

Plastome structure and adaptive evolution of *Calanthe* s.l. species

Yanqiong Chen, Hui Zhong, Yating Zhu, Yuanzhen Huang, Shasha Wu, Zhongjian Liu, Siren Lan and Junwen Zhai

Key Laboratory of National Forestry and Grassland Administration for Orchid Conservation and Utilization at College of Landscape Architecture, Fujian Agriculture and Forestry University, Fuzhou, Fujian, China
Fujian Ornamental Plant Germplasm Resources Innovation & Engineering Application Research Center, Fujian Agriculture and Forestry University, Fuzhou, Fujian, China

ABSTRACT

Calanthe s.l. is the most diverse group in the tribe Collabieae (Orchidaceae), which are pantropical in distribution. Illumina sequencing followed by *de novo* assembly was used in this study, and the plastid genetic information of *Calanthe* s.l. was used to investigate the adaptive evolution of this taxon. Herein, the complete plastome of five *Calanthe* s.l. species (*Calanthe davidii*, *Styloglossum lyroglossa*, *Preptanthe rubens*, *Cephalantheropsis obcordata*, and *Phaius tankervilleae*) were determined, and the two other published plastome sequences of *Calanthe* s.l. were added for comparative analyses to examine the evolutionary pattern of the plastome in the alliance. The seven plastomes ranged from 150,181 bp (*C. delavayi*) to 159,014 bp (*C. davidii*) in length and were all mapped as circular structures. Except for the three *ndh* genes (*ndhC*, *ndhF*, and *ndhK*) lost in *C. delavayi*, the remaining six species contain identical gene orders and numbers (115 gene). Nucleotide diversity was detected across the plastomes, and we screened 14 mutational hotspot regions, including 12 non-coding regions and two gene regions. For the adaptive evolution investigation, three species showed positive selected genes compared with others, *C. obcordata* (*cemA*), *S. lyroglossa* (*infA*, *ycf1* and *ycf2*) and *C. delavayi* (*nad6* and *ndhB*). Six genes were under site-specific positive selection in *Calanthe* s.l., namely, *accD*, *ndhB*, *ndhD*, *rpoC2*, *ycf1*, and *ycf2*, most of which are involved in photosynthesis. These results, including the new plastomes, provide resources for the comparative plastome, breeding, and plastid genetic engineering of orchids and flowering plants.

Submitted 9 June 2020
Accepted 7 September 2020
Published 13 October 2020

Corresponding authors
Siren Lan, lkzx@fafu.edu.cn
Junwen Zhai, zhai-jw@163.com

Academic editor
Alison Nazareno

Additional Information and
Declarations can be found on
page 17

DOI 10.7717/peerj.10051

© Copyright
2020 Chen et al.

Distributed under
Creative Commons CC-BY 4.0

OPEN ACCESS

Subjects Evolutionary Studies, Genomics, Molecular Biology, Plant Science

Keywords *Calanthe*, Sequence divergence, Adaptive evolution, Phylogeny, Molecular marker, Plastome

INTRODUCTION

Orchidaceae is the largest family in angiosperms (ca. 736 genera, ca. 28,000 species; *Chase et al., 2015*; *Christenhusz & Byng, 2016*), with its fascinating biodiversity attracting the interest of numerous botanists. The first plastome of orchids (*Phalaenopsis aphrodite*) was published in 2006 (*Chang et al., 2006*). In total, over 200 plastomes of orchids have been published, including five subfamilies and 51 genera. In general, plastomes are relatively conserved in land plants, but huge divergence was found among different orchid species. Orchids encompass all life forms of plants, including both heterotrophic and autotrophic

species, namely, terrestrial, epiphytic, and saprophytic (also called mycoheterotrophic). The evolutionary direction of the life forms of orchids can be roughly outlined from terrestrial to epiphytic, while saprophytic species independently evolved several times (Sosa *et al.*, 2016). Recent plastome analysis of orchid species has focused on comparisons with partially and fully mycoheterotrophic species (Feng *et al.*, 2016; Barrett & Kennedy, 2018; Unruh *et al.*, 2018; Yuan *et al.*, 2018). The molecular information on these species is of interest not only for the study of phylogenetics but also for evolutionary studies and species conservation.

