Serna-Sánchez et al. Phylogenomics of Cymbidieae and Orchidoideae: Solving phylogenetic ambiguities

Phylogenomics of the plant family Orchidaceae: Solving phylogenetic ambiguities within Cymbidieae and Orchidoideae

Maria Alejandra Serna-Sánchez[a,b], Astrid Catalina Alvarez-Yela[c], Juliana Arcila[a] and Tatiana Arias[a,d,*]

[a] Laboratorio de Biología Comparativa. Corporación para Investigaciones Biológicas (CIB), Cra. 72 A No. 78 B 141, Medellín, Colombia.

[b] Biodiversity, Evolution and Conservation. EAFIT University, Cra. 49, No. 7 sur 50, Medellín, Colombia

[c] Centro de Bioinformática y Biología Computacional (BIOS). Ecoparque Los Yarumos Edificio BIOS, Manizales, Colombia.
[d] Sociedad Colombiana de Orquideología (SCO), Cra. 52 N°73 – 298, Medellín, Colombia
[*] Corresponding Author: T.A.: Corporación para Investigaciones Biológicas, Cra. 72 A No. 78 B 141, Medellín, Colombia. E-mail: tarias@cib.org.co

All data have been deposited in Bioproject (XXXXXXX) and SRA (XXXXXXX, Appendix 1).

ABSTRACT

Phylogenomics for the Orchidaceae backbone have solved most evolutionary relationships among subfamilies, tribes and subtribes but some relationships remain unclear among the tribe Cymbidieae, and the Orchidoideae subfamily. Additionally, two of ten subtribes within Cymbidieae are still without representation. Our goals are to solve some recalcitrant phylogenetic relationships focusing on the Cymbidieae subtribes Stanhopeinae, Maxillariinae Zygopetalinae, Eulophiinae, Catasetinae and Cyrtopodiinae, and to provide stronger bootstrap supports for Codonorchideae within Orchidoideae subfamily. New taxa that have not been previously included in the Orchidaceae phylogeny have been added to this family using chloroplasts data recovered through Genome Survey Sequencing. Eleven orchids genomes were sequenced. Cleaning, normalization and assembly steps were performed to get chloroplasts, with sizes ranging from 132.712 bp to 161.827 bp. Eighty-six whole chloroplasts reported in GenBank were included in a global phylogenetic analysis. New plastid genomes showed similarities in structure, gene order and content (except for *ndh* and transfer RNA genes), but differences in organization of inverted repeat/small single-copy junction were observed. There were expansions or contractions of inverted repeat regions flanking the small single-copy region within Orchidaceae chloroplasts. Phylogenies obtained using whole chloroplasts data confirmed a relationship between sister subtribes Stanhopeinae and Maxillariinae, both sister to Oncidiinae, which is also sister to Eulophiinae. Cymbidiinae is the outermost sister clade to the rest of the Cymbidieae tribe. Our results showed important differences in evolutionary relationships among the Cymbidieae subtribes, with respect to previous studies using 60 coding regions from the chloroplast and the phylogeny obtained by Givnish et al. (2015). The earliest divergent clade in the tribe Cymbideae is Cymbidiinae. Catasetinae + Cyrtopodiinae are sister subtribes to Eulophiinae and Stanhopeinae + Zygopetalinae are sister groups to Maxillariinae + Oncidiinae. For the Orchidoideae, we found Codonorchideae + Orchideae as a strongly supported clade. Newly sequenced chloroplasts are positioned according to previous circumscriptions: genera *Scaphosepalum* and *Teagueia* are placed in Pleurothallidineae, *Sobralia* is located in Sobralieae tribe, and *Goodyera* is found in Goodyerinae within the Cranichideae tribe.

Key words: Cymbidieae, Orchidaceae, Orchidoideae, Phylogenomics, Plant Molecular Biology, Whole Chloroplast Genome

1. Introduction

Orchidaceae, the most diverse and widely distributed flowering family on earth has captivated botanists for a long time. The family has an impressive floral morphology diversity and multiple interactions with fungi, animal and plants (Fay and Chase, 2009). Orchidaceae consists of ca. 25,000 species and 880 genera (Givnish et al., 2015). Many research efforts have been made to understand natural history, evolution and evolutionary relationships among members of this family (Dong et al., 2018; Givnish et al., 2015; Luo et al., 2014; Niu et al., 2017; Pérez-Escobar et al., 2017). Today there are available four nuclear genomes (Chang et al., 2006; Huang et al., 2016; Yuan et al., 2018; Zhang et al., 2017), 287 complete chloroplast genomes and 1,639 Sequence Read Archives for Orchidaceae at NCBI. Phylogenomics have been implemented to solve main relationships among the orchids backbone (Dong et al., 2018; Givnish et al., 2015; Luo et al., 2014; Niu et al., 2017), nevertheless some questions remain unclear. Givnish et al. (2015) published a first well-supported backbone for the Orchidaceae family based on phylogenomics. They used 75 genes of plastid genomes and 39 orchid species to perform a Maximum Likelihood (ML) analysis covering 22 subtribes, 18 tribes and five subfamilies. Recent orchids studies agree on most of the evolutionary relationships between members of subfamilies and tribes (Luo 2014, Niu et al 2017, Dong et al, 2018). However, some relationships remain problematic to solve, especially in terms of subtribes and genera among Orchidaceae. Few relationships within Cymbidieae are not well supported, some subtribes have not been sampled yet and the position of Codonorchideae within the subfamily Orchidoideae is still unclear.

Tribe Cymbidieae comprises around 145 genera and 3,800 species (Chase et al., 2015), 90% occurring in the Neotropics with variation in pollination syndromes (e.i. male euglossine bees and oil bees) and sexual systems (e.i protandry, unisexuality and cleistogamy) (Pérez-Escobar et al., 2017). Cymbidieae belongs to the most diverse and highly speciose subfamily of orchids Epidendroidea and includes ten subtribes, four of which are the most speciose and abundant subclades in the Andean region (Maxillariinae, Oncidiinae, Stanhopeinae and Zygopetaliinae) (Pridgeon, 2009). Taxonomic circumscriptions within Cymbidieae have constantly changed (Dressler, 1993; Whitten et al., 2014), but the latest and more accepted genera and species delimitation among subtribes have been proposed by Chase et al. (2015): Catasetinae (8 genera, 354 species), Coeliopsidinae (3 genera, 19 species), Cymbidiinae (6 genera, 96 species), Cyrtopodiinae (1 genus, 47 species), Eriopsidiinae (1 genus, 5 species), Eulophiinae (13 genera, 302 species), Maxillariinae (12 genera, 819 species), Oncidiinae (65 genera, 1615 species), Stanhopeinae (20 genera, 304 species) and Zygopetalinae (36 genera, 437 species).

Several phylogenies have been generated by morphological and molecular analysis in order to solve relationships within Cymbidieae (Li et al., 2016; Pérez-Escobar et al., 2017; Whitten, Neubig, & Williams, 2014). Relationships among subtribes have recently been inferred using chloroplast coding genes *psaB*, *rbcL*, *matK*, *ycf1* combined with nuclear DNA *Xdh* (Li et al., 2016). In this study, the proposed phylogeny placed Cymbidiinae as the outermost sister clade to the rest of the Cymbidieae tribe. However poor support and incongruent topologies were found among Catasetinae, Eulophiinae and Eriopsidinae subtribes with respect to the topologies obtained by Whitten et al. (2014), Freudenstein & Chase (2015) and Pérez-Escobar et al. (2017). In these phylogenies Eulophiinae and Catasetinae formed a clade. Also, Eriopsidinae was not clearly placed in the results obtained by Li et al. (2016), but it was strongly-supported as the sister group of (Maxillariinae(Stanhopeinae(Coeliopsidinae))) in Freudenstein & Chase (2015) and Pérez-Escobar et al., (2017). In Li et al. (2016), Cyrtopodiinae appears as the second outermost group differing from the topology obtained in Givnish et al. (2015), in which Cyrtopodiinae is clustered with Catasetinae.

Chloroplast genomes of seed plants have been useful for phylogenetic analysis and evolutionary studies. It is due to their inheritance, which in most cases is uniparental, their high gene content, and their slow evolutionary rate in comparison with nuclear genomes (Niu et al., 2017; Smith, 2015). Orchid phylogenomics using the most complete taxonomic sampling to date (Givnish et al., 2015) include eight of ten subtribes belonging to Cymbidieae, but some subtribal relationships are still unresolved: Stanhopeinae (20 genera), Maxillariinae (12 genera), Zygopetalinae (36 genera), Oncidiinae (65 genera) and Eulophiinae (13 genera). A clade formed by Stanhopeinae and Maxillariinae had poor statistical support (BS=62) and their relationship with respect to Zygopetalineae had a moderately support (BS=72). Relationship between sister clades Eulophiinae and a clade containing Stanhopeinae, Maxillariinae, Zygopetalinae, and Oncidiinae also had poor support (BS=42).

Phylogenies are an important evolutionary framework to analyze diversity of pollination syndromes that have been reported in Cymbideae (Cingel, 2001; Pridgeon, 2009). It ranges from species exclusively pollinated by male Euglossini bee (Ramirez et al., 2010) to those pollinated only by oil bees (Pansarin et al., 2009; Renner and Schaefer, 2010).

The subfamily Orchidoideae comprises more than 3,600 species of terrestrial habits, which are present in all continents except by Antarctic. This group contains orchids with a single stamen (monandrous), and with a fertile anther that is erect and basitonic. Circumscriptions proposed by Chase et al. (2015), include Codonorchideae (1 genus, 1 species), Cranichideae (8 subtribes, 98 genera, 1824 species), Diurideae (9 subtribes, 40 genera, 979 species), Orchideae (4 subtribes, 59 genera, 2127 species). Salazar et al. (2003) have used plastid genes *rbcL* and *matK*, the *trnL-F* and nuclear ribosomal ITS regions, in order to check monophyly and subtribal delimitation of Cranichideae. In this research, the most-parsimonious tree of combined analysis of *rbcL*, *matK-trnK*, *trnL-trnF* and ITS sequences showed that Diurideae + Cranichideae, was the sister clade of Codonorchideae, and that Orchideae was the outermost clade within Orchidoideae. Additionally, putative synapomorphies for various clades were identified in their study. Orchideae main synapomorphy is the presence of sectile pollinia, while the rest of Orchidoideae have granular pollinia and apical viscidium (except by Goodyerinae, which seems to had a reversion to a sectile pollinia). A synapomorphy shared by Diurideae and Cranichideae is the presence of binary/bilobed xylem in leaf midrib. The absence of tubers is only common in Cranichideae. Although these synapomorphies were identified against molecular phylogenies, authors emphasized on inadequate previous interpretations of the characters due to the discrepancies generated between the well-supported relationships of the groups they were studying and the current classifications based on morphological characters. For Orchidoideae, Givnish et al. (2015) included four out of four tribes and six of 21 subtribes, but the relationship between Diurideae and Cranichideae was still poorly supported (BS=34) with respect to Codonorchideae.

