might affect the ^{13}C and ^{15}N are translocation from the storage organs. Some studies suggested that the existing N pool affects $\delta^{15}N$ values by diluting the newly acquired N from mycorrhizal fungi (Girlanda et al., 2011; Gonneau et al., 2014). *Cymbidium goeringii* is an evergreen perennial with pseudobulbs, thus $\delta^{13}C$ and $\delta^{15}N$ values in the new shoots might be affected by the isotope ratios of C and N stored in

the pseudobulbs, which reflects the degree of mycoheterotrophy in the past few years, not only the present year.

The number of the pseudobulbs, i.e. plant maturity, showed any effect on the isotope ratios, although the individual with the highest $\delta^{13}C$ and $\delta^{15}N$ values was detected in the individual without any pseudobulbs (Fig. 2-2G, H). After the germination, *C. goeringii* spends underground for several years gaining their nutritions from its mycorrhizal fungi, and produces vegetative shoot after that (Du Puy & Cribb, 2007). Hence, the juvenile shoots may be produced mainly from the C and N from the fungi which show higher ^{13}C and ^{15}N concentrations. As the isotope ratios vary between the plants with one or more pseudobulbs, my results also suggest that the plants after producing the vegetative shoots quickly adjust their degree of mycoheterotrophy to the surrounding environments.

The effect of the shading against the stable isotope ratio

Against my expectations, the shading had no effects on $\delta^{13}C$ and $\delta^{15}N$ values of *C. goeringii*. Since *C. goeringii* have evergreen, pseudobulbous shoots which last for several years (Du Puy & Cribb, 2007), C and N stored in the old shoots may flow into the newer shoots. The isotope ratios of this "old" C and N might reflect in the result at the recent months, but years' trends in degree of mycoheterotrophy of the individuals. In addition to the translocation, shading might also cause an impact on $\delta^{13}C$ values via keeping the humidity inside the frame. Higher humidity reduces water stress, resulting in lower $\delta^{13}C$ value (Damesin & Lelarge, 2003), thus the increase in $\delta^{13}C$ values due to

the influence of mycorrhizal dependency may be counteracted by the influence of shading.

Difference in gene expressions among shaded and unshaded individuals

I could identify several DEGs between shaded and unshaded treatments on leaf and stem specimens, but not in root specimens (Table 2-1 and 2-S1). In the leaf tissues of shaded individuals, genes with the GO term "response to jasmonic acid" are over expressed. Jasmonic acid is one of the phytohormones synthesised in leaves (León & Sánchez-Serrano, 1999), and known to be one of the key factors that positively controls mycorrhizal colonisation (Hause & Schaarschmidt, 2009). In tomatoes and legumes, the level of jasmonic acid synthesis was affected by the density of the plant canopy (Smith & Whitelam, 1997), which enabled the control of AM colonisation according to the environmental light conditions (Nagata et al., 2015). Because major components of symbiosis of orchids with mycorrhizal fungi are known to be shared with AT plants with AM fungi (Miura et al., 2018), jasmonic acid in *C. goeringii* might also play a role in mycorrhizal control as for the AT plants. If this is the case, MX orchid may be controlling the dependency against the mycorrhizal fungi through the regulation of jasmonic acid synthesis by synthesising them in a darker environment. Jasmonic acid is also known to induce the defensive reaction to the various biotic and non-biotic stresses (Campos et al., 2014; Mostofa et al., 2017). The multiples of the genes possibly induced by the jasmonic acid response were also enriched in the leaves under the shading, which are those with GO of "defence response", "response to other organism" , "response to fungus", "response to herbivore", and "defence response to bacterium". These reactions

are congruent with the idea that there are conflicts between host plant and its mycorrhizal fungi, while the fungi might be pathogenic if their growth is excessive (Beyrle et al., 1995). However, jasmonic acid response can occur not only in mycorrhizal responses, but also in responses to microorganisms other than mycorrhizal fungi and abiotic stresses (Campos et al., 2014; Mostofa et al., 2017). Thus stresses related to the shading such as pathogens activity promoted by shading could also induce these genes. The GOs related to other phytohormones such as gibberellins, auxins and strigolactones were not found as over or under expressed. The synthesis of these phytohormones in the leaves might be not regulated by the light conditions, because they are often regulated in the root and affect mycorrhizal symbiosis (Hanlon & Coenen, 2011; Akiyama & Hayashi, 2006). Additionally, GOs related to the transport of nutrients, such as sugars and amino acids, do not exhibit significant over or under expression. However, previous research on albinistic variants of MX plants has demonstrated that their leaves exhibit overexpression of genes involved in nutrient transport, likely as a compensatory mechanism to counteract C deficiency resulting from their inability to perform photosynthesis (Lallemand et al., 2019). This difference should be due to the fact that the shaded individuals may still produce sufficient C by photosynthesis and their severity of C shortage is not as high as to require degrading cellar components in the present experiment.

I found fewer DEGs in stem tissues compared to the leaves. C provided from the mycorrhiza may be supplied to the whole plant through stems, and the jasmonic acids are synthesised in the leaves and transported to the roots via vascular bundles (Ryan & Moura, 2002). Although it is expected that the genes involved in the transporting

process such as “Transport” and “Transporter activity” should respond to the increased C provision from roots and production of jasmonic acid in the leaves, I can not detect any DEGs in stem tissues. Several genes such as RNA-polymerase and elongation factor which are involved in the transcription process were detected as DEGs in stems (Table 2-S1). My samples included the shoot meristem, thus these genes might be affected by the shading and regulating the leaf formations to adapt to the darker environment.

For the root tissues, no DEGs were found between shaded and unshaded individuals, including phytohormone related genes. The study on MX orchid species, *Epipactis helleborine*, showed that there are several DEGs which are related to symbiosis in roots between the green MX and albinistic FM individuals (Suetsugu et al., 2017). The mycorrhizal fungi of *E. helleborine* were restricted to *Wilcoxina* species group (at least for the study), while *C. goeringii* can symbiose with several groups of fungi such as saprotrophic-endophytic rhizoctonias and several ectomycorrhizal fungal taxa (Ogura-Tsujita et al., 2012; Selosse et al., 2022). In some cases, the same plant individuals are colonised by several species of the mycorrhizal fungi (Ogura-Tsujita *et al.*, 2012). I did not investigate the species composition of mycorrhizal fungi in *C. goeringii* in the present study. There might be a difference in composition of the fungi among each individual, and even between the roots of the same individual. If this is the case, the difference in above-ground environment might have less impact than the difference in mycorrhizal fungi composition in each part of the roots. This may explain the lack of DEGs between roots of shaded and unshaded individuals in *C. goeringii*.

Conclusions

Although the stable isotope ratios of C and N had not changed during the shading experiment, changes in gene expressions suggest that the MX orchid *C. goeringii* changed their trophic levels in response to the light conditions of the surrounding environments. My results indicate that the MX orchid *C. goeringii* might control their dependency against mycorrhizal fungi with jasmonic acid regulation in leaves, which could be a common mechanism with the AT plants. This suggests that MH plants adjust their dependency against fungi in response to changes in the environment as well as AT plants.

Acknowledgements

This work is supported in part by the Ichimura Foundation for New Technology. Computations for the RNA-seq were performed on the NIG supercomputer at ROIS National Institute of Genetics. I thank Dr. Hiroshi Yura for supporting the sampling at the Ecological Garden of the Natural History Museum and Institute, Chiba.

References

Akiyama K, Hayashi H. Strigolactones: chemical signals for fungal symbionts and parasitic weeds in plant roots. Ann Bot. 2006;97(6):925-31. doi: 10.1093/aob/mcl063.

Bates D, Maechler M, Bolker B, Walker S. Fitting linear mixed-effects models using lme4. Journal of Statistical Software. 2015;67(1):1-48. doi:10.18637/jss.v067.i01.

Bidartondo MI, Burghardt B, Gebauer G, Bruns TD, Read DJ. Changing partners in the dark: isotopic and molecular evidence of ectomycorrhizal liaisons between forest orchids and trees. Proc R Soc B. 2004;271:1799-1806.

Beyrle HF, Smith SE, Franco CMM, Peterson RL. Colonization of *Orchis morio* protocorms by a mycorrhizal fungus: effects of nitrogen nutrition and glyphosate in modifying the responses. Canadian Journal of Botany. 1995;73(8):1128-1140. doi:10.1139/b95-123.

Bolger AM, Lohse M, Usadel B. Trimmomatic: A flexible trimmer for Illumina Sequence Data. Bioinformatics. 2014;btu170.

Campos ML, Kang JH, Howe GA. Jasmonate-triggered plant immunity. J Chem Ecol. 2014;40:657–675. doi:10.1007/s10886-014-0468-3.

Craine JM, Brookshire ENJ, Cramer MD, Hasselquist NJ, Koba K, Marin-Spiotta E, Wang L. Ecological interpretations of nitrogen isotope ratios of terrestrial plants and soils. Plant and Soil. 2015;396(1):1-26.

Damesin C, Lelarge C. Carbon isotope composition of current-year shoots from *Fagus sylvatica* in relation to growth, respiration and use of reserves. Plant Cell Environ. 2003;26:207–219.

Dearnaley J, Perotto S, Selosse MA. Structure and development of orchid mycorrhizas. In: Martin F. ed. Molecular mycorrhizal symbiosis. New Jersey (NJ), Hoboken: Wiley-Blackwell; 2016. p. 63-86.

Du Puy D. Cribb PJ. The genus *Cymbidium*. Richmond (UK):Royal Botanic Gardens, Kew; 2007.

Farquhar GD, Richards RA. Isotopic composition of plant carbon correlates with water use efficiency of wheat genotypes. Aust. J. Plant Physiol. 1984;11:539–552.

Foo E, Yoneyama K, Hugill CJ, Quittenden LJ, Reid JB. Strigolactones and the regulation of pea symbioses in response to nitrate and phosphate deficiency. Molecular Plant. 2013;6:76–87.

Gebauer G, Meyer M. ^{15}N and ^{13}C natural abundance of autotrophic and mycoheterotrophic orchids provides insight into nitrogen and carbon gain from fungal association. New Phytologist. 2013;160:209–223.

Girlanda M, Segreto R, Cafasso D, Liebel HT, Rodda M, Ercole E, Cozzolino S, Gebauer G, Perotto S. Photosynthetic Mediterranean meadow orchids feature partial mycoheterotrophy and specific mycorrhizal associations. American Journal of Botany, 2011;98(7):1148-1163.

Gomes SIF, Merckx VSFT, Kehl J, Gebauer G. Mycoheterotrophic plants living on arbuscular mycorrhizal fungi are generally enriched in ^{13}C, ^{15}N and ^{2}H isotopes. J Ecol. 2020;108:1250–1261. doi:10.1111/1365-274513381.

Gonneau C, Jersáková J, de Tredern E, Till - Bottraud I, Saarinen K, Sauve M, Roy M, Hájek T, Selosse, MA. Photosynthesis in perennial mixotrophic Epipactis

spp.(Orchidaceae) contributes more to shoot and fruit biomass than to hypogeous survival. Journal of Ecology, 2014;102(5):1183-1194.

Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I, Adiconis X, Fan L, Raychowdhury R, Zeng Q, et al. Full-length transcriptome assembly from RNA-seq data without a reference genome. Nat Biotechnol. 2011;29(7):644-52. doi: 10.1038/nbt.1883.

Hanlon MT, Coenen C. Genetic evidence for auxin involvement in arbuscular mycorrhiza initiation. New Phytologist. 2011;189(3)701-709.

Hause B, Schaarschmidt S. The role of jasmonates in mutualistic symbioses between plants and soil-born microorganisms. Phytochemistry. 2009;70:1589–1599.

Hedden P, Thomas SG. Gibberellin biosynthesis and its regulation. Biochem J. 2012;444:11–25.

Jąkalski M, Minasiewicz J, Caius J, May M, Selosse MA, Delannoy, E. The Genomic Impact of Mycoheterotrophy in Orchids. Frontiers in plant science. 2021;12:632033.

Julou T, Burghardt B, Gebauer G, Berveiller D, Damesin C, Selosse MA. Mixotrophy in orchids: insights from a comparative study of green individuals and nonphotosynthetic individuals of *Cephalanthera damasonium*. New Phytologist. 2005:166(2);639-653.

Kobayashi K, Suetsugu K, Wada H. The leafless orchid *Cymbidium macrorhizon* performs photosynthesis in the pericarp during the fruiting season. Plant and Cell Physiology. 2021;62:472–481.

Lallemand F, Martin‐Magniette ML, Gilard F, Gakière B, Launay‐Avon A, Delannoy É, Selosse MA. In situ transcriptomic and metabolomic study of the loss of

photosynthesis in the leaves of mixotrophic plants exploiting fungi. The Plant Journal. 2019;98(5):826-841.

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

León J, Sánchez-Serrano JJ. Molecular biology of jasmonic acid biosynthesis in plants. Plant physiology and Biochemistry. 1999;37(5):373-380.

Li B, Dewey CN. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. BMC Bioinformatics. 2011;12:323. doi:10.1186/1471-2105-12-323.

Li MH, Liu KW, Li Z, Lu HC, Ye QL, Zhang D, Wang JY, Li YF, Zhong ZM, Liu X, et al. Genomes of Leafy and Leafless *Platanthera* Orchids Illuminate the Evolution of Mycoheterotrophy. Nature Plants. 2022;8(4):373–88. doi:10.1038/s41477-022-01127-9.