Calanthe, the largest genus of the Collabieae tribe (Epidendroideae, Orchidaceae), has a pantropical distribution, being widely distributed in tropical and subtropical Asia, Australia, Madagascar, Africa, Central and South America, and the Caribbean (Chen *et al.*, 2009; Clayton & Cribb, 2013). Since the establishment of *Calanthe* in 1821 (Ker Gawler, 1821), the genus has undergone multiple intra-generic taxonomic revisions lasting for centuries. For example, Lindley (1855) set two subgenera according to the spur length, Bentham (1881) and Pridgeon *et al.* (2005) defined the genus with different criteria based on the pseudobulb and floral characteristics. For the taxonomic classification, morphological data are limited. Genetic information suggests that the origin of *Calanthe* is polyphyletic and includes its relatives (*Cephalantheropsis* and *Phaius*) (Yukawa & Ishida, 2008; Xiang *et al.*, 2014; Zhai *et al.*, 2014). The redefinition of the three genera seems inevitable, and *Calanthe* may be subdivided into three genera, *Calanthe*, *Styloglossum*, and *Preptanthe* (Yukawa & Cribb, 2014; Zhai *et al.*, 2014). These genera, *Calanthe*, *Phaius*, *Styloglossum*, *Cephalantheropsis* and *Preptanthe*, form an independent alliance (which we refer to as *Calanthe* s.l.) within Epidendroideae (Orchidaceae). This alliance can be easily distinguished from other taxa in this family, characterized by plicate leaves, similar sepals and petals, and eight waxy pollinia. However, the latest Orchidaceae classification retained the independent genus status of polyphyly *Calanthe* (Chase *et al.*, 2015). To better understand their phylogenetic relationship, it is necessary for us to identify discrepancies in the genetic information of the major clade of *Calanthe* s.l.

In the current study, we assembled and annotated the plastomes of five *Calanthe* s.l. species for the first time. Additionally, we published and compared *Calanthe* s.l. plastome with the aim to: (1) understand the genetic variation within the *Calanthe* s.l. plastome; (2) identify the characteristics of the plastomes structure, sequence divergence, mutational hotspot regions, repeat regions, and examine them as candidate molecular markers for species classification and further species evolution studies; (3) assess the selective pressure among *Calanthe* s.l. species by identifying genes underlying positive selection; and (4) evaluate the phylogenetic relationships within *Calanthe* s.l.

MATERIALS & METHODS

Plant materials, DNA extraction and sequencing

Specimens of five species, *Calanthe davidii*, *Styloglossum lyroglossa*, *Preptanthe rubens*, *Phaius tankervilleae*, and *Cephalantheropsis obcordate*, represent the five major clades of the *Calanthe* s.l. (according to the phylogenetic relationship detected by Zhai *et al.* (2014))

and they all introduced and cultivated in the Fujian Agriculture and Forestry University, Fujian province, China. Their voucher information is given in [Table S1](#). The modified version of the CTAB method ([Doyle & Doyle, 1987](#)) was applied to extract the genomic DNA. We constructed the short-insert (500 bp) pair-end (PE) library and the sequencing was conducted by the Beijing Genomics Institute (Shenzhen, China) on the Illumina HiSeq 2500 platform, with a read length of 150 bp. At least 5 Gb clean data were obtained for each species.

Genome assembling and annotation

We used the GetOrganelle pipe-line ([Jin et al., 2019](#)) to obtain plastid-like reads, with 18 orchid species as references (listed in [Table S2](#)). Then, the filtered reads were assembled using SPAdes version 3.1.0 ([Bankevich et al., 2012](#)). The final plastid-like reads were filtered by the script of GetOrganelle to obtain pure plastid contigs, and the filtered De Bruijn graphs were viewed and edited using Bandage ([Wick et al., 2015](#)) to manually finalize the complete target genomes. DOGMA ([Wyman, Jansen & Boore, 2004](#)) was used to annotate the plastome, a web-based package based on BLAST searches against a custom database that can identify various genes. Then, we used Geneious prime v2019.03 ([Kearse et al., 2012](#)) to manually align the results, and we examined and confirmed the plastomes with *C. triplicata* as the reference. OrganellarGenomeDRAW (OGDRAW) version 1.3.1 (<https://chlorobox.mpimp-golm.mpg.de/OGDraw.html>) ([Greiner et al., 2019](#)) was used to visualize the structural features of the five species.