Current genomic revolution is providing unrivaled knowledge about plant phylogenetics (Weitemier et al., 2014). Thanks to genomics now it is possible to collect huge amount of molecular information to reconstruct evolutionary histories, understand the state of conservation of populations and explore genetic diversity of many plants groups including orchids. Here we present 11 new orchid chloroplasts, which expand genera representation for the Orchidaceae family and clarify some relations within Cymbideae. Two general approaches were used: 1. Phylogenetic analysis using whole chloroplasts sequences and 2. Phylogenetic analysis using 60 coding regions. Phylogenetic analysis was performed using ML approach. The two different topologies reported here showed how reliable phylogenomics can be to infer relationships between members of a family as diverse as Orchidaceae.

2. Material and methods

2.1 Plant material, sampling and DNA isolation

Whole chloroplasts were sampled from 11 species representing three clades across Cymbidieae (Stanhopeinae, Maxillariinae and Oncidiinae) with five species, one clade across Epidendreae (Pleurothallidinae) with two species, three species within Sobralieae and one clade across Cranichideae (Goodyerinae) with one species. Fresh leaves were stored in silica gel for subsequent DNA extraction using the DNeasy Plant Mini Kit (Qiagen, Inc., Valencia, CA, USA), and a subsequent treatment with RNAase. DNA concentration and purity were initially measured by Nanodrop 2000 UV-Vis spectrophotometer (Thermo Scientific, Wilmington, DE, USA).

2.2 Sequencing on the Illumina platform

Next generation sequencing (NGS) was subcontracted with the Beijing Genomics Institute (BGI). BGI first determined the concentration of DNA using a Qubit 3.0 (Life Technologies® Carlsbad, California, EE.UU.) and evaluated the integrity of the DNA using agarose gel electrophoresis. DNA concentration was never less than 10 ng/mL. Purified genomic DNA (ratio $OD_{260/280}$ between 1.8 to 2.0) was break into smaller fragments of less than 800 bp using Bioruptor 200 (Cosmo Bio Co. Ltd, Tokyo, Japan). Fragment size was checked by electrophoresis; qualified products were purified with a DNA purification kit (QIAGEN). Overhangs fragments were blunt ended using T4 DNA polymerase, Klenow fragment and T4 polynucleotide kinase. Subsequently, a base 'A' was added to the 3 'end of the phosphorylated blunt DNA fragments, and final products were purified. DNA fragments were ligated to adapters, which have the overhang of the base 'T'. Ligation products were gel-purified by electrophoresis to remove all unbound adapters or split adapters that were ligated together. Ligation products were then selectively enriched and amplified by PCR. For each sample, more than 10 million reads of 90 bp were generated.

2.3 Genomes assemblages

A bioinformatic pipeline was developed with the most efficient tools for the analysis of sequences generated by the Genome Survey technique.

2.3.1 Sequences pre-processing

Raw sequences obtained by GSS were quality filtered using Trimmomatic (Bolger et al., 2014) in order to eliminate sequencing artifacts, improve uniformity in the reads content (size >40 pb) and ensure quality (>20) for further analysis. Filtered sequences were processed with BBNorm (Bushnell, 2017) to normalize coverage by down-sampling reads over high-depth areas of the genomes (maximum depth coverage 900x and minimum depth 6x). This step allows obtaining a flat coverage distribution to improve reads assembly quality. Subsequently, overlapping reads were merged into single reads using BBmerge (Bushnell et al., 2017) in order to accelerate the assembly process. Overlapping of paired reads was evaluated with *Flash* (Magoc and Salzberg, 2011) to reduce redundancy. Merged reads were used to carry out the *de novo* assembly with *SPAdes* (Hash length 33,55,77) (Bankevich et al., 2012).

2.3.2 Chloroplast assembly

Assembler *MIRA* 4 (Chevreux et al., 1999) was used to obtain whole chloroplast genomes. This program can map data against a consensus sequence of a reference assembly (simple mapping). *MIRA* has been useful for assembling complicated genomes with many repetitive sequences (Cock et al., 2013; Parakhia et al., 2014; Ward et al., 2012). Additionally, the program improves assemblies with an iterative extension of the reads or contigs based on additional information obtained by overlap of paired reads or by automatic corrections. *MIRA* reduces the number of reads in the Illumina mapping without sacrificing coverage information. The program tracks coverage with respect to each base in the reference and creates a sequence of synthetic length, with the coverage of equivalent reads (Coverage equivalent reads CER). Reads that do not map at a 100% remain as independent entities.

Consensus sequences were generated using *SAMTOOLS* mpileup (Li et al., 2009), which provides a summary of coverage of reads mapped to a reference sequence. In theory, it can call variants by mapping reads to an appropriate reference. For each of the 11 chloroplasts, phylogenetically closed chloroplasts (available in the NCBI) were used as reference (*Masdevallia picturata, Masdevallia coccinea, Cattleya crispata, Goodyera fumata, Oncidium sphacelatum, Sobralia callosa*). Resultant chloroplast genomes were cured using *Geneious* *version* 9 (created by Biomatters) mapping good sequenced quality reads to genomes and obtaining a consensus sequence per each chloroplast.

2.4 Chloroplast annotations

A search for other orchid chloroplasts was carried out through NCBI. Ninety-five chloroplasts from orchids and three from external groups (*Iris sanguinea, Agapanthus coddii* and *Asparagus officinalis*) were recovered. One hundred and six chloroplasts obtained (11 new chloroplasts and 95 from the NCBI) were annotated through the Chlorobox portal of the Max Planck Institute (Tillich et al., 2017). Sequences were uploaded as fasta files and running parameters were established as follow: *BLAST* protein search identity=65%, *BLAST rRNA*, *tRNA*, *DNA* search identity=85%, genetic code = Bacterial/Plant chloroplast, max intron length=3,000, options= allow overlaps. The species *Oncidium sphacelatum* was set as the 'Server Reference' and *Masdevallia coccinea* was set as the 'Custom Reference' for *CDS* and *tRNA*, *rRNA*, primer, other *DNA* or *RNA* specifications.

2.5 Phylogenetic analysis

2.5.1 Whole chloroplasts phylogenies

From the 106 chloroplasts obtained, 97 (11 new chloroplasts and 86 from the NCBI) were used as phylogenetic markers. These were aligned to find the best hypothesis of homology (Chan and Ragan, 2013) using *MAFFT* 7 (Katoh and Standley, 2013). This step was performed at the supercomputing center *APOLO*, EAFIT University, Medellín, Colombia. Phylogenetic reconstruction based on Maximum Likelihood was implemented in *RAxML* 8 (Stamatakis, 2014), using 1,000 bootstraps and under *GTRGAMMA* model.

2.5.2 Coding regions phylogenies

A set of 60 chloroplast genes was used to reconstruct phylogenetic relationships within Orchidaceae. One output of Chlorobox includes an alignment of each gene across all species. Alignments contain 97 or less species per coding region aligned (Table 1). Additional alignments were made to include genes used by Givnish *et al.* (2015), thus obtaining alignments of up to 127 species (Table 1). Each gene was manually checked for start and stop codons.

We conducted an additional ML analysis using 60 coding regions of species belonging to Cymbidieae and six outgroups across external Orchidaceae tribes (*Apostasia wallichii, Habenaria radiata, Neottia ovata, Sobralia callosa, Sobralia mucronata* and *Vanilla planifolia*) using *RAxML* 8 (Stamatakis, 2014) program. The same analysis was conducted for Orchidoideae subfamily, using 12 species of this group but deleting *Chloreae gavilu* and *Dactylorhiza fuchsii* sequences from Givnish et al., (2015), due to their poor quality. Here we used four outgroups (*Cypripedium japonicum*, *Paphiopedilum niveum*, *Vanilla planifolia* and *Apostasia wallichii*). Concatenation of these 60 genes for both, the Cymbidieae tribe and Orchidoideae subfamily, was made using *Geneious* 9. Concatenated protein-coding sequences for all taxa were aligned using *MAFFT* (Katoh and Standley, 2013) and polished.

3. Results

Here 11 new whole chloroplasts of orchids are included, accounting for 94 orchids sampled for the Orchidaceae family (Fig. 1). Five recognized subfamilies, 15 of 21 tribes, 18 of 23 subtribes and 29 genera are included in this study in order to increase taxa representation within the Orchidaceae backbone. From the 11 new chloroplasts reported, five chloroplasts belong to the tribe Cymbidieae, two to Epidendreae, three to Sobralieae and one to Cranichideae. Representativeness within Cymbidieae have increased from eight genera reported by Givnish et al., (2015) to 12 genera included in our work. Two new genera were included in Pleurothallidinae subtribe (Epidendreae). The subtribe Goodyerinae was included here as part of the tribe Cranichideae, with *Goodyera repens* as a new species.

New 11 plastids reported here have enlarged the whole family representation in terms of subtribes and genera. In most of the cases, the statistical support for their inclusion has also been improved. That is the case of *Telipogon glicensteinii*, *Maxillaria Sanderiana*, *Maxillaria nasuta*, *Gongora pleiochroma* and *Otoglossum globuliferum*, which have improved overall branch supports within Cymbidieae. Our results lead to a well-supported tribe phylogeny. New genera are included here: *Scaphosepalum antenniferum* and *Teagueia aliana*, contributing with two new chloroplast to the subtribe Pleurothallidinae, which also lead to high supported branches in the phylogeny (Fig. 2). New species within the tribe Sobralieae are also included in our study (*Sobralia decora*, *Sobralia mandonii* and *Sobralia mucronata*). The inclusion of the subtribe Goodyerinae within the Orchidoideae subfamily have improved statistical support (Figs. 2, 3 and 7).