Liebel HT, Bidartondo MI, Gebauer G. Are carbon and nitrogen exchange between fungi and the orchid Goodyera repens affected by irradiance?. Annals of Botany, 2015;115(2):251-261.

Matsuda Y, Shimizu S, Mori M, Ito SI, Selosse MA. Seasonal and environmental changes of mycorrhizal associations and heterotrophy levels in mixotrophic Pyrola japonica (Ericaceae) growing under different light environments. American Journal of Botany, 2012;99(7):1177-1188.

Merckx V, Stöckel M, Fleischmann A, Bruns TD, Gebauer G. ^{15}N and ^{13}C natural abundance of two mycoheterotrophic and putative partially

mycoheterotrophic species associated with arbuscular mycorrhizal fungi. New Phytol. 2010;188:590–596. doi:10.1111/j.1469-8137.2010.03365.x.

Merckx VSFT, Freudenstein JV, Kissling J, Christenhusz MJM, Stotler RE, Crandall-Stotler B, Wickett N, Rudall PJ, van de Kamer HM, Maas PJM. Taxonomy and classification. In: Merckx VSFT, editor. Mycoheterotrophy: the biology of plants living on fungi. New York (NY): Springer; 2013a. p. 19–101.

Merckx VSFT, Mennes CB, Peay KG, Geml J. Evolution and diversification. In: Merckx VSFT, editor. Mycoheterotrophy: the biology of plants living on fungi. New York (NY): Springer; 2013b. p. 215–244.

Miura C, Yamaguchi K, Miyahara R, Yamamoto T, Fuj, M, Yagame T, Imaizumi-Anraku H, Yamato M, Shigenobu S, Kaminaka H. The mycoheterotrophic symbiosis between orchids and mycorrhizal fungi possesses major components shared with mutualistic plant-mycorrhizal symbioses. Molecular Plant-Microbe Interactions. 2018;31(10):1032-1047.

Mostofa MG, Li W, Nguyen KH, Fujita M, Tran LSP. Strigolactones in plant adaptation to abiotic stresses: An emerging avenue of plant research. Plant, cell & environment. 2018;41(10):2227-2243.

Motomura H, Selosse MA, Martos F, Kagawa A, Yukawa T. Mycoheterotrophy evolved from mixotrophic ancestors: Evidence in Cymbidium (Orchidaceae). Annals of Botany, 2010;106(4):573–581.

Nagata M, Yamamoto N, Shigeyama T, Terasawa Y, Anai T, Sakai T, Inada S, Arma S, Hashiguchi M, Akashi R, Nakayama H, et al. Red/far red light controls arbuscular

mycorrhizal colonization via jasmonic acid and strigolactone signaling. Plant and Cell Physiology. 2015;56(11):2100-2109.

Ogura-Tsujita Y, Yokoyama J, Miyoshi K, Yukawa T. Shifts in mycorrhizal fungi during the evolution of autotrophy to mycoheterotrophy in *Cymbidium* (Orchidaceae). Am J Bot. 2012;99(7):1158–1176.

Preiss K, Adam IK, Gebauer G. Irradiance governs exploitation of fungi: fine-tuning of carbon gain by two partially myco-heterotrophic orchids. Proc Biol Sci. 2010;277(1686):1333-1336. doi:10.1098/rspb.2009.1966.

R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. 2021.

Roy M, Gonneau C, Rocheteau A, Berveiller D, Thomas JC, Damesin C, Selosse MA. Why do mixotrophic plants stay green? A comparison between green and achlorophyllous orchid individuals in situ. Ecological Monographs. 2013:83(1);95-117.

Ruiz‐Lozano JM, Aroca R, Zamarreño ÁM, Molina S, Andreo‐Jiménez B, Porcel R, García-Mina JM, Ruyter-Spira C, López-Ráez JA. Arbuscular mycorrhizal symbiosis induces strigolactone biosynthesis under drought and improves drought tolerance in lettuce and tomato. Plant, Cell & Environment. 2016;39:441–452.

Ryan CA, Moura DS. Systemic wound signaling in plants: A new perception. Proc. Natl. Acad. Sci. USA. 2002;99:6519–6520.

Selosse MA, Petrolli R, Mujica MI, Laurent L, Perez-Lamarque B, Figura T, Bourceret A, Jacquemyn H, Li T, Gao J, et al. The Waiting Room Hypothesis revisited by

orchids: were orchid mycorrhizal fungi recruited among root endophytes?. Annals of Botany, 2022:129(3);259-270.

Selosse MA, Roy M. Green plants that feed on fungi: facts and questions about mixotrophy. Trends in plant science, 2009;14(2):64-70.

Suetsugu K. Autonomous self-pollination and insect visitors in partially and fully mycoheterotrophic species of Cymbidium (Orchidaceae). Journal of Plant Research. 2015:128;115-125.

Suetsugu K, Yamato M, Miura C, Yamaguchi K, Takahashi K, Ida Y, ... & Kaminaka, H. Comparison of green and albino individuals of the partially mycoheterotrophic orchid Epipactis helleborine on molecular identities of mycorrhizal fungi, nutritional modes and gene expression in mycorrhizal roots. Molecular Ecology. 2017;26(6):1652-1669.

Takeda N, Handa Y, Tsuzuki D, Kojima M, Sakakibara H, Kawaguchi M. Gibberellins interfere with symbiosis signaling and gene expression and alter colonization by arbuscular mycorrhizal fungi in *Lotus japonicus*. Plant Physiology. 2015;167(2):545–557. doi:10.1104/pp.114.247700.

Timilsena PR, Barrett CF, Nelson AP, Wafula EK, Ayyampalayam S, McNeal JR, Yukawa T, Givnish TJ, Graham SW, Pires JC, et al. Phylotranscriptomic analyses of mycoheterotrophic monocots show a continuum of convergent evolutionary changes in expressed nuclear genes from three independent nonphotosynthetic lineages. Genome Biology and Evolution. 2022;, evac183. https://doi.org/10.1093/gbe/evac183

Yuan Y, Jin X, Liu J, Zhao X, Zhou J, Wang X, Wang D, Lai C, Xu W, Huang J, et al. The *Gastrodia elata* Genome Provides Insights into Plant Adaptation to Heterotrophy. Nature Communications. 2018;9(1):1615. doi:10.1038/s41467-018-03423-5.

Zhang GQ, Liu KW, Li Z, Lohaus R, Hsiao YY, Niu SC, Wang JY, Lin YC, Xu Q, Chen LJ, et al. The *Apostasia* genome and the evolution of orchids. Nature. 2017;549(7672):379–383. doi:10.1038/nature23897.

Figures and Tables

Table 2-1. Number of differential expressed genes found in each tissue of *Cymbidium goeringii*.

	upregulated	downregulated
Leaf	73	13
Stem	3	5
Root	0	0

Table 2-2. GO enrichment in leaf tissue of shaded individuals of *Cymbidium goeringii*.

+: Upregulated GOs; -: Downregulated GOs.

GO Term (Accession No.)	Reference count	DEG count	Expression	FDR
Biological process				
Response to herbivore (GO:0080027)	7	3	+	1.53E-02
protein complex oligomerization (GO:0051259)	31	4	+	3.48E-02
Response to jasmonic acid (GO:0009753)	267	12	+	1.60E-03
Response to fatty acid (GO:0070542)	269	12	+	8.65E-04
Response to fungus (GO:0009620)	406	12	+	1.77E-02
Defense response to bacterium (GO:0042742)	416	12	+	1.91E-02
Response to bacterium (GO:0009617)	522	14	+	1.16E-02

Response to external biotic stimulus (GO:0043207)	1082	20	+	1.77E-02
Response to other organism (GO:0051707)	1082	20	+	1.58E-02
Response to biotic stimulus (GO:0009607)	1084	20	+	1.46E-02
Biological process involved in interspecies interaction between organisms (GO:0044419)	1092	20	+	1.47E-02
Defence response (GO:0006952)	988	18	+	3.40E-02
Response to external stimulus (GO:0009605)	1371	24	+	1.07E-02
Response to stress (GO:0006950)	2647	37	+	4.39E-03
Cellular component				
protein-containing complex (GO:0032991)	1977	1	-	1.38E-02
Molecular function				

No enriched GO terms				

Figure 2-1. $\delta^{13}C$ and $\delta^{15}N$ values of autotrophic (AT; crosses), mixotrophic (MX; circles), and fully mycoheterotrophic (FM; triangles) plants. The box plots on the right side and the top represent $\delta^{13}C$ values and $\delta^{15}N$ of each trophic level, respectively. Two red circles indicate *Cephalanthera falcata*. See text and figure 2-S1 for the included species.

Figure 2-2. Correlation between the environmental factors and isotope ratios. Simple regression analysis was used to determinate each relationship. **A**, canopy openness and $\delta^{13}C$ ($P < 0.001$); **B**, canopy openness and $\delta^{15}N$ ($P = 0.92$); **C**, temperature and $\delta^{13}C$ ($P = 0.92$), **D**; temperature and $\delta^{15}N$ ($P = 0.88$); **E**, precipitation and $\delta^{13}C$ ($P = 0.57$); **F**, precipitation and $\delta^{15}N$ ($P = 0.28$); **G**, bulb numbers and $\delta^{13}C$ ($P = 0.53$); and **H**, bulb numbers and $\delta^{15}N$ ($P = 0.097$).

Figure 2-3. $\delta^{13}C$ and $\delta^{15}N$ values of leaf tissue in shaded and unshaded individuals of *C. goeringii*. The left plots are the values from the first day of the shading, and right plots are from the last day of the shading. Null models were selected based on AIC value. **A**, $\delta^{13}C$; **B** , $\delta^{15}N$.

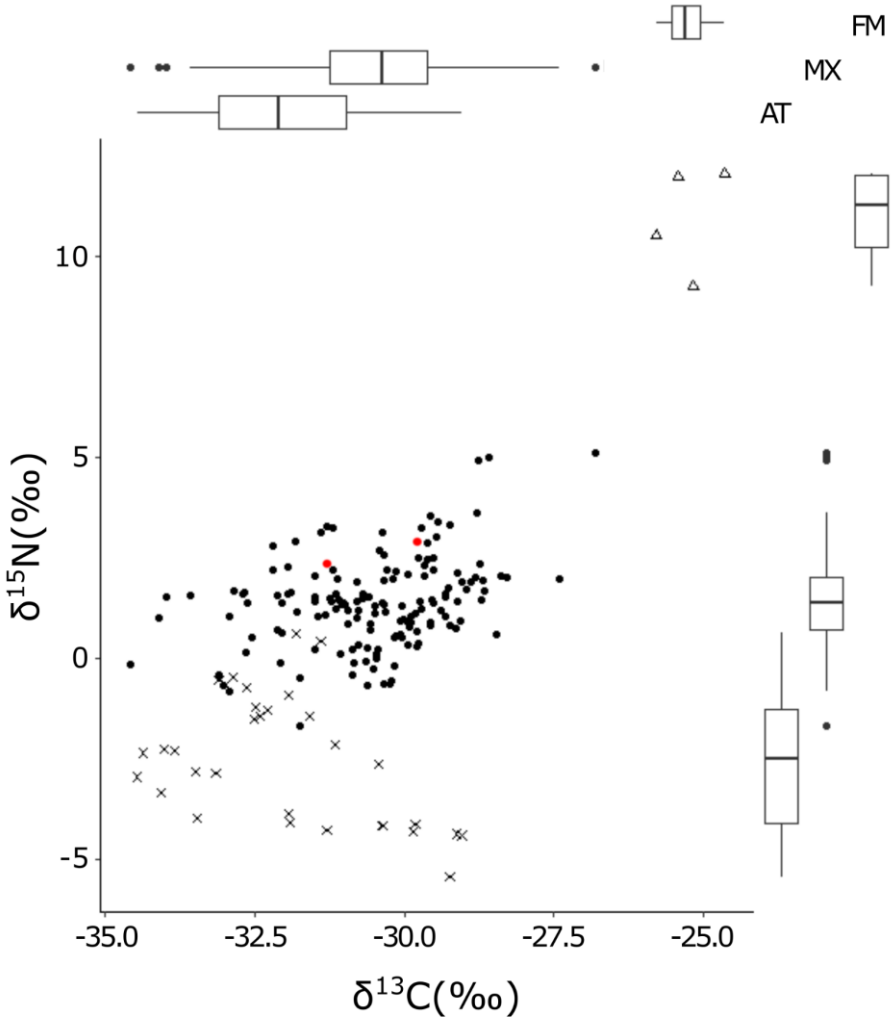


Figure 2-1.