Genome comparison and analysis

Combined with two published *Calanthe* plastome (*C. triplicata* [KF753635](#) and *C. delavayi*, [MK388860](#)), seven complete plastomes of *Calanthe* s.l. provided the possibility of comparative analysis within the relatives. In addition, We also compared the two plastome-assembled versions of *C. davidii* (D1, [MG708353](#)) and *C. davidii* (D2, assembled in the present study). Mauve alignment ([Darling et al., 2004](#)) was employed to analyse the plastome DNA rearrangement of seven species. The junction regions between the IR, SSC, and LSC of seven species were compared using the online program IRscope (<https://irscope.shinyapps.io/irapp/>) ([Amiryousefi, Hyvönen & Pocza, 2018](#)). Meanwhile, the sequence identity of the seven species were compared and plotted using the program mVISTA (<http://genome.lbl.gov/vista/mvista/submit.shtml>) with Shuffle-LAGAN ([Brudno et al., 2003](#); [Frazer et al., 2004](#)) with *C. triplicata* ([KF753635](#)) as reference to show inter- and intraspecific variations.

To identify the mutational hotspot regions and genes, we calculated nucleotide diversity (π) across the whole plastome. 76 CDS and 56 IGS shared by seven species were used for analysing and were extracted by PhyloSuit ([Zhang et al., 2020](#)). Seven plastome sequences were aligned using MAFFT v7.407 ([Katoh & Standley, 2013](#)) and the nucleotide diversity was detected using DnaSP v6.12.03 (DNA Sequences Polymorphism) ([Rozas et al., 2017](#)) with sliding window strategy. The step size was set to 200 bp, with a 600 bp window length.

Repeat sequence analysis

The online software REPuter (<https://bibiserv.cebitec.uni-bielefeld.de/reputer/>) was used to identify the repeat sequences (Kurtz et al., 2001), including palindromic direct and reverse repeats. The parameters were set as follows: (1) The maximum and minimum computed repeat sizes were limited to 50 and 30, respectively; (2) Hamming distance of 3. Tandem repeat sequences were identified with the tandem repeats finder (Benson, 1999). The alignment parameters were set as 2 match, 7 mismatch, and 7 indels. We identified repeats with condition of 80 minimum alignment score, maximum period sizes in 500 bp, and maximum TR array sizes of 2 million. Perl script MISA (MICroSATellite identification tool) was used to detect simple sequence repeats (SSRs) loci of the plastome, with a threshold of mono-, di-, tri-, tetra-, penta-, and hexa-nucleotide SSRs, respectively (Thiel et al., 2003).

Gene selection pressure analysis

To identify the gene divergences changes within seven species, nonsynonymous (K_a) and synonymous substitution rates (K_s), and the ratio K_a/K_s were calculated using the K_aK_s _calculator (Wang et al., 2010). We analysed all CDS gene regions (76 genes), except for *ndhC*, *ndhF* and *ndhK* as these regions were lost in *C. delavayi*, and *ycf15* as pseudogenes in *S. lyroglossa* and *P. rubens*. The seven species pairwise comparison generated 21 species pairs in total. The parameters were set as genetic code tab11 (bacterial and plant plastid code), method of calculation YN. When the genes had no substitutions on the alignment or there was 100% match, the K_s value was set as 0; in this case, the K_a/K_s value was shown as “NA” in the results. We replaced all “NA” results with 0.

We also evaluated the role of site-specific selection in 76 genes of seven species. The maximum likelihood tree was constructed based on the aligned concatenated CDS genes data set of the seven species using IQ-tree (Nguyen et al., 2015). We calculated the nonsynonymous (dN) substitution, synonymous (dS) substitution rates, and the dN/dS ratio (ω) in order to estimate the selection pressure with site-specific model (the option of the analyses was set to seqtype = 1, model = 0, and Nssites = 0, 1, 2, 3, 7, and 8) with CodeML program in PAML 4.9 (Yang & Nielsen, 2002; Yang, 2007). The likelihood ratio tests (LRTs) P -values of under three pairs of site models were calculated to detect positive selection ($p < 0.05$), including: M0 (one-ratio) vs. M3 (discrete), M1 (nearly neutral) vs. M2 (positive selection), and M7 (β) vs. M8 (β and ω).