3.1. Chloroplasts annotations

Whole chloroplasts sequences belonging to 97 species (11 sequenced here and 86 reported in NCBI) were annotated for 75 protein-coding genes. Five additional genes were recovered when concatenating this data matrix with the protein coding regions matrix used by Givnish et al. (2015), giving a total number of 80 genes for 124 orchid species and three outgroups. From these 80 genes, 20 were found to be problematic (*accD, ndhA, ndhB, ndhC, ndhD, ndhE, ndhF, ndhG, ndhH, ndhI, ndhJ, ndhK, petA, petB, petD, rpl16, rpoC1, rpoC2,* *rps12, ycf1*), thus they were not included in the final alignment, which had a final sequence length of 41,942 bp.

3.1.1 Molecular description of 11 chloroplast genomes

Complete chloroplast genomes of 11 new orchid species reported here have sizes ranging between 132,712 and 161,827 bp. The smaller chloroplast genome corresponds to *Maxillaria sanderiana* (132,712 bp) and the largest corresponds to *Sobralia mucronata* (161.827 bp) (Fig. 1). GC content was similar among all 11 chloroplasts and it ranges between 37 to 38,6%. The *M. sanderiana* chloroplast contains 123 different genes, of which 99 were single-copy and 24 were duplicated. Of these genes, 62 are protein-coding genes, four are *rRNA* genes and 33 are *tRNA* (Fig. 1 & Table 2). Chloroplast of *Telipogon glicensteinii* contains 113 genes (91 single-copy and 22 duplicated) coding for, 57 proteins, four *rRNAs* and 30 *tRNAs* (Table 2). All new chloroplasts reported here have *rRNA* genes (*rRNA*4.5, *rRNA*5, *rRNA*16S, *rRNA*23S) and approximately 13 *tRNA* genes are located in the inverted repeat regions (Fig. 1).

We identified that the *ndhF* gene is lost in five of the 11 new chloroplasts (*Gongora pleiochroma, Maxillaria nasuta, Maxillaria sanderiana, Otoglossum globuliferum* and *Telipogon glicensteinii*). The *trn* genes (*trnT-UGU, trnI-AAU, trnG-UCC*) were also commonly lost in the plastid genomes of seven of the 11 new orchid chloroplast. The chloroplast of *Sobralia mucronata* has all *trn* genes, *Maxillaria sanderiana* lacks *trnT-UGU* and *trnI-AAU*, and *Sobralia decora* and *Sobralia mandonii* lack just *trnG-UCC*. The gene *ndhK* is lost in *Gongora pleiochroma* and *Telipogon glicensteinii*. The chloroplast genome with more genes lost is *Telipogon glicensteinii*, which lacks *ndhC, ndhF, ndhJ, ndhK, trnT-UGU, trnI-AAU, trnG- UCC* and *trnL-CAG*. The 11 chloroplasts have portions of the genes *rpl22* and *ycf1* duplicated, contributing with the expansion and contraction among inverted regions flanking the small single-copy region of the chloroplast.

3.2. Phylogenomic analyses based on whole chloroplast

The phylogenomic analyses obtained using complete chloroplast strongly-supported previously published family-wide phylogenies (Givnish et al., 2015; Niu et al., 2017) (Fig. 3).

All bootstraps values were of 100, except for the relationship between Cymbidieae and Vandeae tribes (BS = 71).

3.3. Phylogenomic analyses based on coding regions

The analysis performed using 60 concatenated protein-coding regions for 21 Cymbidieae taxa, yielded a strongly-supported phylogeny. Bootstrap values for main branches were 100% (Fig. 4). Genera *Otoglossum*, *Oncidium*, *Erycina* and *Telipogon* form a unique clade belonging to the subtribe Oncidiinae (BS=100). Oncidiinae is a sister clade of Maxillariinae (BS=97). According to our sampling, the subtribes Stanhopeinae and Zigopetalinae are sister groups (BS = 90) (Figs. 5 and 6).

In our analysis, values supporting relationships between tribes improved with the inclusion of *Goodyera* (Cranichideae) and *Habenaria* (Orchideae) to Orchidoideae subfamily This leaded to a well-supported phylogeny in which Diurideae is the sister tribe of Cranichideae (BS=94), and Codonorchideae is a sister group to Orchideae (BS=86) (Fig. 7). All bootstrap values were equal or higher than 86.

4. Discussion

The overall results presented here are highly congruent with those produced by previous analyses using whole chloroplasts and a few gene markers (*ITS*, *Xdh*, *psaB*, *rbcL*, *matK*, *ycf1*and/or *trnL-F* intergenic spacer). We obtained a phylogenetic tree for Orchidaceae that is well-resolved and better supported than any produced so far by plastid DNA.

In our phylogenomic analysis, tribes, subtribes and genera that were not reported before were included: the tribe Cypripedieae, the subtribes Elleanthinae, Goodyerinae, Habenariinae, Cypripediinae, Paphiopedilinae and the genera *Otoglossum, Erycina, Telipogon, Scaphosepalum, Teagueia, Calanthe, Bletilla, Elleanthus, Goodyera, Habenaria, Cypripedium* and *Paphiopedilum*. The Eulophiinae subtribe did appear in Givnish et al., (2015), but it was not included in Pérez-Escobar et al., (2017), and it has also been reported here.

4.1. Chloroplast annotation

A total of 80 protein genes were annotated for the 124 orchids. Around 30 transfer RNAs genes, and four ribosomal genes were also identified. Annotated genes belong to the photosystems I and II, cytochrome b/f complex, *ATP* synthase, *NADH* dehydrogenase, RubisCO large subunit, *RNA* polymerase, ribosomal proteins, *clpP*, *matK*, hypothetical chloroplast reading frames (*ycf*), transfer RNAs and ribosomal RNAs. It is common to find transfer RNAs genes, ribosomal RNAs genes, ribosomal protein genes, *ndhB* and *ycf2* genes within the inverted repeated regions (IR) of orchids chloroplasts. While, genes such as *ycf1*, ribosomal protein genes, photosystem genes and the majority of the *ndh* genes are commonly found along the short single copy region (SSC) (Fig. 1). Finally, the rest of the protein coding genes are found in the long single copy region (LSC), as well as other transfer RNAs genes.

4.2. Orchid chloroplast evolution

Comparing orchid chloroplasts with the *Nicotiana tabacum* chloroplast reported at NCBI, some differences were identified. In terms of total genes content, *N. tabacum* chloroplast has 144 genes, while in orchids chloroplasts the gene content is around 120. Protein coding genes are more abundant in *N. tabacum* than in orchids, being 98 and around 62 respectively. Two protein genes found in orchids chloroplasts (*infA* and *pbf1*) were not found in *N. tabacu*m, and six protein genes (*ndhB*, *rpl2*, *rpl23*, *rps12*, *rps7* and *ycf2*) were found as duplicated genes within the IR regions in both chloroplasts. Much has been said about the movement of the *ndh* genes between the chloroplast and the nucleus. The *N. tabacum* chloroplast has 11 *ndh* genes (*ndhA*, *ndhB*, *ndhC*, *ndhD*, *ndhE*, *ndhF*, *ndhG*, *ndhH*, *ndhI*, *ndhJ*, *ndhK*), as well as the chloroplast genome of *Apostasia wallichii*, which has showed to have and transcribe all 11 *ndh* genes and has been predicted to encode for functional proteins (Givnish et al., 2010). For some other orchids, not all those 11 genes are present, as in the case of *Gongora pleiochroma*, where just eight *ndh* genes are present (*ndhA*, *ndhB*, *ndhC*, *ndhD*, *ndhE*, *ndhG*, *ndhH*, *ndhI*). These findings indicate that probably the orchid common ancestor contained a complete functional set of *ndh* genes.

Diverse patterns of junctions between IR and SSC regions are seen in the 11 orchids sequenced here. Some chloroplasts have portions of the genes *rpl22* and *ycf1* within the IR region. Those genes seem to be repeated in some orchids, contributing with the expansion and contraction among the inverted regions, which flank the small single-copy of the chloroplasts. Studies around chloroplasts content, have also found both loss and retention of *ndh* genes among orchids (Chris Blazier et al., 2011; Lin et al., 2015). Few *ndh* genes are thought to encode for functional *ndh* proteins in *Oncidium* and *Cymbidium* (Wu et al., 2010; Yang et al., 2013). *ndh* genes function is thought to be related with land plants adaptation and photosynthesis (Martín and Sabater, 2010). However, Lin et al., (2015) found that no significant differences in biogeography or growth conditions (including light and water requirements) were observed between orchid chloroplasts where *ndh* genes were lost and orchid chloroplasts where *ndh* genes are present. Mechanisms leading to shifts in IR boundaries and the variable loss or retention of *ndh* genes are still unclear (Chris Blazier et al., 2011; Niu et al., 2017).

4.3. Evolutionary relationships of Orchidaceae

Twenty protein-coding genes were identified as problematic and were removed from the reported phylogenies. Problematics are due to the genes were out of the reading frame and some of them have multiple stop codons. Few species could be aligned to the *ycf1* gene, which if included, may have caused noise in the phylogenetic analysis. Some of these genes have also been removed from other orchid phylogenies previously reported (Chang et al., 2006; Logacheva et al., 2011; Yang et al., 2013).

Our results differ from those obtained in previous researches (Givnish et al., 2015; Salazar et al., 2003) in the categorization of Codonorchideae. This tribe appeared as the sister group of Diurideae + Cranichideae. Differences also were observed here with the position of Orchideae as the outermost tribe in Orchidoideae. Nevertheless, our results are in agreement with the phylogeny reported by (Pridgeon et al., 2001), which used *rbcL* gene and Parsimony in PAUP (Swofford, 2003) to infer Orchidoideae topology.

Our whole chloroplasts analysis led to similar results as the reported by Givnish et al. (2015) and Niu et al. (2017). Sampling within genera (Stanhopeinae, Maxillariinae, Oncidiinae, Eulophiinae and Cymbidiinae) conserved the same topologies but with all bootstrap values higher than previously published (Figs. 5 and 6). Our results provide resolution among Cymbideae subtribes, however we are still constrained by the lack of representatives for the subtribes Eriopsidiinae and Coleopsidinae. In our phylogeny obtained using 60 chloroplast coding regions, the relationships among Stanhopeinae and Zigopetalinae, and Oncidiinae with Maxillariinae differ from previous studies (Givnish et al., 2015; Pérez-Escobar et al., 2017). Also, our coding genes phylogeny analysis disagrees with whole chloroplast phylogeny presented here (Fig. 6). When using whole chloroplasts, Stanhopeinae remains as a sister group to Maxillariinae. However, when using only coding regions, Stanhopeinae is defined as sister to Zigopetalinae, and both are sister subtribes to the Maxillariinae + Oncidiinae clade (Fig 6).