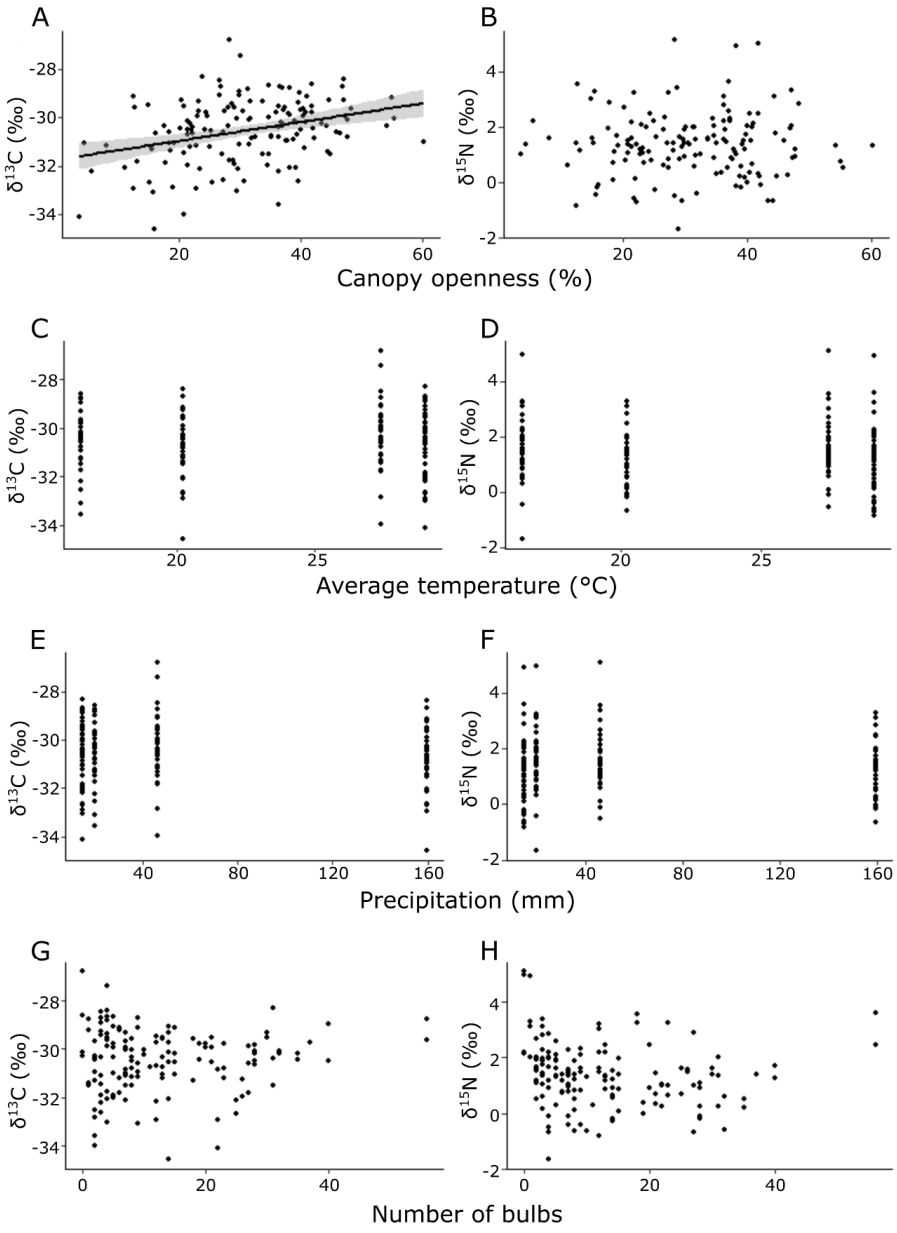

A
B
C
D
E
F
G
H
δ13C (‰)
δ15N (‰)
Canopy openness (%)
Average temperature (°C)
Precipitation (mm)
Number of bulbs

Figure 2-2.

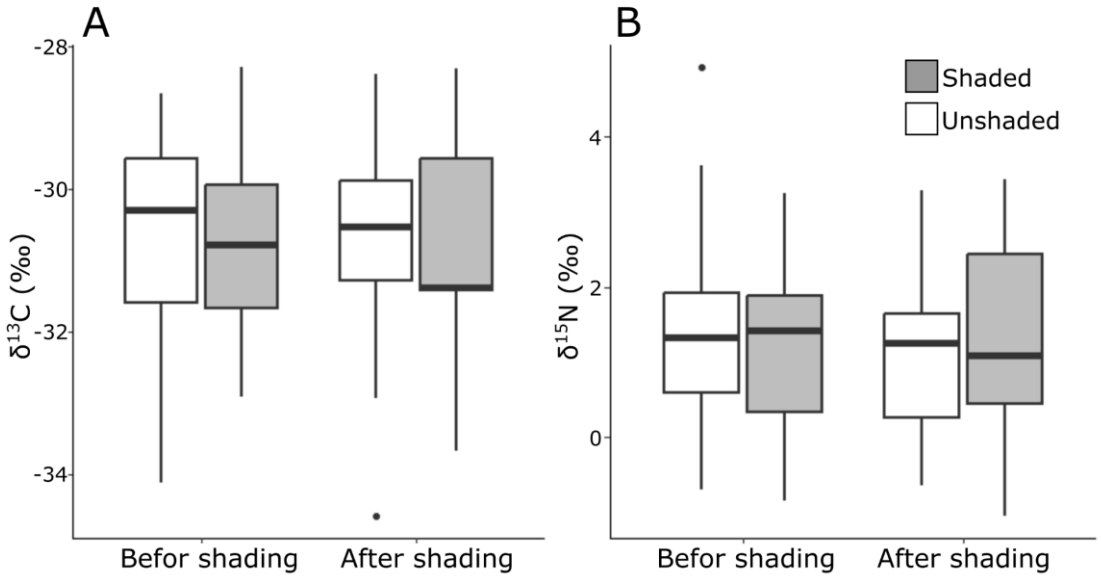

A
B
-28
-30
-32
-34
δ13C (‰)
4
2
0
δ15N (‰)
Shaded
Unshaded
Befor shading
After shading
Befor shading
After shading

Figure 2-3.

Table 2-S1. List of differential expressed genes found in each tissue of *Cymbidium goeringii*. +: overexpressed gene; -: underexpressed gene.

Sequence ID	Expression	Best hit (BLAST search against Arabidopsis proteins)	Description (based on Araport 11)
Leaf			
TRINITY_DN23135_c0_g1	+	AT2G44800	2-oxoglutarate (2OG) and Fe(II)-dependent oxygenase superfamily protein
TRINITY_DN7279_c0_g1	+	AT3G11180	2-oxoglutarate (2OG) and Fe(II)-dependent oxygenase superfamily protein
TRINITY_DN9705_c0_g1	+	AT5G13490	ADP/ATP carrier 2
TRINITY_DN128_c1_g1	+	AT5G16120	alpha/beta-Hydrolases superfamily protein
TRINITY_DN3161_c0_g2	+	AT5G37710	alpha/beta-Hydrolases superfamily protein
TRINITY_DN6606_c1_g1	+	AT5G37710	alpha/beta-Hydrolases superfamily protein

TRINITY_DN19712_c0_g1	+	AT5G53050	alpha/beta-Hydrolases superfamily protein
TRINITY_DN32488_c0_g1	+	AT1G02660	alpha/beta-Hydrolases superfamily protein
TRINITY_DN12862_c0_g1	+	AT3G10960	AZA-guanine resistant1
TRINITY_DN27386_c0_g1	-	AT1G62710	beta vacuolar processing enzyme
TRINITY_DN5169_c0_g1	+	AT2G24240	BTB/POZ domain with WD40/YVTN repeat-like protein
TRINITY_DN11464_c0_g1	+	AT5G05240	cation-transporting ATPase
TRINITY_DN11923_c0_g1	-	AT5G13930	Chalcone and stilbene synthase family protein
TRINITY_DN5838_c0_g1	-	AT5G13930	Chalcone and stilbene synthase family protein
TRINITY_DN39192_c1_g1	-	AT3G22840	Chlorophyll A-B binding family protein
TRINITY_DN4828_c2_g1	-	AT4G37970	cinnamyl alcohol dehydrogenase 6
TRINITY_DN878_c0_g1	-	AT5G43310	COP1-interacting protein-like protein

TRINITY_DN4933_c0_g1	+	AT4G24460	CRT (chloroquine-resistance transporter)-like transporter 2
TRINITY_DN1121_c0_g1	+	AT2G36130	Cyclophilin-like peptidyl-prolyl cis-trans isomerase family protein
TRINITY_DN1287_c8_g1	-	AT4G23160	cysteine-rich RECEPTOR-like kinase
TRINITY_DN33119_c0_g1	+	AT3G01900	cytochrome P450
TRINITY_DN4375_c0_g1	-	AT2G41120	DUF309 domain protein
TRINITY_DN772_c2_g1	+	AT2G38410	ENTH/VHS/GAT family protein
TRINITY_DN23374_c0_g1	+	AT3G23240	ethylene response factor 1
TRINITY_DN16988_c0_g1	+	AT4G35930	F-box family protein
TRINITY_DN24344_c0_g1	+	AT1G70140	formin 8
TRINITY_DN7249_c0_g1	-	AT4G32940	gamma vacuolar processing enzyme
TRINITY_DN7611_c0_g1	+	AT1G78440	gibberellin 2-beta-dioxygenase
TRINITY_DN9070_c0_g1	+	AT3G62760	Glutathione S-transferase family protein
TRINITY_DN5559_c1_g1	+	AT5G32470	heme oxygenase-like

TRINITY_DN7506_c0_g1	+	AT4G37680	heptahelical protein 4
TRINITY_DN1747_c0_g1	+	AT4G21870	HSP20-like chaperones superfamily protein
TRINITY_DN14523_c0_g1	+	AT5G01210	HXXXD-type acyl-transferase family protein
TRINITY_DN2123_c6_g3	+	AT4G11430	hydroxyproline-rich glycoprotein family protein
TRINITY_DN159_c0_g1	+	AT1G76070	hypothetical protein
TRINITY_DN7893_c0_g1	+	AT3G23230	Integrase-type DNA-binding superfamily protein
TRINITY_DN571_c0_g1	+	AT5G64750	Integrase-type DNA-binding superfamily protein
TRINITY_DN18410_c0_g1	+	AT3G21720	isocitrate lyase
TRINITY_DN8370_c0_g1	+	AT3G45140	lipoxygenase 2
TRINITY_DN3520_c1_g1	+	AT4G34950	Major facilitator superfamily protein
TRINITY_DN25857_c0_g2	+	AT2G45040	Matrixin family protein
TRINITY_DN2424_c0_g2	+	AT1G64660	methionine gamma-lyase
TRINITY_DN1986_c0_g1	+	AT5G55090	mitogen-activated protein kinase kinase kinase 15

TRINITY_DN3642_c1_g1	+	AT1G49920	MuDR family transposase
TRINITY_DN7538_c0_g1	-	AT3G13080	multidrug resistance-associated protein 3
TRINITY_DN7349_c0_g1	+	AT3G13540	myb domain protein 5
TRINITY_DN7349_c0_g2	+	AT3G13540	myb domain protein 5
TRINITY_DN2166_c0_g1	+	AT1G69850	nitrate transporter 1:2
TRINITY_DN5066_c1_g1	+	AT3G20660	organic cation/carnitine transporter4
TRINITY_DN648_c0_g1	+	AT4G11650	osmotin 34
TRINITY_DN1788_c1_g1	+	AT1G78230	Outer arm dynein light chain 1 protein
TRINITY_DN2528_c0_g1	+	AT5G48840	pantoate-beta-alanine ligase
TRINITY_DN1485_c3_g2	+	AT1G78780	pathogenesis-related family protein
TRINITY_DN17817_c0_g1	+	AT5G24070	Peroxidase superfamily protein
TRINITY_DN4086_c0_g2	+	AT1G52200	PLAC8 family protein
TRINITY_DN63210_c0_g1	+	AT2G35930	plant U-box 23
TRINITY_DN6371_c0_g1	+	AT2G35930	plant U-box 23

TRINITY_DN2090_c0_g1	-	AT3G61870	plant/protein
TRINITY_DN6765_c0_g1	-	AT1G33970	P-loop containing nucleoside triphosphate hydrolases superfamily protein
TRINITY_DN7130_c0_g1	+	AT1G76600	poly polymerase
TRINITY_DN20614_c0_g1	+	AT1G28390	Protein kinase superfamily protein
TRINITY_DN4865_c0_g1	+	AT1G54820	Protein kinase superfamily protein
TRINITY_DN70959_c0_g2	+	AT1G07160	Protein phosphatase 2C family protein
TRINITY_DN2263_c0_g1	+	AT1G55230	proteinase inhibitor I4
TRINITY_DN3134_c0_g2	+	AT4G14680	Pseudouridine synthase/archaeosine transglycosylase-like family protein
TRINITY_DN4784_c0_g1	+	AT5G21930	P-type ATPase of Arabidopsis 2
TRINITY_DN12136_c0_g1	+	AT5G50400	purple acid phosphatase 27
TRINITY_DN2130_c1_g1	+	AT1G34060	Pyridoxal phosphate (PLP)-dependent transferases superfamily protein

TRINITY_DN3337_c1_g1	+	AT3G15060	RAB GTPase homolog A1G
TRINITY_DN4723_c0_g1	-	AT4G21470	riboflavin kinase/FMN hydrolase
TRINITY_DN19539_c0_g1	+	AT5G20885	RING/U-box superfamily protein
TRINITY_DN2272_c0_g2	+	AT3G24255	RNA-directed DNA polymerase (reverse transcriptase)-related family protein
TRINITY_DN1524_c0_g1	+	AT2G41380	S-adenosyl-L-methionine-dependent methyltransferases superfamily protein
TRINITY_DN1046_c0_g1	+	AT1G32740	SBP (S-ribonuclease binding protein) family protein
TRINITY_DN4289_c0_g1	+	AT4G10170	SNARE-like superfamily protein
TRINITY_DN4801_c2_g1	+	AT4G32480	sugar phosphate exchanger
TRINITY_DN13033_c4_g1	+	AT4G16740	terpene synthase 03
TRINITY_DN331_c1_g1	+	AT1G61120	terpene synthase 04
TRINITY_DN591_c3_g1	+	AT1G74950	TIFY domain/Divergent CCT motif family protein
TRINITY_DN4787_c0_g2	+	AT2G36750	UDP-glucosyl transferase 73C1