The Cymbidieae phylogenies proposed by Freudenstein & Chase (2015), Li et al., (2016), Pérez-Escobar et al., (2017) differ from the one presented here through coding regions analysis. Differences are found in the placement of the subtribes Maxillariinae (sister to Stanhopeinae), Zygopetalinae (sister to Maxillarinae and Stanhopeinae) and Eulophinae, which is sister to Catasetinae in studies reported by Freudenstein & Chase (2015) and Pérez-Escobar et al., (2017). Li et al., (2016) and Pérez-Escobar et al., (2017) found Dipodiinae (*Dipodium*) as the second most basal subtribe sister to the rest of Cymbidieae. However, the genus *Dipodium* has been previously included within Eulophiinae (Chase et al., 2015) and it is not represented in our phylogeny.

Phylogenetic relationships within the tribe Cymbidieae have changed through the years according to the available data and approximations taken, either morphological and/or genetic. In (Dressler, 1993), Cymbidieae contained seven subtribes (Goveniinae, Bromheadiinae, Eulophiinae, Theostelinae, Cyrtopodiinae, Acriopsidinae and Catasetinae), and circumscriptions were very different as the known today. A later study has shown that Cymbidieae could comprise up to 11 subtribes (Li et al., 2016). But the latest study (Pérez-Escobar et al., 2017), reported ten well-supported and circumscribed subtribes: (((Stanhopeinae,Coeliopsidinae)(Maxillarinae(Eriopsidinae(Zygopetalinae(Oncidiinae(Cyrtopodinae(Catasetinae, Eulophinae)(Cymbidiinae). Some topological differences can be identified with respect to our study. Here, relationships among most derived subtribes showed Stanhopeinae as a sister group to Zygopetalinae and Maxillariinae as the sister subtribe of Oncidiinae. Also, the position of Eulophiinae within Catasetinae and Cyrtopodiinae, does not agree with our findings, because Eulophiinae was placed as sister group to the most derived Cymbidieae subtribes, and Catasetinae was clustered together with Cyrtopodiinae (Figs. 5 and 6).

Most of the Cymbidieae species are epiphytes, however almost all subtribes (except Catasetinae) also have terrestrial species. Evolutionary transitions from terrestrial to epiphytic habits has played an important role in orchids diversification, once epiphytism is achieved, the rates of orchid diversification increase (Freudenstein and Chase, 2015). Those subtribes that are richest in species (Oncidiinae = 1615, Maxillariinae = 819, Zygopetalinae = 437 and Catasetinae = 354) may be more speciose, partly because of the adoption of the epiphytic habit. These can be due to the movement into mountainous areas (Givnish et al., 2015), and because of changes in the rate of uplift of the Andes (Pérez-Escobar et al., 2017). Unlike other subtribes, most Eulophiinae species are terrestrial and widely distributed in the Old-World tropics of Africa, Asia and Australasia, with few taxa in the Neotropics. However, the Madagascan genera *Cymbidiella*, *Eulophiella*, *Grammangis* and *Paralophia* are all epiphytes (Bone et al., 2015). Nevertheless in Eulophiinae, more speciose genera are terrestrial (*Eulophia*: 200 species and *Oeceoclades*: 38 species).

Here we have presented, for first time, a well-supported phylogeny for the backbone of Orchidoideae. The phylogeny obtained using complete chloroplasts yielded a strongly-supported topology: Diurideae + Cranichideae and Orchideae as the outermost group, lacking a representative of Codonorchideae. Our approach using 60 coding regions, supports findings of (Pridgeon et al., 2001), in which Diurideae and Cranichideae are sisters groups, as well as Codonorchideae and Orchideae. Our findings differ from Givnish *et al.* (2015) and Salazar et al. (2003), in which Diurideae + Cranichideae form a clade – same as here – but this clade is a sister group to Codonorchideae and Orchideae is placed as the outermost tribe within Orchidoideae (Fig. 7).

All Orchidoideae members have terrestrial habits and cosmopolitan distribution. The most speciose subtribe is Orchidinae (Orchideae) with 1,811 species. Records on pollination, have shown that *Dactylorhiza* genus is pollinated by dipterans and beetles, which are attracted by scent (Gutowski, 1990). At the same time, *Habenaria* is pollinated by moths (Smith and Snow, 1976). Inflorescences within Orchidoideae are commonly terminal and racemose, but in the case of the monotypic tribe Codonorchideae (one genus = *Codonorchis*), those characters are not present. In fact, *Codonorchis* present a single flower. This genus is only present in the South of the Andes and Paraná state.

Rhizanthellinae and Thelymitrinae are grouped together within the Diurideae tribe. They share the geographical distribution, being common in Southeast Asia, Japan, New Zealand and Australia. Both count with inflorescences, but the way they resemble in each tribe is very different. The monotypic group Rhizanthellinae (one genus = *Rhizanthella*) has a very particular inflorescence. It seems to be a solitary florescence but when it blooms under the leaf litter (which is also a unique character), tiny and densely grouped flowers can be observed. The inflorescences in Thelymitrinae are quite different from the rest of the subtribes within Orchidoideae. In its case, the size of the flowers is considerably bigger (1 to 6 cm, compared to 1 cm or less in other subtribes).

5. Conclusions

This study presents a well-resolved and better supported phylogeny for the Orchidaceae family than any produced so far by plastid DNA analyses. Here we report the complete chloroplast sequences of 11 orchid species: *G. pleiochroma, M. nasuta, M. sanderiana, O. globuliferum, T. glicensteinii, S. antenniferum, T. aliana, S. decora, S. mandonii, S. mucronata* and *G. repens*. These 11 plastomes differ in the IR boundaries and the loss/retention of *ndh* genes. For deep branches within Cymbidieae subtribe and Orchidoideae subfamily, supports were improved, leading to the first well-supported phylogeny for Orchidoideae. Comparison of two approaches to infer phylogenies from chloroplast, showed different topologies due to different taxa sampling. Although sampling was sufficient to resolve the relationships between the major clades in the family, the lack of sampling of several key genera (*Zygopetalum*, *Catasetum* and *Cyrtopodium)* and representatives for Eriopsidiinae and Coleopsidinae subtribes, will adress future work on whole chloroplast analysis.

Acknowledments

We would like to thank Esteban Urrea for helping with bioinformatics pipelines. We thank Janice Valencia for comments about the paper. We also thank Juan David Pineda Cardenas for advising about computational resources used through EAFIT and Juan Carlos Correa for computational advices at BIOS. Finally, we would like to thank IDEA WILD for supporting with photographical equipment.

Fig. 1. Smallest and largest chloroplast genomes found in eleven orchids sequenced here: *Maxillaria sanderiana* and *Sobralia mucronata* respectively. Genes shown inside the circle are transcribed clockwise, and those outside the circle are transcribed counterclockwise.

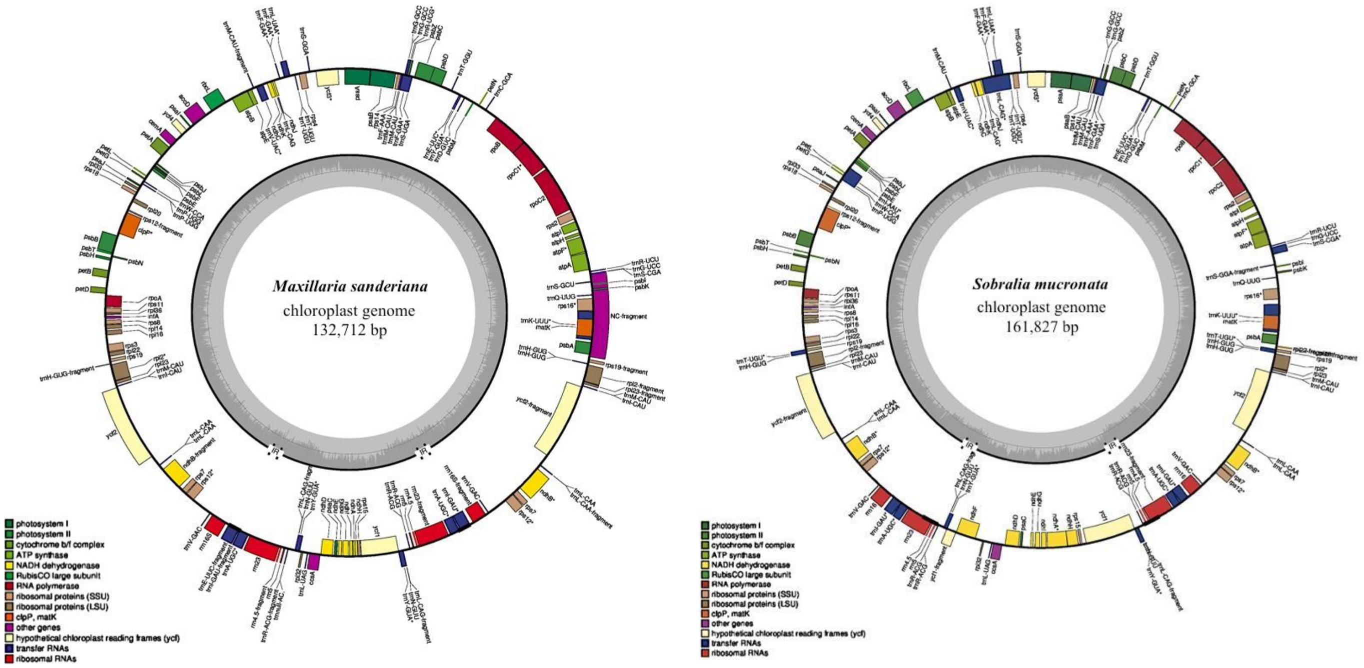


Table 1
Comparison between the set of genes alignments

Gene	127 species*	97 species**
accD	54	54
atpA	111	85
atpB	114	87
atpE	115	90
atpF	113	86
atpH	116	88
atpI	117	90
ccsA	68	43
cemA	103	74
clpP	86	59
infA	117	90
matK	74	47
ndhA	48	35
ndhB	92	59
ndhC	59	41
ndhD	26	4
ndhE	88	70
ndhF	1	-
ndhG	50	29
ndhH	36	19
ndhI	17	8
ndhJ	18	-
ndhK	19	4
pbf1	91	90
petA	117	93
petB	27	-
petD	28	-
petG	117	91
petL	114	90
petN	113	91
psaA	109	85
psaB	114	88
psaC	115	91
psaI	109	87
psaJ	100	75
psbA	113	85
psbB	105	90
psbC	115	90
psbD	113	87
psbE	119	92
psbF	117	90
psbH	111	89
psbI	114	91
psbJ	117	92
psbK	114	88
psbL	117	90
psbM	106	83
psbN	110	90
psbT	109	91
psbZ	92	91
rbcL	105	77
rpl14	121	97
rpl16	29	-
rpl2	117	90
rpl20	83	62
rpl22	89	61
rpl23	115	87
rpl32	96	77
rpl33	116	89
rpl36	117	95
rpoA	105	79
rpoB	111	87
rpoC1	110	84
rpoC2	66	67
rps11	119	91
rps12	28	-
rps14	122	96
rps15	109	88
rps16	62	36
rps18	110	87
rps19	118	92
rps2	103	74
rps3	112	85
rps4	117	94
rps7	121	92
rps8	118	93
ycf1	10	7
ycf2	111	84
ycf3	109	83
ycf4	95	76

* Includes 86 whole chloroplasts from NCBI, 11 new chloroplasts and 30 species sampled in Givnish *et al*, 2015.