TRINITY_DN562_c0_g1	+	AT2G36760	UDP-glucosyl transferase 73C2
TRINITY_DN2116_c2_g2	+	AT1G68450	VQ motif-containing protein
TRINITY_DN5231_c1_g1	+	AT1G80840	WRKY DNA-binding protein 40
TRINITY_DN1953_c0_g1	+	AT5G64810	WRKY DNA-binding protein 51
TRINITY_DN378_c3_g1	+	AT5G13080	WRKY DNA-binding protein 75
TRINITY_DN9795_c0_g1	+	AT5G13080	WRKY DNA-binding protein 75
Stem			
TRINITY_DN20977_c0_g2	-	AT2G25010	Aminotransferase-like plant mobile domain family protein
TRINITY_DN38558_c0_g1	+	AT5G12110	elongation factor 1-beta 1
TRINITY_DN6710_c0_g1	-	AT1G50610	Leucine-rich repeat protein kinase family protein
TRINITY_DN1044_c0_g1	+	AT3G59220	pirin
TRINITY_DN1903_c0_g1	+	AT3G49500	RNA-dependent RNA polymerase 6
TRINITY_DN15967_c0_g2	-	AT5G51080	RNase H family protein

TRINITY_DN19086_c0_g1	-	AT4G37925	subunit NDH-M of NAD(P)H:plastoquinone dehydrogenase complex
TRINITY_DN877_c0_g1	-	AT4G13980	winged-helix DNA-binding transcription factor family protein

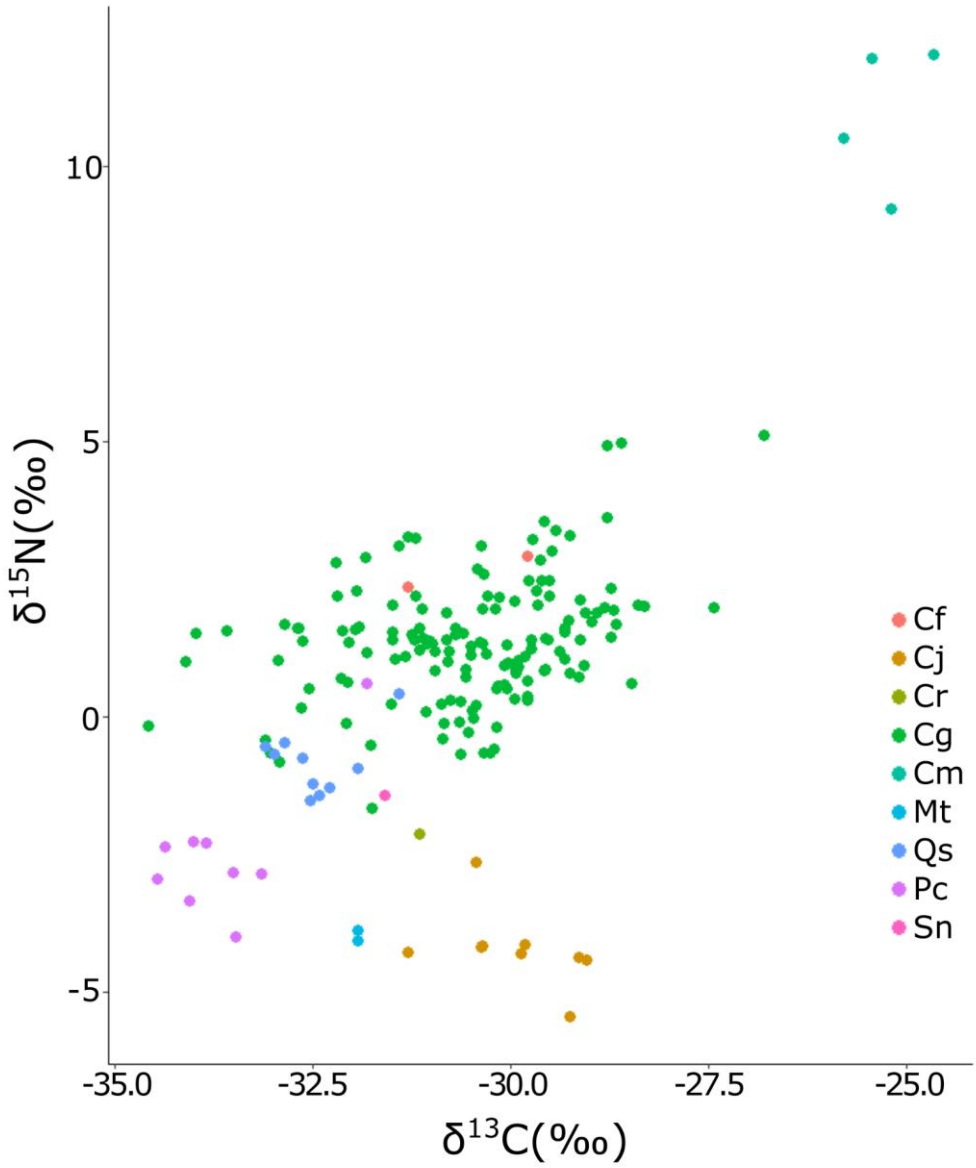


Figure 2-S1. $\delta^{13}C$ and $\delta^{15}N$ values of the plants collected for the present study.

Abbreviation: Cf, *Cephalanthera falcata*; Cj, *Cleyera japonica*; Cr, *Carpinus tschonoskii*; Cg, *Cymbidium goeringii*; Cm, *Cymbidium macrorhizon*; MT, *Machilus thunbergii*; Qs, *Quercus serrata*; Pc, *Pleioblastus chino*; and Sn, *Sasa nipponica*.

Chapter 3: The evolution and genetic background of mycoheterotrophy in Cymbidium

Abstract

Fully mycoheterotrophic (FM) plants are the plants that totally rely its carbon needs to their mycorrhizal fungi. They exhibit some interesting features such as leafless morphology, defective photosynthesis, and specialized mycorrhizal fungi. Despite these features, little is known about the genetic backgrounds of FM plants. Here, I assembled and compared the genome of three related species, including two partial mycoheterotrophic (Mixotrophic; MX) orchids *Cymbidium goeringii* and *C. nagifolium*, and fully mycoheterotrophic (FM) orchid, *C. macrorhizon* to investigate the changes in the genomes that occurred during the evolution of FM plants. I provide the genome assemblies of *C. goeringii* which size is 2.2 Gb and contains 17,709 contigs with 27,753 protein-coding genes. I also provided the draft genome assemblies of *C. nagifolium* and *C. macrorhizon* by mapping their sequences to the genome of *C. goeringii.* Comparative analysis between the genomes of these three species shows that in the genome of FM species *C. macrorhizon*, there are many mutations occurring in several genes which might be related to photosynthesis, leaf morphogenesis, root morphogenesis, and mycorrhizal symbiosis, compared to MX species. This may be attributed to the morphology and physiology of FM features in *C. macrorhizon*, which have no leaves and perform only limited photosynthesis. My study finds out the potential molecular basis underlying the FM species, providing insights into the

genetic backgrounds of evolution of mycoheterotrophic plants.

Introduction

The characteristic features of plants are the photosynthetic ability with specific organs of leaves and chloroplast. The plants carrying out the photosynthesis are called autotrophic (AT) plants with obtaining their carbon needs for their growth. However, some species can obtain their carbon needs from their mycorrhizal fungi, and those plants are so-called mycoheterotrophic (MH) plants.

Among the MH plants, some plant species, however, do not have photosynthetic organs and obtain the whole carbon needs by parasitizing mycorrhizal fungi without performing photosynthesis (Leake, 1994). These plants are called fully mycoheterotrophic (FM) plants, and about 880 species are known as FM (Merckx et al., 2013). There are also other plants called partially mycoheterotrophic (mixotrophic; MX) plants, which are photosynthetic but also obtain carbon from fungi (Selosse & Roy, 2009). They range in physiology, ecology and morphology from AT-like to FM-like, and are considered to be a preadaptation for FM (Selosse & Roy, 2009).

Both MX and FM plants have been studied for their special ecology and morphology compared to AT plants. Phylogenetic studies suggested that the most FM plants evolved from AT plants, with MX plants as the intermediate step in terms of morphology, physiology and mycorrhizal compositions (Jacquemyn & Merckx, 2019). Previous studies focused on physiology showed that the carbon supplied to MH plants ultimately comes from photosynthesis in surrounding AT plants (Selosse & Roy, 2009). Physiological backgrounds of mycoheterotrophy are specialized but have similarity

with AT plants, suggesting an application of their system (Suetsugu et al., 2017; Yuan et al., 2018; Miura et al., 2018; Li et al., 2022). The genomic studies in MH orchids suggest that genes such as MADSBOX genes, photoreceptor genes, and auxin transporter genes might be involved with the specialized morphology in MH plants (Zhang et al., 2017; Li et al., 2022).

Orchidaceae comprise more than 25,000 species, almost all of which are at least MH during their germination stage (Jacquemyn & Merckx, 2019). There are more than 300 species that lack leaves and are expected to be FM (Jacquemyn & Merckx, 2019). These species transited from non-FM species at least 30 times independently (Wang et al., 2021). Hence, Orchidaceae could be one of the model groups to investigate the evolution of mycoheterotrophy. Recent study comparing the genomes of closely related MX and FM orchids revered the potential molecular basis underlying mycoheterotrophy (Li *et al.*, 2022). They found that FM plant genome shows increased substitution rates and gene loss, especially for the photoreceptor and auxin transporter genes. However, they compared only two species,which might cause the failure of direction of the genes that are under selection in FM plants. It is also not clear if there are common mechanisms underlying these parallel evolutions among Orchidaceae.

Cymbidium is an orchid genus comprising about 75 species that vary in their nutrient modes (Du Puy & Cribb, 2007; Motomura *et al.*, 2010). The genus includes two almost FM and about 20 putative MX, and the rest are considered as initial mycoheterotrophic species (Du Puy & Cribb, 2007; Motomura *et al.*, 2010; Ogura-Tsujita et al., 2012). Recent researches had assembled the genomes of three putative MX *Cymbidium* species, although they mainly focused on floral morphology, not the

evolution of mycoheterotrophy (Yang et al., 2021; Ai et al., 2021; Sun et al., 2021; Chung et al., 2022).

Here, I present draft assembled genomes of three closely related species including two MX species *C. goeringii* and *C. nagifolium* (Motomura *et al.*, 2010; Chagi et al., in prep., See Chapter 1), and an almost FM species *C. macrorhizon* without leaves (Motomura *et al.*, 2010; Kobayashi *et al.*, 2021; Fig. i). *Cymbidium goeringii* and *C. nagifolium* has normal leaves, and has association with both saprophytic and ectomycorrhizal fungi (Motomura *et al.*, 2010; Ogura-Tsujita *et al.*, 2012). *Cymbidium macrorhizon* lacks leaves, and has association only with ectomycorrhizal fungi, although it is closely related to *C. nagifolium* (Yukawa *et al.*, 2002; Motomura *et al.*, 2010; Ogura-Tsujita *et al.*, 2012). By comparing the genomes of these three sister species, I could find out the genetic backgrounds in evolution of mycoheterotrophy in *Cymbidium*. I compared each gene of these species to find out what kind of genes are involved in the process of evolution of mycoheterotrophy. Further, I categorized the genes by its expression patterns, and compared their substitution rates to find out the patterns in evolutionary trends among the tissues.

Materials and Methods

Sample preparation

For genome sequencing, the leaves of a single individual of *Cymbidium goeringii* from the wild population at the temperate broadleaf deciduous forest located in the Ecological Garden of the Natural History Museum and Institute, Chiba (Chuo-ku, Chiba-Shi, Chiba-ken, Japan; E35.60°, N140.14°) were sampled. For the *C. macrorhizon*, the scapes of multiple individuals were sampled from the same locality as *C. goeringii*, because of leafless features of this species. The leaves of a single individual of *C. nagifolium* from a private collection of an author (K. Chagi) were sampled. High-molecular-weight DNA was extracted by NucleoBond High Molecular Weight DNA Kit (Macherey-Nagel, Germany) according to the manufacturer's protocol. The extracted DNA was quality-checked using agarose gel electrophoresis, Qubit 2.0 Fluorometer (Invitrogen, U.S.A.), and NanoDrop Lite (Thermo Fisher Scientific, U.S.A.).

I sequenced the transcriptome to be used to estimate the completeness of the assembled genome sequence. For transcriptome sequencing, three different *C. goeringii* tissues (leaves, stems, and roots) from multiple individuals were also collected at the same locality (including the same individual used for genome sequencing). Total RNA was extracted using Maxwell 16 LEV Plant RNA Kit with the Maxwell 16 Research Instrument (Promega, U.S.A.) according to the manufacturer's instructions. RNA concentrations were measured using a Qubit 2.0 Fluorometer (Invitrogen, U.S.A.), and electrophoresis on agarose gel was performed to check the RNA degradation. RNA purity was estimated using a BioSpec-nano (Shimadzu, Japan).