** Includes 86 whole chloroplasts from NCBI and the 11 new whole chloroplast

Table 2
Comparison of major features of eleven orchid chloroplast genomes

Species	**Accession number**	**Size (bp)**	**LSC* lenght (bp)**	**SSC** lenght (bp)**	**IR*** lenght (bp)**	**Number of different genes**	**Duplicated genes in IR**	**Protein-coding genes**	***tRNA* genes**	***rRNA* genes**	**GC content (%)**
Gongora pleiochroma	XXXXXXXX	146,990	82,808	13,005	25,442	117	22	61	30	4	37.3
Maxillaria sanderiana	XXXXXXXX	132,712	74,195	8,638	24,807	123	24	62	33	4	38.6
Maxillaria nasuta	XXXXXXXX	144,213	81,128	12,357	25,251	121	22	64	31	4	37.7
Otoglossum globuliferum	XXXXXXXX	145,149	82,340	11,902	25,447	121	22	64	31	4	37.3
Telipogon glicensteinii	XXXXXXXX	143,414	80,462	11,785	25,559	113	22	57	30	4	37.0
Scaphosepalum antenniferum	XXXXXXXX	156,106	84,789	19,973	25,802	118	22	62	30	4	37.0
Teagueia aliana	XXXXXXXX	155,682	83,712	18,225	27,562	119	24	62	29	4	37.2
Sobralia decora	XXXXXXXX	160,230	87,540	20,449	26,282	120	24	61	31	4	37.3
Sobralia mandonii	XXXXXXXX	160,062	87,346	19,454	27,313	120	24	61	31	4	37.4
Sobralia mucronata	XXXXXXXX	161,827	88,602	19,845	27,311	122	24	64	30	4	37.1
Goodyera repens	XXXXXXXX	151,361	81,945	17,583	26,305	122	22	64	32	4	37.6

* Long Single Copy (LSC) section of the chloroplast
** Short Single Copy (SSC) section of the chloroplast
*** Inverted Repeats (IR) of the chloroplast

Fig. 2. Chloroplast phylogeny for Orchidaceae based on ML analysis of sequence variation in 94 orchids and 3 Asparagales outgroups whole chloroplasts. Colored boxes correspond to new chloroplasts sequences, the rest are the chloroplasts found in NCBI. Bootstraps support values are shown above each branch

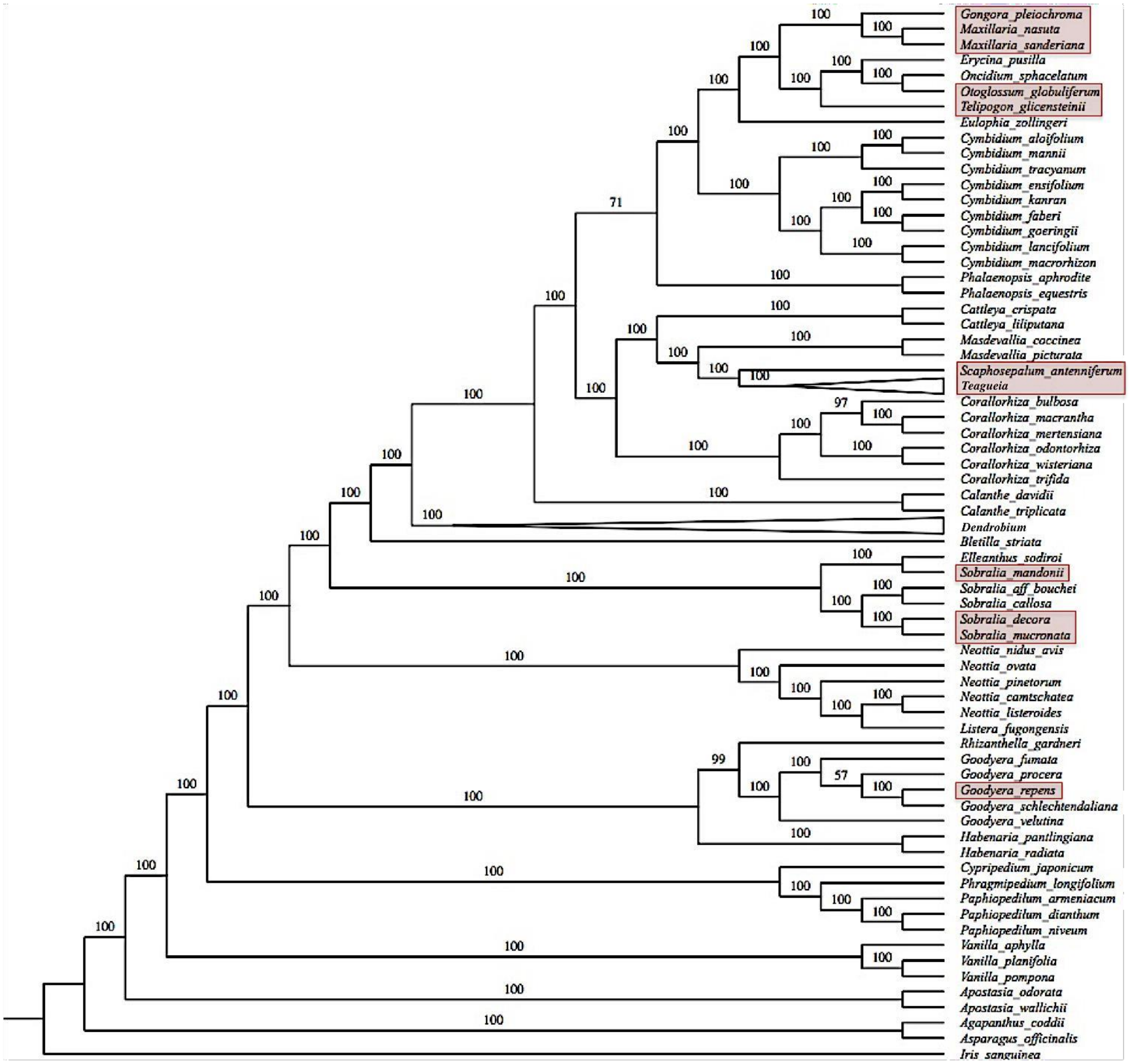


Fig. 3. Comparison between A) Givnish et al., 2015 phylogeny and B) best-scoring ML phylogeny presented here based on 60 coding regions. Colored boxes correspond to tribes, and bold words to subfamilies.

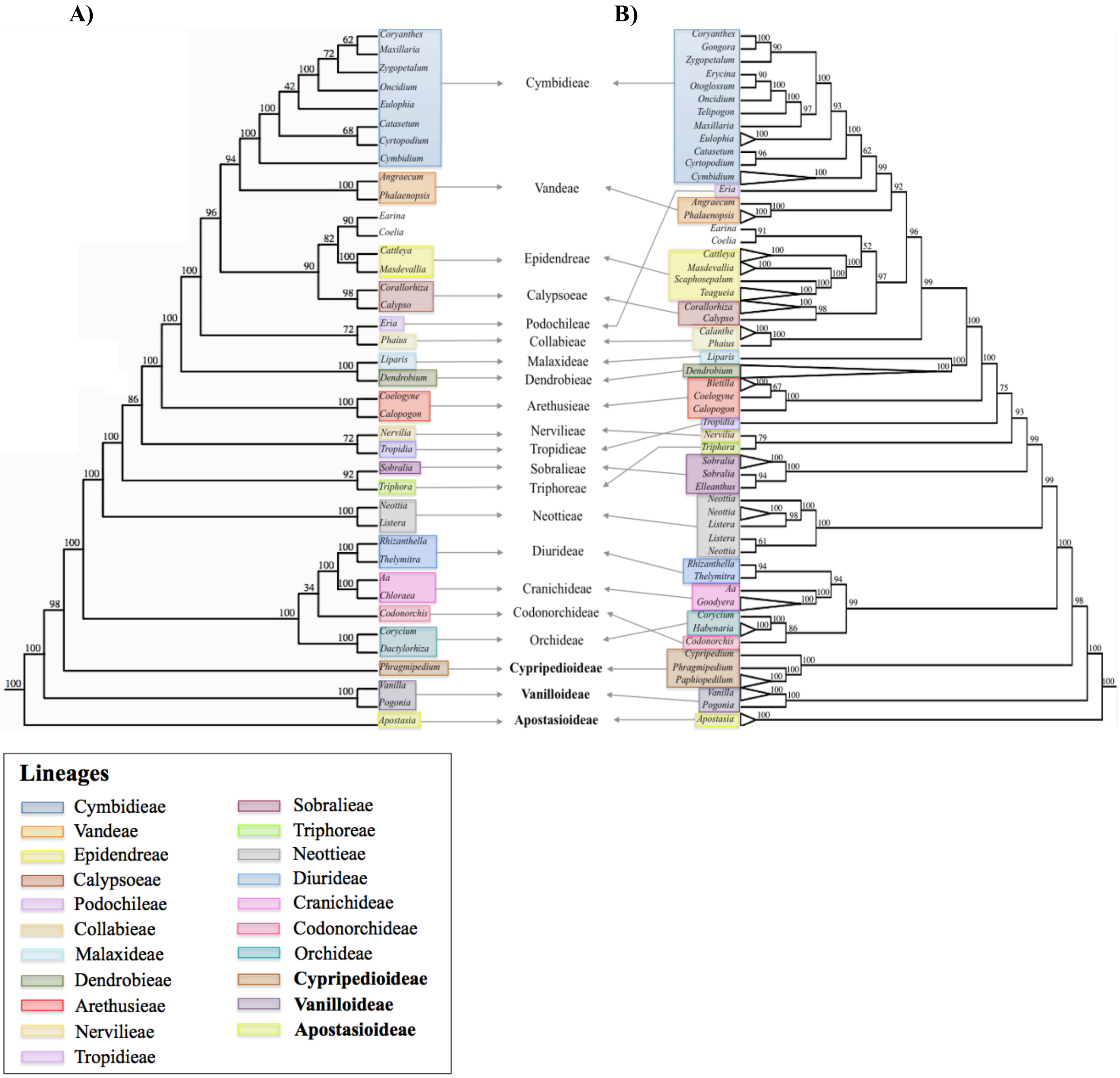


Fig. 4. Cymbidieae phylogeny based on ML analysis of sequence variation: 60 genes in 21 species and 6 outgroups. Bootstrap support values are shown above each branch.