Genome sequencing, assembly and genome-quality evaluation

To generate the short but abundant DNA reads, the qualified DNA of *C. goeringii, C. macrorhizon*, and *C. nagifolium* were constructed using KAPA Library Prep Kits for Illumina (Kapa Biosystems Inc, U.S.A.). Paired-end (150 bp) whole genome sequencing was performed on the Illumina NovaSeq6000 platform (Illumina, U.S.A.). To generate the long DNA reads, the qualified DNA of *C. goeringii* was sequenced by MinION mk1B equipped by flow cell R9.4.1 (Oxford Nanopore Technologies, U.K.), using a Ligation-sequencing kit (SQK-LSK109) according to the manufacturer's protocol. The generated fast5 data were basecalled by Guppy v5.0.16 (https://community.nanoporetech.com/guppy) and converted into HAC DNA reads. I used flye v2.8.1 to assemble the genome. Racon v1.4.7 was used for four iterations to polish the assembly result (Vaser et al., 2017). Then the polished assembly was applied to Medaka v1.3.3 (https://github.com/nanoporetech/medaka) to create consensus sequences. The short read DNA of *C. goeringii* were mapped to the consensus sequences using minimap2 v2.17, and NTedit v1.3.5 was performed to polish the consensus sequences (Li, 2018; Warren et al., 2019). Purge Haplotigs v1.2.5 were used to remove the haplotigs and create the primary contigs (Roach et al., 2018). These primary contigs were used as the *C. goeringii* genome in my present study. I performed BUSCO analysis using gene sets of Embryophyta and Liliopsida to evaluate the completeness of the assembled genome using gVolante (Nishimura *et al.*, 2017; Manni et al., 2021). The short read DNA of *C. nagifolium* and *C. macrorhizon* were mapped to this genome of *C. goeringii* using minimap2, and variant-called by DeepVariant v1.3.0 (Poplin et al., 2018).

From the quality checked RNA, the cDNA library was constructed using TruSeq RNA Sample Prep Kits. Paired-end (150 bp) RNA sequencing (RNA-seq) was performed on the Illumina NovaSeq6000 platform. These RNA-seq reads were mapped to the genome of *C. goeringii* using STAR v2.7.9 and checked the percentage of uniquely mapped reads to evaluate the quality of the *C. goeringii* genome (Dobin et al., 2013).

Genome annotation

For de novo prediction of repetitive elements in *C. goeringii*, I used RepeatModeler v 2.0.1 (Jullien et al., 2020) with default parameters. RepeatMasker v 4.1.2 (http://www.repeatmasker.org) was used to identify repeats in the assembled genome.

I used Breaker v2.1.6 to predict the coding genes in the *C. goeringii* genome (Brůna et al., 2021). The RNA-seq reads mapped on the repeat-masked genome and protein sequences of plants downloaded from OrthoDB (https://v100.orthodb.org/download/odb10_plants_fasta.tar.gz) were utilized as the hints file. Among the predicted genes by Braker, those with start and stop codon were regarded as validly detected genes. The gene expressions of leaf, stem, and root were estimated by RNA-seq data from each tissue. The gene was annotated by BLAST searches against all protein sequences of *Arabidopsis thaliana* (Araport11) with e-value $<$ 0.0001 as the threshold.

Mutation patterns of genes by function

SnpEff v5.1 were used to detect variants which are SNPs and indels for both *C. nagifolium* and *C. macrorhizon* compared with *C. goeringii* (Cingolani et al., 2012). The numbers of SNPs and indels were combined as a number of variants. In each coding sequence (CDS), the number of variants per CDS length for *C. nagifolium* and *C. macrorhizon,* respectively, were examined for: (1) each gene, (2) each tissue with RNA-seq data; (3) each GO term involves morphogenesis of leaf, root and flower; and (4) each GO term belonging to biological process.

Detection of genes under selection in FM species

After removing the genes with indels, orthologous genes from the three focal species were aligned, and used to calculate the dN/dS (ω). ω was calculated for each gene of *C. nagifolium* and *C. macrorhizon*, respectively, as the query and genes of *C. goeringii* as the reference. To detect the genes under selection only in FM species, *C. macrorhizon*, but not in MX species, *C. goeringii* and *C. nagifolium*, I compared branch models "0" (branches with single ω) and "2" (branches with independent ω) with codeml implied in paml v4.8 (Yang, 2007). I provided a newick format tree with *C. nagifolium* and *C. macrorhizon* as a sister species and *C. goeringii* as an outgroup as the starting points for the estimation referring to Yukawa *et al.* (2002). The genes with $P < 0.05$ were selected as the genes under selection only in *C. macrorhizon*.

Results

Genome assembly and annotation of *C. goeringii*

Total of 71 Gb and 93 Gb of DNA were obtained from the MinION and Illumina sequencer, respectively. The assembled genome was 2.2 Gb composed of about 18,000 contigs with an N50 of 240 Kb, and mean coverage of 22. The BUSCO of the genome was 96% and 93% for embryophyta and liliopsida databases, respectively. Mapping rate of RNA sequences from each tissue was on average 87% (leaf, 89%; stem, 91%; and root, 80%). Total of 1.8 Gb of repetitive elements, which constitutes 82% of the assembled genome, were detected. I found 27,753 genes in this genome. Detailed statistics of the genome were reported in Table 3-S1. There were 9,014, 7,617, and 1,108 genes expressed in leaf, stem and root, respectively. Among them, 867 genes were expressed in all the three tissues and 2,047 genes were expressed in leaf, 1,576 in stem, 114 in root specifically.

Interspecific comparison of genes

For *C. nagifolium*, there were 53,634,320 SNPs and 4,296,058 indels compared to the *C. goeringii* genome. If limited to the coding regions, there were 381,273 SNPs and 5,546 indels (Table 3-S2). For *C. macrorhizon*, there were 51,759,843 SNPs and 4,250,281 indels compared to the *C. goeringii* genome. There were 404,608 SNPs and 5,932 indels if limited to the coding regions (Table 3-S2). The variants per CDS length were higher in *C. macrorhizon* than in *C. nagifolium* in genes, which includes *CRRSP55*, *TPS03* and *HCT*

(Table 3-1). When the genes expressed in each tissue were compared based on the number of variants per CDS length, the genes expressed only in leaves and stems, respectively, showed a lower rate than those in roots (Table 3-2). The gene expressed in all these three tissues showed the similar rate of variants with those in leaves and stems. This trend is similar in both *C. nagifolium* and *C. macrorhizon*. The variants per CDS length for genes involved in the GO terms of root development (GO:0048364) and root morphogenesis (GO:0010015) are higher than those in leaf development (GO:0048366), leaf morphogenesis (GO:0009965), flower development (GO:0009908), flower morphogenesis (GO:0048439), and average of all genes in both *C. nagifolium* and *C. macrorhizon* (Table 3-3). For *C. macrorhizon*, the variants per CDS length for each GO term belonging to biological process was highest in DNA replication preinitiation complex assembly (GO: 0071163), followed by regulation of DNA-templated DNA replication initiation (GO: 0030174), sesquiterpenoid biosynthetic process (GO: 0016106), alpha-amino acid biosynthetic process (GO: 1901607), and ent-kaurene oxidation to kaurenoic acid (GO: 0010241) (Table 3-4).

By comparing ω of each gene among the species, I found 602 and 525 genes which were under positive or negative selection, respectively, in *C. macrorhizon* (Table 3-5). This includes the gene such as *MDO1* and *CP33B* for the genes under positive or relaxed selection, and *EMB2219* and *NHX6* for those under negative selection. There were no enriched GO terms in these gene sets.

Discussion

Genome assembly and annotation

Recently, Sun *et al.* (2021), and Chung *et al.* (2021) assembled the *C. goeringii* genome as 4.1 Gb and 4.0 Gb, respectively. These were larger than my assemblage with 2.2 Gb. The fewer repetitive elements could be detected in my study with 1.8 Gb compared with those in Sun *et al.* (2021), 3.2 Gb; and Chung *et al.* (2021), 3.6 Gb. On the other hand, my assembled genome showed the highest completeness in BUSCO (my study, 96%; Sun *et al.*, 2021, 87%; Chung *et al.*, 2021, 88%), which suggests that my input of long read DNA is not sufficient to assemble the repetitive elements, but is enough to reconstruct the genes. The previous studies did not perform polishing by long read DNA (Sun *et al.*, 2021; Chung *et al.*, 2021), and this could lead to a higher error rate in the assembled genome, resulting in lower BUSCO prediction in those studies. The higher error rate in the assembly could also affect the gene prediction, thus these studies might misidentify some genes. The mapping rates of RNA sequences from each tissue was generally high, but the mapping rate of the sequences from the root was relatively low. This might be because the root tissues have more contaminants such as mycorrhizal fungi and other soil microorganisms, and RNA from these organisms occupied the output sequences (Suetsugu et al., 2017).

The higher evolutionary rates in *C. macrorhizon* genome

The number of SNPs and indels are greater in *C. nagifolium* than in *C. macrorhizon* for the most of genomic regions, however, in CDS regions, those were greater in *C. macrorhizon,* the almost FM species, than in *C. nagifolium,* the MX species. The previous genomic study in MH orchid showed that the substitution rates in CDS are higher in FM species (Li *et al.*, 2022). My results are congruent with this phenomenon, suggesting the presence of common mechanisms in MH plants that increase the mutation rate in genes. Theoretical model suggests that parasites should evolve faster than their hosts due to the host-parasite arms race (Haraguchi & Sasaki, 1996). Similar phenomena are found in the genomes of several parasitic plants in angiosperms (Bromham et al., 2013). Therefore, the arms race between the FM plants and its mycorrhizal fungi might affect the higher substitution rate in FM plants.

Genes that might involved in mycoheterotrophy

The genes such as ortholog of *CRRSP55* and *TPS03* showed greater number of variations in *C. macrorhizon* than in *C. nagifolium* if compared with *C. goeringii*. *CRRSP55* is involved in a karrikin response in *A. thaliana* (Nelson *et al.*, 2010), and karrikin is known to be involved in establishment of mycorrhizal fungi in rice (Gutjahr *et al.*, 2015). *TPS03* is a gene involved in jasmonic acid response and sesquiterpenoid biosynthesis in *A. thaliana,* and controlled by rhizobacterial colonization (Pangesti *et al.*, 2015). Jasmonic acid and sesquiterpenoid such as strigolactone are involved in mycorrhizal symbiosis in several plants including MX and FM orchid (Hause & Schaarschmidt, 2009; Yuan et al., 2018; Chagi et al., in press). The changes in these genes which should be involved in mycorrhizal symbiosis might reflect the fact that *C.*

macrorhizon has specific mycorrhizal components compared to other MH *Cymbidium* species (Ogura-Tsujita *et al.*, 2012). The GO terms such as sesquiterpenoid biosynthetic process (GO: 0016106) and ent-kaurene oxidation to kaurenoic acid (GO: 0010241) had greater variants than other GOs in both *C. nagifolium* and *C. macrorhizon*. Sesquiterpenoids particularly strigolactone induces the branching of hyphae of mycorrhizal fungi and promotes mycorrhizal symbiosis (Akiyama & Hayashi, 2006). Ent-kaurene oxidase is involved in gibberellin synthesis, and upregulated in the AM infected root of *Lotus japonicus* (Takeda *et al.*, 2015). *Cymbidium macrorhizon* has relationships only with ectomycorrhizal fungi, and *C. goeringii* can symbiose with both ectomycorrhizal and saprophytic fungi, while *C. nagifolium* can symbiose both types of fungi but biased toward ectomycorrhizal fungi (Ogura-Tsujita *et al.*, 2012). The greater variants in these GOs in both *C. nagifolium* and *C. macrorhizon* suggest that the species composition of the mycorrhizal fungi could be regulated by strigolactone and gibberellin. The gene that might affect the morphological characters related to FM also had greater variations in *C. macrorhizon* than in *C. nagifolium*. *HCT* affect auxin transport, and silencing it results in stunted growth in *A. thaliana* and *Nicotiana benthamiana* (Hoffmann *et al.*, 2004; Kriegshauser *et al.*, 2021). Auxin transportation related genes are contracted in some FM orchids and may be related to the leafless morphology of them (Sun *et al.*, 2021). Hence, *HCT* might affect the leafless morphology in *C. macrorhizon*.

When the number of variations were compared among the genes expressed in different plant organs, the greater variation was observed in genes expressing only in roots than for those expressed only in stems and leaves in both *C. macrorhizon* and

C. nagifolium. The number of variations was also greater for genes involved in root morphogenesis and development than for genes involved in leaf or flower morphogenesis and development. In other words, in *C. macrorhizon*, many mutations occur in genes that are involved in roots, while fewer mutations occur in genes involved in leaves despite the absence of them. This indicates that genes involved in leaf formation may have to be conservative while those in roots are not. Because leaves and flowers share genes that are necessary for their development (Ditta *et al*., 2004), genes involved in leaf might have to be stable so as not to affect the floral morphology. The greater variants in genes involved in root might reflect the changes in mycorrhizal compositions in *C. macrorhizon* and *C. nagifolium* (Ogura-Tsujita *et al*., 2012)

The several genes that are located in chloroplast such as *CP33B* showed high ω (> 1) especially in *C. macrorhizon*. Since *C. macrorhizon* lacks leaves and perform reduced photosynthesis (Kobayashi *et al*., 2021), selection might be relaxed for photosynthesis related genes. The non-synonymous substitution in these genes might affect the reduced efficiency in photosynthesis of *C. macrorhizon* (Kobayashi *et al*., 2021). The genes involved in shoot formation, such as Leucine-rich repeat protein kinase family protein and *MDO1* also had higher ω (> 1) in *C. macrorhizon*. These genes might affect the lack of vegetative shoot and rhizome formation in *C. macrorhizon*, thus under the positive or relaxed selections. On the other hand, genes with ω smaller than 1 include the gene such as *HESP*, *EMB2219* and *NHX6*. These genes are essential for regulation of circadian rhythm, seed formations and regulation of intracellular pH (Delis et al., 2016; Shevtsov et al., 2018; Reguera et al., 2015), respectively, thus it might be lethal if changes occur in those genes. This might be the

reason that these genes are under negative selection in *C. macrorhizon*. There are no enriched GOs in both gene sets with $\omega > 1$ and $\omega < 1$ in *C. macrorhizon*, which suggests that at least for the early stages of the evolution of FM, various functions may be evolving in a coordinated manner rather than an intensive evolution of specific functions.