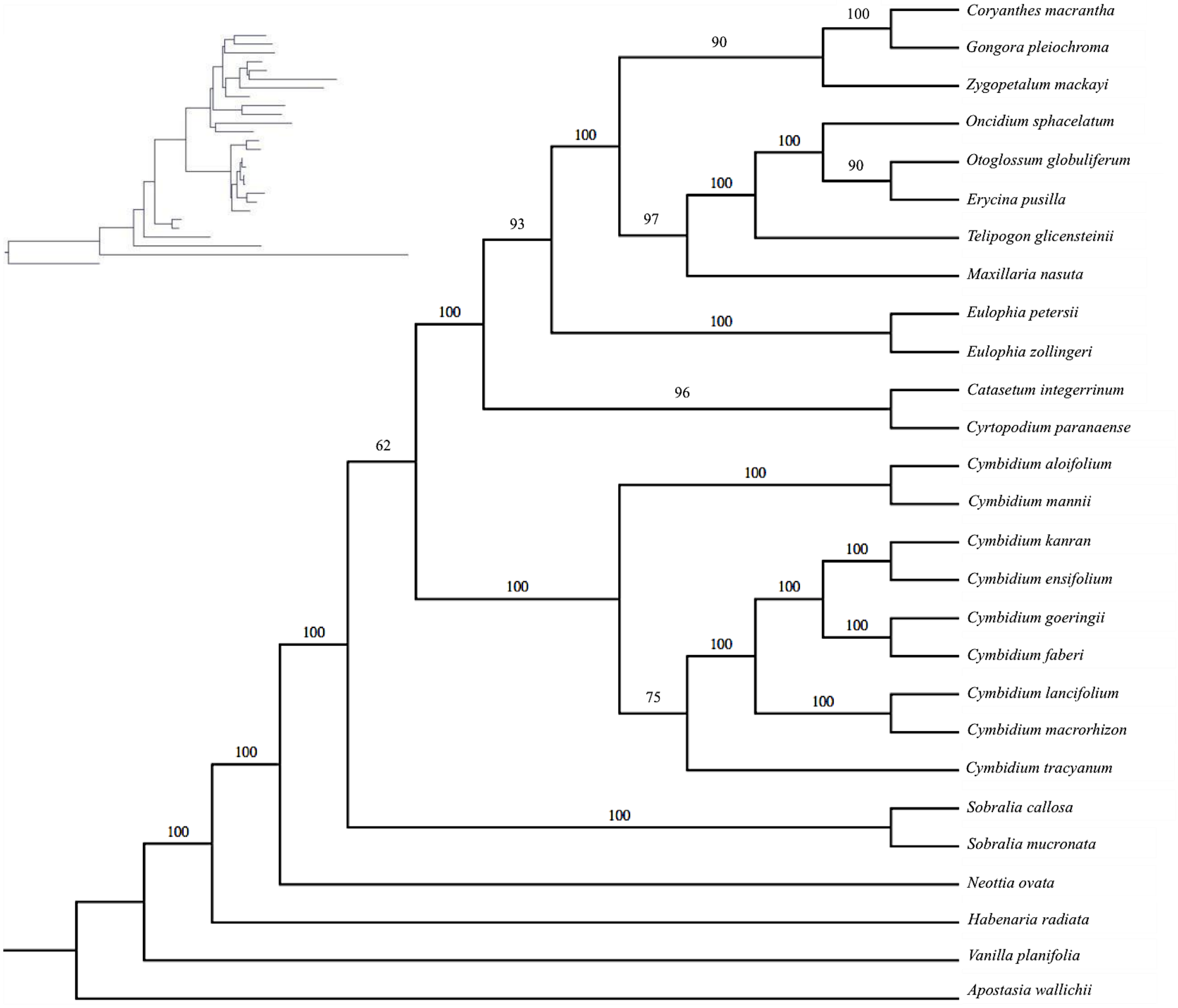


Fig. 5. Comparison between A) Cymbidieae phylogeny achieved by Givnish et al., 2015 and B) Zoom of Cymbidieae tribe from all Orchidaceae best-scoring ML phylogeny based on 60 genes. Colored boxes correspond to subtribes.

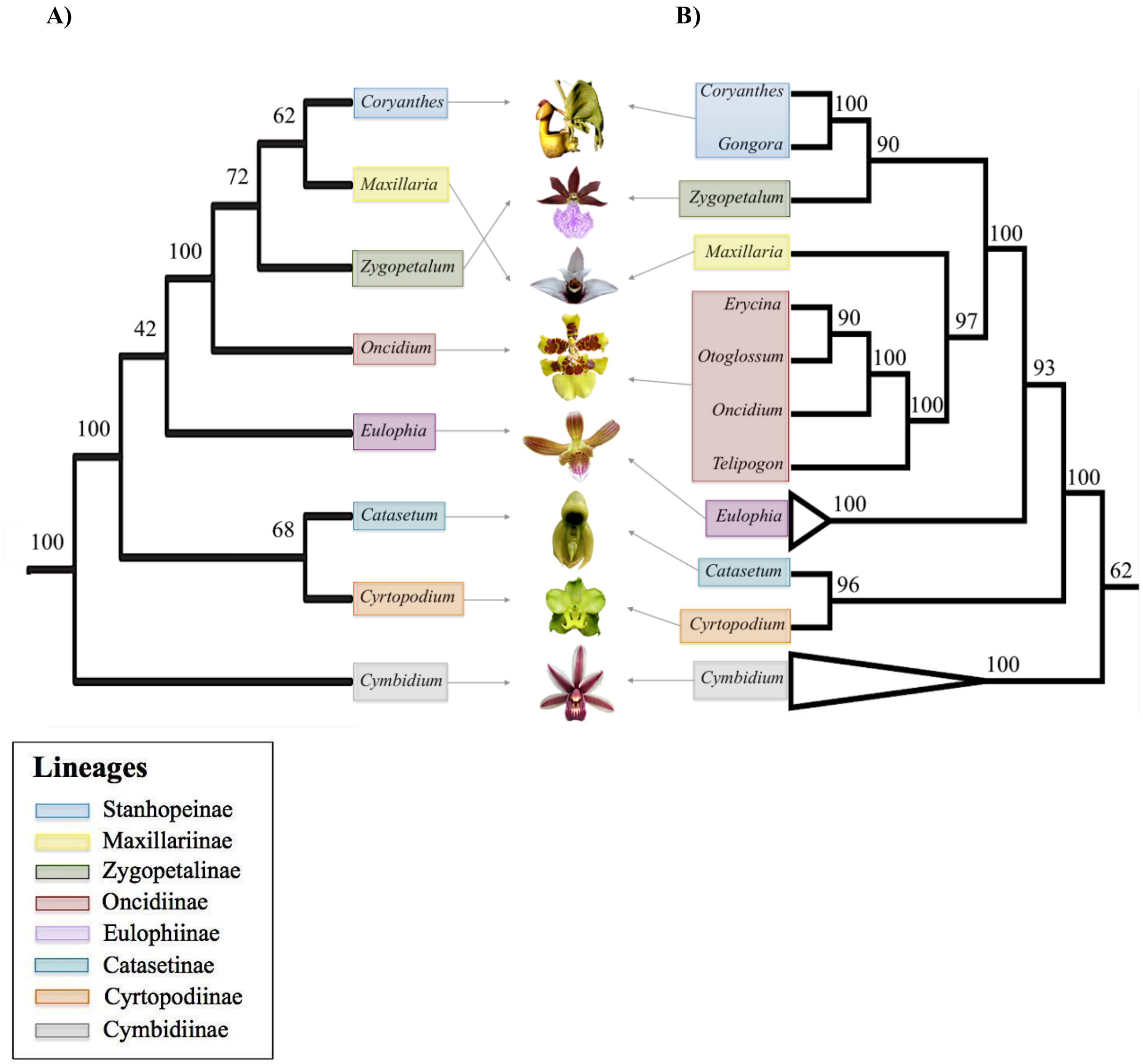


Fig. 6. Comparison between A) Whole chloroplast phylogeny and B) Zoom of Cymbidieae tribe from all Orchidaceae best-scoring ML phylogeny based on 60 genes. Colored boxes correspond to subtribes.