In conclusion, by sequencing and analysing the genomes of two MX and one FM orchids, *C. goeringii*, *C. nagifolium*, and *C. macrorhizon*, I reveal the potential molecular basis underlying the leafless morphology in FM species, and also, providing insights into the evolution of mycoheterotrophic plants.

Acknowledgements

This work is supported in part by the Ichimura Foundation for New Technology and the Sasakawa Scientific Research Grant [2021-5035]. Computations were performed on the NIG supercomputer at the ROIS National Institute of Genetics. I thank Dr. Hiroshi Yura for supporting the sampling at the Ecological Garden of the Natural History Museum and Institute, Chiba. I also thank Dr. Yu Okamura for his advice on Nanopore sequencing and other genome analysis techniques.

References

Ai, Y., Li, Z., Sun, W. H., Chen, J., Zhang, D., Ma, L., ... & Liu, Z. J. (2021). The *Cymbidium* genome reveals the evolution of unique morphological traits. Horticulture research, 8.

Akiyama, K., Hayashi, H. (2006). Strigolactones: chemical signals for fungal symbionts and parasitic weeds in plant roots. Ann Bot, 97(6):925-31.

Bromham, L., Cowman, P. F., & Lanfear, R. (2013). Parasitic plants have increased rates of molecular evolution across all three genomes. BMC evolutionary biology, 13, 1-11.

Brůna, T., Hoff, K. J., Lomsadze, A., Stanke, M., & Borodovsky, M. (2021). BRAKER2: automatic eukaryotic genome annotation with GeneMark-EP+ and AUGUSTUS supported by a protein database. NAR genomics and bioinformatics, 3(1), lqaa108.

Cingolani, P., Platts, A., Wang, L. L., Coon, M., Nguyen, T., Wang, L., ... & Ruden, D. M. (2012). A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of Drosophila melanogaster strain w1118; iso-2; iso-3. fly, 6(2), 80-92.

Delis, C., Krokida, A., Tomatsidou, A., Tsikou, D., Beta, R. A., Tsioumpekou, M., ... & Papadopoulou, K. K. (2016). AtHESPERIN: a novel regulator of circadian rhythms with poly (A)-degrading activity in plants. RNA biology, 13(1), 68-82.

Ditta, G., Pinyopich, A., Robles, P., Pelaz, S., & Yanofsky, M. F. (2004). The SEP4 gene of Arabidopsis thaliana functions in floral organ and meristem identity. Current biology, 14(21), 1935-1940.

Dobin, A., Davis, C. A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., ... & Gingeras, T. R. (2013). STAR: ultrafast universal RNA-seq aligner. Bioinformatics, 29(1), 15-21.

Gutjahr, C., Gobbato, E., Choi, J., Riemann, M., Johnston, M. G., Summers, W., ... & Paszkowski, U. (2015). Rice perception of symbiotic arbuscular mycorrhizal fungi requires the karrikin receptor complex. Science, 350(6267), 1521-1524.

Haraguchi, Y., & Sasaki, A. (1996). Host–parasite arms race in mutation modifications: indefinite escalation despite a heavy load?. Journal of theoretical biology, 183(2), 121-137.

Hoffmann, L., Besseau, S., Geoffroy, P., Ritzenthaler, C., Meyer, D., Lapierre, C., ... & Legrand, M. (2004). Silencing of hydroxycinnamoyl-coenzyme A shikimate/quinate hydroxycinnamoyltransferase affects phenylpropanoid biosynthesis. The Plant Cell, 16(6), 1446-1465.

Jacquemyn, H., & Merckx, V. S. (2019). Mycorrhizal symbioses and the evolution of trophic modes in plants. Journal of Ecology, 107(4), 1567-1581.

Jullien M. Flynn, Robert Hubley, Clément Goubert, Jeb Rosen, Andrew G. Clark, Cédric Feschotte, and Arian F. Smit. RepeatModeler2 for automated genomic discovery of transposable element families. PNAS April 28, 2020 117 (17) 9451-9457; first published April 16, 2020.

Reguera, M., Bassil, E., Tajima, H., Wimmer, M., Chanoca, A., Otegui, M. S., ... & Blumwald, E. (2015). pH regulation by NHX-type antiporters is required for receptor-

mediated protein trafficking to the vacuole in Arabidopsis. The Plant Cell, 27(4), 1200-1217.

Kriegshauser, L., Knosp, S., Grienenberger, E., Tatsumi, K., Gütle, D. D., Sørensen, I., ... & Renault, H. (2021). Function of the HYDROXYCINNAMOYL-CoA: SHIKIMATE HYDROXYCINNAMOYL TRANSFERASE is evolutionarily conserved in embryophytes. The Plant Cell, 33(5), 1472-1491.

Leake, J. R. (1994). The biology of myco‐heterotrophic ('saprophytic') plants. New Phytologist, 127(2), 171-216.

Li, H. (2018). Minimap2: pairwise alignment for nucleotide sequences. Bioinformatics, 34(18), 3094-3100.

Li, M. H., Liu, K. W., Li, Z., Lu, H. C., Ye, Q. L., Zhang, D., ... & Liu, Z. J. (2022). Genomes of leafy and leafless Platanthera orchids illuminate the evolution of mycoheterotrophy. Nature Plants, 8(4), 373-388.

Manni, M., Berkeley, M. R., Seppey, M., & Zdobnov, E. M. (2021). BUSCO: assessing genomic data quality and beyond. Current Protocols, 1(12), e323.

Merckx, V. S., Freudenstein, J. V., Kissling, J., Christenhusz, M. J., Stotler, R. E., Crandall-Stotler, B., ... & Maas, P. J. (2013). Taxonomy and classification. Mycoheterotrophy: the biology of plants living on fungi, 19-101.

Nelson, D. C., Flematti, G. R., Riseborough, J. A., Ghisalberti, E. L., Dixon, K. W., & Smith, S. M. (2010). Karrikins enhance light responses during germination and seedling development in Arabidopsis thaliana. Proceedings of the National Academy of Sciences, 107(15), 7095-7100.

Nishimura O, Hara Y, Kuraku S., gVolante for standardizing completeness assessment of genome and transcriptome assemblies Bioinformatics. 2017; 33(22):3635-3637. doi: 10.1093/bioinformatics/btx445.

Ogura‐Tsujita, Y., Yokoyama, J., Miyoshi, K., & Yukawa, T. (2012). Shifts in mycorrhizal fungi during the evolution of autotrophy to mycoheterotrophy in Cymbidium (Orchidaceae). American Journal of Botany, 99(7), 1158-1176.

Pangesti, N., Weldegergis, B. T., Langendorf, B., van Loon, J. J., Dicke, M., & Pineda, A. (2015). Rhizobacterial colonization of roots modulates plant volatile emission and enhances the attraction of a parasitoid wasp to host-infested plants. Oecologia, 178(4), 1169-1180.

Poplin, R., Chang, P. C., Alexander, D., Schwartz, S., Colthurst, T., Ku, A., ... & DePristo, M. A. (2018). A universal SNP and small-indel variant caller using deep neural networks. Nature biotechnology, 36(10), 983-987.

Roach, M. J., Schmidt, S. A., & Borneman, A. R. (2018). Purge Haplotigs: allelic contig reassignment for third-gen diploid genome assemblies. BMC bioinformatics, 19(1), 1-10.

Selosse MA, Roy M. Green plants that feed on fungi: facts and questions about mixotrophy. Trends in plant science, 2009;14(2):64-70.

Shevtsov, S., Nevo-Dinur, K., Faigon, L., Sultan, L. D., Zmudjak, M., Markovits, M., & Ostersetzer-Biran, O. (2018). Control of organelle gene expression by the mitochondrial transcription termination factor mTERF22 in Arabidopsis thaliana plants. PLoS One, 13(7), e0201631.

Suetsugu K, Yamato M, Miura C, Yamaguchi K, Takahashi K, Ida Y, ... & Kaminaka, H. Comparison of green and albino individuals of the partially mycoheterotrophic orchid Epipactis helleborine on molecular identities of mycorrhizal fungi, nutritional modes and gene expression in mycorrhizal roots. Molecular Ecology. 2017;26(6):1652-1669.

Sun, Y., Chen, G. Z., Huang, J., Liu, D. K., Xue, F., Chen, X. L., ... & Liu, Z. J. (2021). The Cymbidium goeringii genome provides insight into organ development and adaptive evolution in orchids. Ornamental Plant Research, 1(1), 1-13.

Takeda, N., Handa, Y., Tsuzuki, S., Kojima, M., Sakakibara, H., & Kawaguchi, M. (2015). Gibberellins interfere with symbiosis signaling and gene expression and alter colonization by arbuscular mycorrhizal fungi in *Lotus japonicus*. Plant physiology, 167(2), 545-557.

Vaser, R., Sović, I., Nagarajan, N., & Šikić, M. (2017). Fast and accurate de novo genome assembly from long uncorrected reads. Genome research, 27(5), 737-746.

Wang, D., Jacquemyn, H., Gomes, S. I., Vos, R. A., & Merckx, V. S. (2021). Symbiont switching and trophic mode shifts in Orchidaceae. New Phytologist, 231(2), 791-800.

Warren, R. L., Coombe, L., Mohamadi, H., Zhang, J., Jaquish, B., Isabel, N., ... & Birol, I. (2019). ntEdit: scalable genome sequence polishing. Bioinformatics, 35(21), 4430-4432.

Yang, Z. (2007). PAML 4: phylogenetic analysis by maximum likelihood. Molecular biology and evolution, 24(8), 1586-1591.

Yang, F. X., Gao, J., Wei, Y. L., Ren, R., Zhang, G. Q., Lu, C. Q., ... & Zhu, G. F. (2021). The genome of Cymbidium sinense revealed the evolution of orchid traits. Plant biotechnology journal, 19(12), 2501-2516.

Yuan, Y., Jin, X., Liu, J., Zhao, X., Zhou, J., Wang, X., ... & Huang, L. (2018). The *Gastrodia elata* genome provides insights into plant adaptation to heterotrophy. Nature Communications, 9(1), 1615.

Zhang, G. Q., Liu, K. W., Li, Z., Lohaus, R., Hsiao, Y. Y., Niu, S. C., ... & Liu, Z. J. (2017). The *Apostasia* genome and the evolution of orchids. Nature, 549(7672), 379-383.

Tables

Table 3-1. The list of genes that showed higher variation rate in *Cymbidium macrorhizon* compared to *C. nagifolium*. Only the top 25 genes are shown.

gene name	Orthologous gene in *Arabidopsis*	Gene name in *Arabidopsis*	difference in number of variants between *C. macrorhizon* and *C. nagifolium* per CDS length
gene_13560	ATMG00860.1	(*ORF158*)	0.204
gene_14815	AT5G48540.1	-	0.169
g20018	AT4G16740.1	*TERPENE SYNTHASE 03* (*TPS03*)	0.137
gene_8295	AT2G02520.1	-	0.110
gene_686	AT3G01410.1	-	0.109
gene_10738	ATMG00810.1	(*ORF240B*)	0.109
g42311	AT4G22758.1	-	0.108
g41774	AT5G48930.1	*HYDROXYCINNAMOYL-COA SHIKIMATE/QUINATE HYDROXYCINNAMOYL TRANSFERASE* (*HCT*)	0.103

g68439	AT5G01120.1	-	0.102
gene_4582	ATMG00810.1	(*ORF240B*)	0.099
g69709	AT5G08670.1	-	0.097
g52707	AT1G24706.2	(*THO2*)	0.096
g50427	AT4G16740.1	*TERPENE SYNTHASE 03* (*TPS03*)	0.096
g38524	AT5G01140.1	-	0.094
g6077	AT1G51090.1	(*ATHMAD1*)	0.088
g45852	AT1G08080.1	*ALPHA CARBONIC ANHYDRASE 7* (*ACA7*)	0.087
g41773	AT5G48930.1	*HYDROXYCINNAMOYL-COA SHIKIMATE/QUINATE HYDROXYCINNAMOYL TRANSFERASE* (*HCT*)	0.086
g6094	AT4G16380.1	*HEAVY METAL ASSOCIATED PROTEIN 35* (*ATHMP35*)	0.085
gene_12577	AT1G42190.1	-	0.083
g18094	AT5G12080.1	*MECHANOSENSITIVE CHANNEL OF SMALL CONDUCTANCE-LIKE 10* (*MSL10*)	0.083
gene_8333	AT4G36950.1	*MITOGEN-ACTIVATED PROTEIN KINASE KINASE KINASE 21* (*MAPKKK21*)	0.081
gene_6908	AT4G23160.1	*CYSTEINE-RICH RECEPTOR-LIKE PROTEIN KINASE 8* (*CRK8*)	0.080

gene_10160	AT1G42190.1	-	0.079
g56710	AT1G18750.1	*AGAMOUS-LIKE 65* (*AGL65*)	0.078
g71502	AT3G26430.1	(*GGL20*)	0.078

Table 3-2. Number of variants per exon length for the genes expressed in leaf, stem, and root specifically.

	Number of genes	*C. nagifolium*	*C. macrorhizon*
leaf only	2,047	14/Kbp	14/Kbp
stem only	1,576	13/Kbp	14/Kbp
root only	114	18/Kbp	18/Kbp
all organs	867	13/Kbp	14/Kbp

Table 3-3. Number of variants per exon length for the genes involved in each GO term related to the development of leaf, root and flower, respectively.