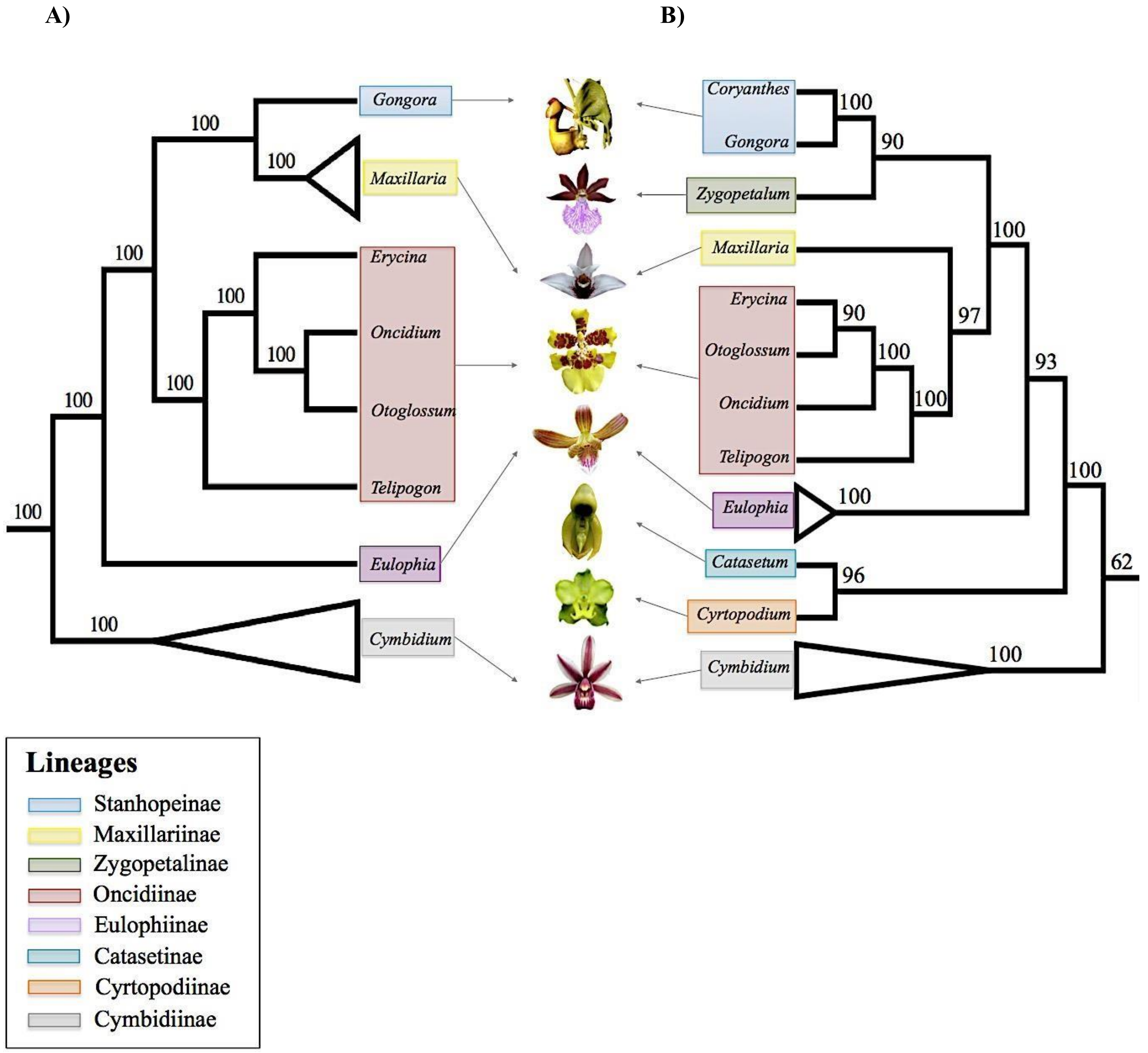
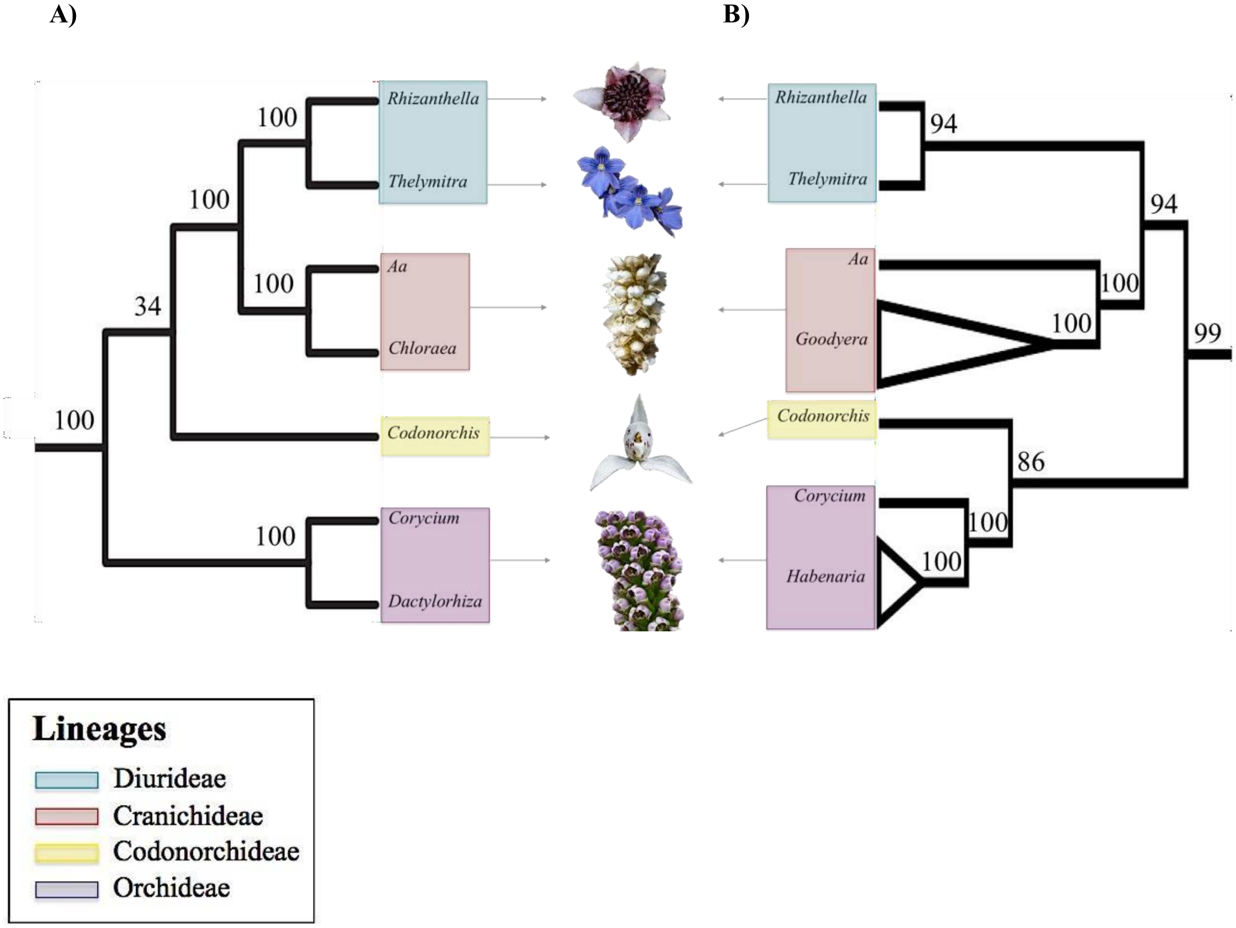


Fig. 7. Comparison between A) Orchidoideae phylogeny by Givnish et al., 2015 and B) best-scoring ML phylogeny based on 60 genes. Colored boxes correspond to tribes.



References

Bankevich, A., Nurk, S., Antipov, D., Gurevich, A.A., Dvorkin, M., Kulikov, A.S., Lesin, V.M., Nikolenko, S.I., Pham, S., Prjibelski, A.D., Pyshkin, A.V., Sirotkin, A.V., Vyahhi, N., Tesler, G., Alekseyev, M.A., Pevzner, P.A., 2012. SPAdes: A New Genome Assembly Algorithm and Its Applications to Single-Cell Sequencing. J. Comput. Biol. 19, 455–477. https://doi.org/10.1089/cmb.2012.0021

Bolger, A.M., Lohse, M., Usadel, B., 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 30, 2114–2120. https://doi.org/10.1093/bioinformatics/btu170

Bone, R.E., Cribb, P.J., Buerki, S., 2015. Phylogenetics of Eulophiinae (Orchidaceae: Epidendroideae): evolutionary patterns and implications for generic delimitation: Evolutionary patterns in Eulophiinae. Bot. J. Linn. Soc. 179, 43–56. https://doi.org/10.1111/boj.12299

Bushnell, 2017. BBMap/BBTools [WWW Document]. URL https://sourceforge.net/projects/bbmap/files/ (accessed 4.28.19).

Bushnell, B., Rood, J., Singer, E., 2017. BBMerge – Accurate paired shotgun read merging via overlap. PLOS ONE 12, e0185056. https://doi.org/10.1371/journal.pone.0185056

Chan, C.X., Ragan, M.A., 2013. Next-generation phylogenomics. Biol. Direct 8. https://doi.org/10.1186/1745-6150-8-3

Chang, Ching-Chun, Lin, H.-C., Lin, I.-P., Chow, T.-Y., Chen, H.-H., Chen, W.-H., Cheng, C.-H., Lin, C.-Y., Liu, S.-M., Chang, Chien-Chang, Chaw, S.-M., 2006. The Chloroplast Genome of Phalaenopsis aphrodite (Orchidaceae): Comparative Analysis of Evolutionary Rate with that of Grasses and Its Phylogenetic Implications. Mol. Biol. Evol. 23, 279–291. https://doi.org/10.1093/molbev/msj029

Chase, M.W., Cameron, K.M., Freudenstein, J.V., Pridgeon, A.M., Salazar, G., van den Berg, C., Schuiteman, A., 2015. An updated classification of Orchidaceae: Updated Classification of Orchidaceae. Bot. J. Linn. Soc. 177, 151–174. https://doi.org/10.1111/boj.12234

Chevreux, B., Wetter, T., Suhai, S., 1999. Genome sequence assembly using trace signals and additional sequence information., in: German Conference on Bioinformatics. Citeseer, pp. 45–56.

Chris Blazier, J., Guisinger, M.M., Jansen, R.K., 2011. Recent loss of plastid-encoded ndh genes within Erodium (Geraniaceae). Plant Mol. Biol. 76, 263–272. https://doi.org/10.1007/s11103-011-9753-5

Cingel, N.A.V.D., 2001. An Atlas of Orchid Pollination: European Orchids. CRC Press.

Cock, P.J.A., Grüning, B.A., Paszkiewicz, K., Pritchard, L., 2013. Galaxy tools and workflows for sequence analysis with applications in molecular plant pathology. PeerJ 1, e167. https://doi.org/10.7717/peerj.167

Dong, W.-L., Wang, R.-N., Zhang, N.-Y., Fan, W.-B., Fang, M.-F., Li, Z.-H., 2018. Molecular Evolution of Chloroplast Genomes of Orchid Species: Insights into Phylogenetic Relationship and Adaptive Evolution. Int. J. Mol. Sci. 19, 716. https://doi.org/10.3390/ijms19030716

Dressler, R.L., 1993. Phylogeny and Classification of the Orchid Family. Cambridge University Press.