GO term	Number of genes	*C. nagifolium*	*C. macrorhizon*
Leaf development (GO:0048366)	217	16/Kbp	16/Kbp
Leaf morphogenesis (GO:0009965)	74	11/Kbp	12/Kbp
Root development (GO:0048364)	448	20/Kbp	19/Kbp
Root morphogenesis (GO:0010015)	278	19/Kbp	20/Kbp
Flower development (GO:0009908)	1,045	15/Kbp	14/Kbp
Flower morphogenesis (GO:0048439)	3	9/Kbp	9/Kbp
All genes	25,247	19/Kbp	19/Kbp

Table 3-4. Number of variants per exon length for the genes involved in each GO term, listed in order of the frequency of mutations in *C. macrorhizon*. Only GO terms with rate > 50/Kbp were shown.

GO term	gene number	*C. nagifolium*	*C. macrorhizon*
DNA replication preinitiation complex assembly (GO:0071163)	5	156/Kbp	178/Kbp
branched-chain amino acid biosynthetic process (GO:0009082)	18	94/Kbp	148/Kbp
regulation of DNA-dependent DNA replication initiation (GO:0030174)	6	127/Kbp	144/Kbp
sesquiterpenoid biosynthetic process (GO:0016106)	32	124/Kbp	133/Kbp
adenosine to inosine editing (GO:0006382)	7	100/Kbp	130/Kbp
valine biosynthetic process (GO:0009099)	26	88/Kbp	117/Kbp
alpha-amino acid biosynthetic process (GO:1901607)	7	116/Kbp	110/Kbp
isoleucine biosynthetic process (GO:0009097)	29	81/Kbp	109/Kbp
ent-kaurene oxidation to kaurenoic acid (GO:0010241)	7	101/Kbp	106/Kbp

benzoate metabolic process (GO:0018874)	3	63/Kbp	100/Kbp
allantoin transport (GO:0015720)	4	74/Kbp	95/Kbp
uracil transport (GO:0015857)	4	74/Kbp	95/Kbp
response to herbivore (GO:0080027)	44	96/Kbp	93/Kbp
DNA replication checkpoint signalling (GO:0000076)	12	78/Kbp	89/Kbp
FAD metabolic process (GO:0046443)	3	59/Kbp	88/Kbp
regulation of DNA double-strand break processing (GO:1903775)	6	67/Kbp	81/Kbp
regulation of histone H2B conserved C-terminal lysine ubiquitination (GO:2001173)	6	67/Kbp	81/Kbp
regulation of systemic acquired resistance (GO:0010112)	11	72/Kbp	80/Kbp
response to insect (GO:0009625)	41	72/Kbp	80/Kbp
para-aminobenzoic acid metabolic process (GO:0046482)	6	58/Kbp	77/Kbp
regulation of secondary metabolic process (GO:0043455)	15	80/Kbp	75/Kbp
cellular response to molecule of bacterial origin (GO:0071219)	12	67/Kbp	74/Kbp
meiotic joint molecule formation (GO:0000709)	3	52/Kbp	71/Kbp

meiotic strand invasion involved in reciprocal meiotic recombination (GO:0010774)	3	52/Kbp	71/Kbp
diterpenoid biosynthetic process (GO:0016102)	75	69/Kbp	70/Kbp
pyrimidine nucleobase salvage (GO:0043100)	5	56/Kbp	69/Kbp
negative regulation of flavonoid biosynthetic process (GO:0009964)	4	41/Kbp	68/Kbp
meiotic DNA integrity checkpoint signalling (GO:0044778)	4	61/Kbp	67/Kbp
regulation of salicylic acid biosynthetic process (GO:0080142)	19	58/Kbp	65/Kbp
negative regulation of MAPK cascade (GO:0043409)	2	56/Kbp	65/Kbp
plus-end directed microtubule sliding (GO:0031535)	5	44/Kbp	65/Kbp
vesicle transport along microtubule (GO:0047496)	5	44/Kbp	65/Kbp
cellular amino acid biosynthetic process (GO:0008652)	12	68/Kbp	64/Kbp
glucosinolate biosynthetic process (GO:0019761)	13	67/Kbp	62/Kbp
monoterpene biosynthetic process (GO:0043693)	6	36/Kbp	62/Kbp
developmental maturation (GO:0021700)	6	91/Kbp	61/Kbp
pH reduction (GO:0045851)	2	79/Kbp	61/Kbp

regulation of ion transport (GO:0043269)	1	52/Kbp	60/Kbp
regulation of small molecule metabolic process GO:0062012)	18	64/Kbp	60/Kbp
small nucleolar ribonucleoprotein complex assembly (GO:0000491)	2	39/Kbp	58/Kbp
magnesium ion transmembrane transport (GO:1903830)	20	53/Kbp	58/Kbp
chromosome organization (GO:0051276)	29	51/Kbp	57/Kbp
regulation of trichome patterning (GO:1900032)	2	35/Kbp	56/Kbp
regulation of proanthocyanidin biosynthetic process (GO:2000029)	12	25/Kbp	56/Kbp
plastid to vacuole vesicle-mediated transport (GO:1904962)	1	44/Kbp	55/Kbp
positive regulation of apoptotic process (GO:0043065)	1	62/Kbp	54/Kbp
thiamine metabolic process (GO:0006772)	10	42/Kbp	54/Kbp
negative regulation of cell differentiation (GO:0045596)	5	47/Kbp	54/Kbp
2,4,6-trinitrotoluene catabolic process (GO:0046256)	6	57/Kbp	53/Kbp
positive regulation of telomere maintenance (GO:0032206)	2	62/Kbp	53/Kbp
positive regulation of telomere maintenance via telomerase (GO:0032212)	2	62/Kbp	53/Kbp
programmed cell death in response to reactive oxygen species (GO:0097468)	4	91/Kbp	53/Kbp

monoterpenoid biosynthetic process (GO:0016099)	14	73/Kbp	53/Kbp
arginine catabolic process (GO:0006527)	3	49/Kbp	52/Kbp
medium-chain fatty acid metabolic process (GO:0051791)	5	60/Kbp	52/Kbp
regulation of fatty acid beta-oxidation (GO:0031998)	3	53/Kbp	52/Kbp
diacylglycerol catabolic process (GO:0046340)	1	63/Kbp	51/Kbp
monoacylglycerol catabolic process (GO:0052651)	1	63/Kbp	51/Kbp
regulation of cell diameter (GO:0060305)	1	54/Kbp	51/Kbp

Table 3-5. The list of genes with $\omega > 1$ or < 1 only in *C. macrorhizon*. Only the top 25 genes with $\omega > 1$ and < 1 were shown, respectively.

gene name	Orthologous gene in *Arabidopsis*	Gene name in *Arabidopsis*	ω value
gene_10356	-	-	> 1
gene_11623	AT2G35410	(*CP33B*)	> 1
gene_12111	AT1G76400	*OLIGOSACCHARYLTRANSFERASE 1B* (*OST1B*)	> 1
gene_12428	AT2G43130	(*ARA4*)	> 1
gene_13692	AT3G07890	-	> 1
gene_13905	AT1G05060	-	> 1
gene_14426	AT4G10850	(*SWEET7*)	> 1
gene_14502	AT4G17486	-	> 1
gene_15206	AT1G63430	-	> 1
gene_17030	AT1G56260	*MERISTEM DISORGANIZATION 1* (*MDO1*)	> 1
gene_17523	AT1G73910	*ACTIN-RELATED PROTEINS 4A* (*ARP4A*)	> 1
gene_18171	AT3G49780	*PHYTOSULFOKINE 4 PRECURSOR* (*PSK4*)	> 1
gene_18549	AT4G19710	*ASPARTATE KINASE-HOMOSERINE DEHYDROGENASE II* (*AK-HSDH II*)	> 1
gene_18712	AT2G37195	-	> 1
gene_21097	AT1G17530	*TRANSLOCASE OF INNER MITOCHONDRIAL MEMBRANE 23* (*TIM23-1*)	> 1
gene_21605	-	-	> 1
gene_21808	AT1G23190	*PHOSPHOGLUCOMUTASE 3* (*PGM3*)	> 1
gene_22636	AT1G51090	(*ATHMAD1*)	> 1

gene_22890	AT2G33510	(*CFL1*)	> 1
gene_23940	AT4G14723	*CHALLAH-LIKE 2* (*CLL2*)	> 1
gene_26418	AT3G59140	*ATP-BINDING CASSETTE C10* (*ABCC10*)	< 1
gene_20035	AT3G55340	*PHRAGMOPLASTIN INTERACTING PROTEIN 1* (*PHIP1*)	< 1
gene_9149	AT3G16340	*ATP-BINDING CASSETTE G29* (*ABCG29*)	< 1
gene_6439	AT2G39930	*ISOAMYLASE 1* (*ISA1*)	< 1
gene_13460	AT4G37000	*ACCELERATED CELL DEATH 2* (*ACD2*)	< 1
gene_27713	AT1G18390	*LEAF RUST 10 DISEASE-RESISTANCE LOCUS RECEPTOR- LIKE PROTEIN KINASE-LIKE 1.2* (*LRK10L1.2*)	< 1
gene_24330	AT1G18390	*LEAF RUST 10 DISEASE-RESISTANCE LOCUS RECEPTOR- LIKE PROTEIN KINASE-LIKE 1.2* (*LRK10L1.2*)	< 1
gene_386	AT5G62680	*NRT1/ PTR FAMILY 2.11* (*NPF2.11*)	< 1
gene_15008	AT4G35540	*POLLEN-EXPRESSED TRANSCRIPTION FACTOR 2* (*PTF2*)	< 1
gene_21197	AT5G60700	-	< 1
gene_15865	AT3G53480	*ATP-BINDING CASSETTE G37* (*ABCG37*)	< 1
gene_15630	AT4G00330	*CALMODULIN-BINDING RECEPTOR-LIKE CYTOPLASMIC KINASE 2* (*CRCK2*)	< 1
gene_19632	AT1G53840	*PECTIN METHYLESTERASE 1* (*PME1*)	< 1
gene_6417	AT1G04500	-	< 1
gene_491	AT4G12780	*AUXILIN-LIKE1* (*AUXILIN-LIKE1*)	< 1
gene_18457	AT3G55320	*ATP-BINDING CASSETTE B20* (*ABCB20*)	< 1
gene_6977	AT3G14360	*OIL BODY LIPASE 1* (*ATOBL1*)	< 1
gene_23675	AT2G17200	(*DSK2B*)	< 1
gene_5517	AT1G31500	*HESPERIN* (*HESP*)	< 1

gene_25836	AT4G01080	*TRICHOME BIREFRINGENCE-LIKE 26* (*TBL26*)	< 1

Table 3-S1. Summary of *Cymbidium goeringii* genome.

Genome size	2,186,357,068 bp
Contig number	17,709
Contig N50	2,266,437 bp
Mean coverage	22
BUSCO	Embryophyta C:89.4%[S:86.4%,D:3.0%],F:6.3%,M:4.3%,n: 1614; Liliopsida C:83.9%[S:81.2%,D:2.7%],F:9.5%,M:6.6%,n: 3236
Repetitive elements	1,783,444,476 bp
Gene number	27,753

Table 3-S2. Detailed statistics of number of variants in *Cymbidium* goeringii genome.

		C. nagifolium	*C. macrorhizon*
All region	SNPs	53,634,320	51,759,843
	Insertions	2,154,915	2,034,169
	Deletions	2,141,143	2,216,112
Coding region	SNPs	381,273	404,608
	Insertions	2,828	2,804
	Deletions	2,718	3,127

General discussion

Evolution of mycoheterotrophy in genus *Cymbidium*

The ancestral characteristic assessment in Chapter 1 indicates the evolutionary process of FM evolution within the genus *Cymbidium*. In the genus *Cephalanthera*, another genus that includes both MX and FM plants in Orchidaceae, there is also a discussion about the evolution of FM species from MX ancestors (Julou et al., 2005; Sakamoto et al., 2016; Tsukaya, 2018). As the hypothesis of the FM emergence, one theory posits that leafless individuals emerged from albinistic variants of MX species (Julou et al., 2005), while the other one suggests that they lost the capability for photosynthesis subsequent to losing their leaves (Sakamoto et al., 2016). The second scenario is plausible according to my results for genus *Cymbidium*. It has been suggested that the most recent common ancestor (MRCA) of genus *Cymbidium* were epiphytic in the symbiotic relationship with saprophytic "*Rhizoctonia*" (Fig. 1-4; Yukawa et al., 2002; Ogura-Tsujita et al., 2012). Thereafter, throughout the process of speciation, terrestrial species which evolved from epiphytic AT ancestors emerged four or five times independently within three distinct clades (Fig. 1-1, 1-2, 1- 3, 1-4). These terrestrial species had a mycorrhizal association with saprophytic "*Rhizoctonia*", however, among them, capability for symbiosis with ectomycorrhizal fungi were developed only in the MRCA of section *Jensoa* and section *Pachyrhizanthe* (Fig. 1-4; Ogura-Tsujita et al., 2012). Furthermore, certain lineages within section *Pachyrhizanthe* have undergone an evolutionary adaptation to increase their dependency on fungi and have emerged as FM species (Fig. 1-4), having exclusive relationships with ectomycorrhizal fungi and perform limited photosynthetic capabilities (Ogura-Tsujita et al., 2012; Suetsugu et al., 2018; Kobayashi et al., 2021). The changes in mycorrhizal partners during the evolution of FM were also found in several lineages (Jacquemyn & Merckx, 2019; Wang et al., 2021), suggesting that the mycorrhizal fungi with better ability to

provide C to the host plants are chosen for the partner of FM plants (Selosse et al., 2022; Wang et al., 2021). During this process, leaves have been degenerated. Some species within the section *Pachyrhizanthe* are deciduous during anthesis (Liu et al., 2006), thus it is possible that FM species that are entirely leafless, endogenous and only emerging when anthesis, have evolved from such species.

The comparisons between MX and FM *Cymbidium* genome in Chapter 3 give several insights into the genetic backgrounds of FM species. Among the shifts between MX and FM, many mutations occurred in genes involved in photosynthesis and leaf morphogenesis (Table 3-1; 3-4). In the genome of FM species *C. macrorhizon*, there are several genes related to photosynthesis and leaf morphogenesis with ω larger than 1 (Table 3-5), suggesting those genes are under positive or relaxed selection. This may be attributed to the morphology and physiology of FM features in *C. macrorhizon*, which have no leaves and perform only limited photosynthesis (Suetsugu et al., 2018; Kobayashi et al., 2021). The previous study that made hybrid individuals between MX and FM *Cymbidium* found unstable morphology of hybrids, while some individuals lack leaves, another bear leaves as it grows (Ogura-Tsujita et al., 2014). This suggests the possibility that the mutation which causes the loss of leaves does not appear to be caused by a mutation in a single or limited number of genes. It also suggests that not only the changes in coding regions, but also in the regulatory regions might have an impact on the leafless morphology of FM species. The study on the plastid genome of FM species, *C. macrorhizon,* found that the most plastid genes were apparently functional (Kim et al., 2018). However, I found several nucleic genes that code proteins located in plastids are under positive or relaxed selections (Table 3-5). It is known that *C. macrorhizon* performs incomplete photosynthesis (Kobayashi et al., 2021), and it is possible that the changes in these nucleic genes, not plastid genes, are affecting the process. Transcriptomic

analysis in mycorrhiza of *C. goeringii* and *C. macrorhizon* might provide deeper insights in this viewpoint.

Symbiosis between *Cymbidium* and its mycorrhizal fungi

The transcriptomic analysis in Chapter 2 suggests that the MX species *C. goeringii* may regulate its dependence on mycorrhizal fungi in response to changes in light conditions through jasmonic acid signaling (Table 2-2). This system is also present in AT plants (Nagata et al., 2015), implying that MX species, which rely on fungi for carbon, may control mycorrhizal symbiosis through a mechanism that is partially similar to that of autotrophic plants, which maintain a symbiotic relationship with fungi. Other phytohormones such as auxin and gibberellin are known to involved in the mycorrhizal relation in Orchid species (Miura et al., 2018), thus *Cymbidium* may also utilize these phytohormones to regulate the mycorrhizal symbiosis. Performing symbiotic germination might provide further insights into the regulation of mycorrhizal symbiosis in *Cymbidium*.

In genus *Cymbidium*, the mycorrhizal components between AT, MX, and FM species are different (Ogura-Tsujita et al., 2012). However, it is unclear if the system of controlling mycorrhizal symbiosis has changed during the transition of mycorrhizal partners. On the other hand, my genomic comparison between MX and FM species in Chapter 3 gives some insights to this point. Many mutations were found in genes that are expected to be expressed in the roots and involve in mycorrhizal symbiosis of *C. macrorhizon* and its sister species, *C. nagifolium* if they are compared to *C. goeringii* (Table 3-2), and these three species all have different mycorrhizal components with each other (Ogura-Tsujita et al., 2012). This suggests that there are changes in root systems in tandem with changes in their mycorrhizal components. Overall, my study suggests that the systems controlling the mycorrhizal symbiosis in MX and FM plants have not changed significantly

compared to autotrophic AT plants, but that minor adjustments may have been made to accommodate different mycorrhizal partners.

Comparing heterotrophic plants: Mycoheterotrophic plants and parasitic plants

In general, heterotrophic plants can be divided into two categories (Meerckx, 2013): parasitic plants that depend on other plants for their nutrients, and mycoheterotrophic plants that derive nutrients from fungi. Both parasitic plants and MH plants have undergone similar evolutionary changes, including the loss of leaves and photosynthetic ability (Leake, 1994; Westwood et al., 2010; Meerckx, 2013). These characteristics are attributed to their ability to rely on other organisms, reducing the need for photosynthesis (Westwood et al., 2010; Meerckx, 2013). Among parasitic plants, there exist holoparasitic plants which rely totally on their host plants, and hemiparasitic species which maintain their ability to photosynthesize while simultaneously relying on host plants (Westwood et al., 2010). Therefore, hemiparasitic plants within parasitic plants correspond to the presence of MX plants within MH plants (Meerckx, 2013). These phenomena suggest an occurrence of convergent evolution between parasitic plants and MH plants. Thus, a comparative examination of these two plants might shed light on the evolution of plants that rely on other organisms during their life.

The genomic studies in parasitic plants found that there is convergent gene loss in genes related to photosynthesis, organelle development, and nutrient uptake (Sun et al., 2018; Yoshida et al., 2019). Similar losses are also found in MH plants (Yuan et al., 2018; Xu et al., 2021; Li et al., 2022). My findings also suggest that the genes related to photosynthesis are under relaxed selection in FM orchid (Chapter 3). These convergent gene losses in both parasitic and MH plants reflect their property of not requiring photosynthesis (Xu et al., 2021). Moreover, the genome of parasitic plants show higher substitution rate compared to the related non-parasitic species

(Bromham et al., 2013; Yoshida et al., 2019). This phenomenon is comparable with the similar findings in the genomes of MH plants (Xu et al., 2021; Li et al., 2022). My findings in Chapter 3, the higher variation number in exon of FM species compared to MX species, are congruent with this phenomenon. Theoretical model derived that the higher evolution rates in parasites are needed as parasites should evolve faster than evolution of their hosts' counteracts (Haraguchi & Sasaki, 1996). Some research has substantiated this hypothesis in regards to parasitic plants (Bromham et al., 2013), but not in MH plants. By testing this hypothesis on MH plants and comparing those to parasitic plants, it may be possible to understand the universal mechanisms related to host-parasite arm-race and adjustment of evolutionary speeds.

In conclusion, through a comprehensive analysis of the genus *Cymbidium*, incorporating phylogenetic, transcriptomic, and genomic approaches, my study uncovered the genetic backgrounds of the evolutionary trajectory of the FM species within the genus *Cymbidium*. The FM species have evolved from their ancestral AT form through the intermediate MX species. During this evolutionary process, the mycorrhizal components, the mechanisms controlling its symbiotic relationship, the morphological development systems, and the photosynthetic capability all evolved in tandem.

Acknowledgements

I would like to appreciate Prof. Masashi Murakami (Chiba University) for his support on my work. I also like to thank Dr. Yuma Takahashi (Chiba University) for his kind help and suggestion on my work.

I thank Dr. Tomohisa Yukawa (National Museum of Nature and Science, Tokyo) for his support for the phylogenetic analysis, valuable comments on my study, and helping orchid sampling in Tsukuba Botanical Garden for my work in Chapter 1. I thank Dr. Hiroshi Yura (Natural History Museum and Institute, Chiba.) for supporting the sampling at the Ecological Garden of the Natural History Museum and Institute, Chiba. I also thank Dr. Yu Okamura (University of Tokyo) especially for his advice on Nanopore sequencing and other genome analysis techniques for my work in Chapter 3, and for his kind suggestion on my work.

This work is supported in part by the Ichimura Foundation for New Technology and the Sasakawa Scientific Research Grant [2021-5035]. Computations were performed on the NIG supercomputer at the ROIS National Institute of Genetics.

References

Bromham, L., Cowman, P. F., & Lanfear, R. (2013). Parasitic plants have increased rates of molecular evolution across all three genomes. BMC evolutionary biology, 13, 1-11.

Cai, L., Arnold, B. J., Xi, Z., Khost, D. E., Patel, N., Hartmann, C. B., ... & Davis, C. C. (2021). Deeply altered genome architecture in the endoparasitic flowering plant *Sapria himalayana* Griff.(Rafflesiaceae). Current Biology, 31(5), 1002-1011.

Jacquemyn, H., & Merckx, V. S. (2019). Mycorrhizal symbioses and the evolution of trophic modes in plants. Journal of Ecology, 107(4), 1567-1581.

Julou, T., Burghardt, B., Gebauer, G., Berveiller, D., Damesin, C., & Selosse, M. A. (2005). Mixotrophy in orchids: insights from a comparative study of green individuals and nonphotosynthetic individuals of *Cephalanthera damasonium*. New Phytologist, 166(2), 639-653.

Kim, H. T., Shin, C. H., Sun, H., & Kim, J. H. (2018). Sequencing of the plastome in the leafless green mycoheterotroph *Cymbidium macrorhizon* helps us to understand an early stage of fully mycoheterotrophic plastome structure. Plant Systematics and Evolution, 304, 245-258.

Kobayashi, K., Suetsugu, K., & Wada, H. (2021). The leafless orchid *Cymbidium macrorhizon* performs photosynthesis in the pericarp during the fruiting season. Plant and Cell Physiology, 62(3), 472-481.

Leake, J. R. (1994). The biology of myco - heterotrophic ('saprophytic') plants. New Phytologist, 127(2), 171-216.

Li, M. H., Liu, K. W., Li, Z., Lu, H. C., Ye, Q. L., Zhang, D., ... & Liu, Z. J. (2022). Genomes of leafy and leafless *Platanthera* orchids illuminate the evolution of mycoheterotrophy. Nature Plants, 8(4), 373-388.

Liu, Z.J., Chen, S.C., & Ru, Z.Z. (2006). The genus *Cymbidium* in China. Science Press, Beijing.

Merckx, V. S. (2013). Mycoheterotrophy: an introduction. Mycoheterotrophy: The biology of plants living on fungi, 1-17.

Miura C, Yamaguchi K, Miyahara R, Yamamoto T, Fuj, M, Yagame T, Imaizumi-Anraku H, Yamato M, Shigenobu S, Kaminaka H. The mycoheterotrophic symbiosis between orchids and mycorrhizal fungi possesses major components shared with mutualistic plant-mycorrhizal symbioses. Molecular Plant-Microbe Interactions. 2018;31(10):1032-1047.

Ogura-Tsujita, Y., Miyoshi, K., Tsutsumi, C., & Yukawa, T. (2014). First flowering hybrid between autotrophic and mycoheterotrophic plant species: breakthrough in molecular biology of mycoheterotrophy. Journal of plant research, 127, 299-305.

Ogura - Tsujita, Y., Yokoyama, J., Miyoshi, K., & Yukawa, T. (2012). Shifts in mycorrhizal fungi during the evolution of autotrophy to mycoheterotrophy in *Cymbidium* (Orchidaceae). American Journal of Botany, 99(7), 1158-1176.

Sakamoto, Y., Ogura-Tsujita, Y., Ito, K., Suetsugu, K., Yokoyama, J., Yamazaki, J., ... & Maki, M. (2016). The tiny-leaved orchid *Cephalanthera subaphylla* obtains most of its carbon via mycoheterotrophy. Journal of plant research, 129, 1013-1020

Selosse, M. A., Petrolli, R., Mujica, M. I., Laurent, L., Perez-Lamarque, B., Figura, T., ... & Martos, F. (2022). The Waiting Room Hypothesis revisited by orchids: were orchid mycorrhizal fungi recruited among root endophytes?. Annals of Botany, 129(3), 259-270.

Suetsugu, K., Ohta, T., & Tayasu, I. (2018). Partial mycoheterotrophy in the leafless orchid *Cymbidium macrorhizon*. American journal of botany, 105(9), 1595-1600.

Sun, G., Xu, Y., Liu, H., Sun, T., Zhang, J., Hettenhausen, C., ... & Wu, J. (2018). Large-scale gene losses underlie the genome evolution of parasitic plant *Cuscuta australis*. Nature communications, 9(1), 2683.

Tsukaya, H. (2018). How leaves of mycoheterotrophic plants evolved–from the view point of a developmental biologist. New Phytologist, 217(4), 1401-1406.

Wang, D., Jacquemyn, H., Gomes, S. I., Vos, R. A., & Merckx, V. S. (2021). Symbiont switching and trophic mode shifts in Orchidaceae. New Phytologist, 231(2), 791-800.

Westwood, J. H., Yoder, J. I., Timko, M. P., & Depamphilis, C. W. (2010). The evolution of parasitism in plants. Trends in plant science, 15(4), 227-235.

Xu, Y., Lei, Y., Su, Z., Zhao, M., Zhang, J., Shen, G., ... & Wu, J. (2021). A chromosome - scale *Gastrodia elata* genome and large - scale comparative genomic analysis indicate convergent evolution by gene loss in mycoheterotrophic and parasitic plants. The Plant Journal, 108(6), 1609-1623.

Yoshida, S., Kim, S., Wafula, E. K., Tanskanen, J., Kim, Y. M., Honaas, L., ... & Shirasu, K. (2019). Genome sequence of *Striga asiatica* provides insight into the evolution of plant parasitism. Current Biology, 29(18), 3041-3052.

Yuan, Y., Jin, X., Liu, J., Zhao, X., Zhou, J., Wang, X., ... & Huang, L. (2018). The *Gastrodia elata* genome provides insights into plant adaptation to heterotrophy. Nature Communications, 9(1), 1615.