Fay, M.F., Chase, M.W., 2009. Orchid biology: from Linnaeus via Darwin to the 21st century. Ann. Bot. 104, 359–364. https://doi.org/10.1093/aob/mcp190

Freudenstein, J.V., Chase, M.W., 2015. Phylogenetic relationships in Epidendroideae (Orchidaceae), one of the great flowering plant radiations: progressive specialization and diversification. Ann. Bot. 115, 665–681. https://doi.org/10.1093/aob/mcu253

Geneious version (9) created by Biomatters. <http://www.geneious.com/>.

Givnish, T.J., Ames, M., McNeal, J.R., McKain, M.R., Steele, P.R., dePamphilis, C.W., Graham, S.W., Pires, J.C., Stevenson, D.W., Zomlefer, W.B., Briggs, B.G., Duvall, M.R., Moore, M.J., Heaney, J.M., Soltis, D.E., Soltis, P.S., Thiele, K., Leebens-Mack, J.H., 2010. Assembling the Tree of the Monocotyledons: Plastome Sequence Phylogeny and Evolution of Poales [1]. Ann. Mo. Bot. Gard. 97, 584–616. https://doi.org/10.3417/2010023

Givnish, T.J., Spalink, D., Ames, M., Lyon, S.P., Hunter, S.J., Zuluaga, A., Iles, W.J.D., Clements, M.A., Arroyo, M.T.K., Leebens-Mack, J., Endara, L., Kriebel, R., Neubig, K.M., Whitten, W.M., Williams, N.H., Cameron, K.M., 2015. Orchid phylogenomics and multiple drivers of their extraordinary diversification. Proc. R. Soc. B Biol. Sci. 282, 20151553. https://doi.org/10.1098/rspb.2015.1553

Gutowski, J.M., 1990. Pollination of the orchid Dactylorhiza fuchsii by longhorn beetles in primeval forests of Northeastern Poland. Biol. Conserv. 51, 287–297. https://doi.org/10.1016/0006-3207(90)90114-5

Huang, J.-Z., Lin, C.-P., Cheng, T.-C., Huang, Y.-W., Tsai, Y.-J., Cheng, S.-Y., Chen, Y.-W., Lee, C.-P., Chung, W.-C., Chang, B.C.-H., Chin, S.-W., Lee, C.-Y., Chen, F.-C., 2016. The genome and transcriptome of *Phalaenopsis* yield insights into floral organ development and flowering regulation. PeerJ 4, e2017. https://doi.org/10.7717/peerj.2017

Katoh, K., Standley, D.M., 2013. MAFFT Multiple Sequence Alignment Software Version 7: Improvements in Performance and Usability. Mol. Biol. Evol. 30, 772–780. https://doi.org/10.1093/molbev/mst010

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R., 1000 Genome Project Data Processing Subgroup, 2009. The Sequence Alignment/Map format and SAMtools. Bioinformatics 25, 2078–2079. https://doi.org/10.1093/bioinformatics/btp352

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, 37. https://doi.org/10.11646/phytotaxa.246.1.3

Lin, C.-S., Chen, J.J.W., Huang, Y.-T., Chan, M.-T., Daniell, H., Chang, W.-J., Hsu, C.-T., Liao, D.-C., Wu, F.-H., Lin, S.-Y., Liao, C.-F., Deyholos, M.K., Wong, G.K.-S., Albert, V.A., Chou, M.-L., Chen, C.-Y., Shih, M.-C., 2015. The location and translocation of ndh genes of chloroplast origin in the Orchidaceae family. Sci. Rep. 5. https://doi.org/10.1038/srep09040

Logacheva, M.D., Schelkunov, M.I., Penin, A.A., 2011. Sequencing and Analysis of Plastid Genome in Mycoheterotrophic Orchid Neottia nidus-avis. Genome Biol. Evol. 3, 1296–1303. https://doi.org/10.1093/gbe/evr102

Luo, J., Hou, B.-W., Niu, Z.-T., Liu, W., Xue, Q.-Y., Ding, X.-Y., 2014. Comparative Chloroplast Genomes of Photosynthetic Orchids: Insights into Evolution of the

Orchidaceae and Development of Molecular Markers for Phylogenetic Applications. PLoS ONE 9, e99016. https://doi.org/10.1371/journal.pone.0099016

Magoc, T., Salzberg, S.L., 2011. FLASH: fast length adjustment of short reads to improve genome assemblies. Bioinformatics 27, 2957–2963. https://doi.org/10.1093/bioinformatics/btr507

Martín, M., Sabater, B., 2010. Plastid ndh genes in plant evolution. Plant Physiol. Biochem. 48, 636–645. https://doi.org/10.1016/j.plaphy.2010.04.009

Niu, Z., Xue, Q., Zhu, S., Sun, J., Liu, W., Ding, X., 2017. The Complete Plastome Sequences of Four Orchid Species: Insights into the Evolution of the Orchidaceae and the Utility of Plastomic Mutational Hotspots. Front. Plant Sci. 8. https://doi.org/10.3389/fpls.2017.00715

Pansarin, L.M., De Moraes Castro, M., Sazima, M., 2009. Osmophore and elaiophores of *Grobya amherstiae* (Catasetinae, Orchidaceae) and their relation to pollination. Bot. J. Linn. Soc. 159, 408–415. https://doi.org/10.1111/j.1095-8339.2009.00953.x

Parakhia, M.V., Tomar, R.S., Malaviya, B.J., Dhingani, R.M., Rathod, V.M., Thakkar, J.R., Golakiya, B.A., 2014. Draft Genome Sequence of the Endophytic Bacterium Enterobacter spp. MR1, Isolated from Drought Tolerant Plant (Butea monosperma). Indian J. Microbiol. 54, 118–119. https://doi.org/10.1007/s12088-013-0429-5

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 Phytol. 215, 891–905. https://doi.org/10.1111/nph.14629

Pridgeon, A., 2009. Genera Orchidacearum Vol. 5, Vol. 5,. Oxford University Press, Oxford.

Pridgeon, A.M., Cribb, P.J., Chase, M.W., 2001. Genera Orchidacearum: Volume 2. Orchidoideae. OUP Oxford.

Ramirez, S.R., Roubik, D.W., Skov, C., Pierce, N.E., 2010. Phylogeny, diversification patterns and historical biogeography of euglossine orchid bees (Hymenoptera: Apidae). Biol. J. Linn. Soc. 100, 552–572.

Renner, S.S., Schaefer, H., 2010. The evolution and loss of oil-offering flowers: new insights from dated phylogenies for angiosperms and bees. Philos. Trans. R. Soc. B Biol. Sci. 365, 423–435. https://doi.org/10.1098/rstb.2009.0229

Salazar, G.A., Chase, M.W., Soto Arenas, M.A., Ingrouille, M., 2003. Phylogenetics of Cranichideae with emphasis on Spiranthinae (Orchidaceae, Orchidoideae): evidence from plastid and nuclear DNA sequences. Am. J. Bot. 90, 777–795. https://doi.org/10.3732/ajb.90.5.777

Smith, D.R., 2015. Mutation Rates in Plastid Genomes: They Are Lower than You Might Think. Genome Biol. Evol. 7, 1227–1234. https://doi.org/10.1093/gbe/evv069

Smith, G.R., Snow, G.E., 1976. Pollination Ecology of Platanthera (Habenaria) Ciliaris and P. blephariglottis (Orchidaceae). Bot. Gaz. 137, 133–140. https://doi.org/10.1086/336852

Stamatakis, A., 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics 30, 1312–1313. https://doi.org/10.1093/bioinformatics/btu033

Swofford, D., 2003. PAUP*: phylogenetic analysis using parsimony, version 4.0b10.

Tillich, M., Lehwark, P., Pellizzer, T., Ulbricht-Jones, E.S., Fischer, A., Bock, R., Greiner, S., 2017. GeSeq – versatile and accurate annotation of organelle genomes. Nucleic Acids Res. 45, W6–W11. https://doi.org/10.1093/nar/gkx391

Ward, J.A., Ponnala, L., Weber, C.A., 2012. Strategies for transcriptome analysis in nonmodel plants. Am. J. Bot. 99, 267–276. https://doi.org/10.3732/ajb.1100334

Weitemier, K., Straub, S.C.K., Cronn, R.C., Fishbein, M., Schmickl, R., McDonnell, A., Liston, A., 2014. Hyb-Seq: Combining target enrichment and genome skimming for plant phylogenomics. Appl. Plant Sci. 2, 1400042. https://doi.org/10.3732/apps.1400042

Whitten, W.M., Neubig, K.M., Williams, N.H., 2014. Generic and Subtribal relationShipS in neotropical cymbidieae (orchidaceae) baSed on matK/ycf1 plaStid data. Lankesteriana 13. https://doi.org/10.15517/lank.v13i3.14425

Wu, F.-H., Chan, M.-T., Liao, D.-C., Hsu, C.-T., Lee, Y.-W., Daniell, H., Duvall, M.R., Lin, C.-S., 2010. Complete chloroplast genome of Oncidium Gower Ramsey and evaluation of molecular markers for identification and breeding in Oncidiinae 12.

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 Evol. Biol. 13, 84. https://doi.org/10.1186/1471-2148-13-84

Yuan, Y., Jin, X., Liu, J., Zhao, X., Zhou, J., Wang, X., Wang, D., Lai, C., Xu, W., Huang, J., Zha, L., Liu, D., Ma, X., Wang, L., Zhou, M., Jiang, Z., Meng, H., Peng, H., Liang, Y., Li, R., Jiang, C., Zhao, Y., Nan, T., Jin, Y., Zhan, Z., Yang, J., Jiang, W., Huang, L., 2018. The Gastrodia elata genome provides insights into plant adaptation to heterotrophy. Nat. Commun. 9. https://doi.org/10.1038/s41467-018-03423-5

Zhang, G.-Q., Liu, K.-W., Li, Z., Lohaus, R., Hsiao, Y.-Y., Niu, S.-C., Wang, J.-Y., Lin, Y.-C., Xu, Q., Chen, L.-J., Yoshida, K., Fujiwara, S., Wang, Z.-W., Zhang, Y.-Q., Mitsuda, N., Wang, M., Liu, G.-H., Pecoraro, L., Huang, H.-X., Xiao, X.-J., Lin, M., Wu, X.-Y., Wu, W.-L., Chen, Y.-Y., Chang, S.-B., Sakamoto, S., Ohme-Takagi, M., Yagi, M., Zeng, S.-J., Shen, C.-Y., Yeh, C.-M., Luo, Y.-B., Tsai, W.-C., Van de Peer, Y., Liu, Z.-J., 2017. The Apostasia genome and the evolution of orchids. Nature 549, 379–383. https://doi.org/10.1038/nature23897