Phylogenetic analysis

Phylogenetic analysis was conducted using plastome sequences of the seven *Calanthe* s.l. taxa mentioned above, with *Bletilla striata* (GenBank accession No. [KT588924](https://www.ncbi.nlm.nih.gov/nuccore/KT588924)) used as an outgroup. We generated four data sets for phylogenetic inference: (1) complete plastome sequences; (2) coding regions (CDS); (3) intergenic regions (IGS); and (4) 13 highly variable regions screened in the present study (*nad6* gene was lost in *Bletilla striata*, so *nad6-ndhI* was excluded in the analysis). We extracted the 76 CDS and 46 IGS shared by eight species from plastomes using PhyloSuit (Zhang et al., 2020). MAFFT v7.107 (Katoh & Standley, 2013) was conducted for sequence alignments.

Table 1 The basic characteristics of the plastome of eight *Calanthe* s.l. species.

Species	<i>Calanthe triplicata</i>	<i>Calanthe davidii</i> (D1)	<i>Calanthe davidii</i> (D2)	<i>Calanthe delavayi</i>	<i>Phaius tankervilleae</i>	<i>Cephalantheropsis obcordata</i>	<i>Styloglossum lyroglossa</i>	<i>Preptanthe rubens</i>
Accession number	KF753635	MG925365	MN708353	MK388860	MN708349	MN708351	MN708350	MN708352
Genome size(bp)	158,759	153,629	159,014	150,181	158,229	157,919	156,036	158,215
LSC length(bp)	87,305	86,045	87,857	83,411	86,638	86,650	85,421	87,498
SSC length(bp)	18,476	15,672	18,589	16,338	18,357	18,420	18,149	18,397
IR length(bp)	26,489	25,956	26,284	25,216	26,617	26,424	26,233	26,160
Coding(bp)	79,578	72,495	79,572	73,731	79,671	79,609	79,109	79,208
Non-Coding(bp)	79,181	81,134	79,442	76,450	78,558	78,310	76,927	79,007
Number of genes	136(115)	136(115)	136(115)	133(112)	136(115)	136(115)	136(115)	136(115)
Number of protein-coding genes	88(80)	88(80)	88(80)	85(77)	88(79)	88(80)	86(79)	86(79)
Number of tRNA genes	38(30)	38(30)	38(30)	38(30)	38(30)	38(30)	38(30)	38(30)
Number of rRNA genes	8(4)	8(4)	8(4)	8(4)	8(4)	8(4)	8(4)	8(4)
GC content (%)	36.70	36.90	36.60	36.90	37.00	36.80	36.90	36.70
GC content in LSC (%)	34.40	34.50	34.40	34.50	34.80	34.50	34.60	34.30
GC content in SSC (%)	29.70	30.20	29.60	29.40	29.90	29.70	29.90	29.70
GC content in IR (%)	43.00	43.10	43.10	43.30	43.00	43.10	43.10	43.30

Table 2 Genes length (bp) difference between two versions of *Calanthe davidii* complete plastome.

Version	<i>nad6</i>	<i>ndhA</i>	<i>ndhC</i>	<i>ndhD</i>	<i>ndhE</i>	<i>ndhF</i>	<i>ndhI</i>	<i>ndhK</i>	<i>ycf2</i>
D1 (MG925365)	120	869	171	987	306	1,827	297	138	4,650
D2 (MN708353)	531	2,215	363	1,509	321	2,259	504	774	6,813

A maximum likelihood (ML) tree was constructed using IQ-tree with 1,000 bootstrap replicates. The optimal nucleotide substitution model was found with ModelFinder module in IQ-tree (Nguyen et al., 2015). Bayesian inference (BI) was performed with MrBayes v3.2 (Ronquist et al., 2012). The Markov chain Monte Carlo (MCMC) analysis was run for 50,000,000 generations. The stationarity was regarded as having been reached when the average standard deviation of split frequencies remained below 0.01. Trees were sampled at every 1,000 generations with the first 25% discarded as burn-in. The remaining trees were used to build a 50% majority-rule consensus tree.

RESULTS

Features of the plastome

We reassembled and annotated the plastome of *C. davidii* (we use the abbreviations D1 (MG925365) and D2 (MN708353) to represent the different versions of the *C. davidii* plastome). In the present study, the reassembled version of *C. davidii* (D2) is hugely different from the published plastome. A comparison between these two genomes indicated that the gene numbers and orders are identical, while huge differences were found in the genome size, gene length, and GC content (Tables 1 and 2). We used the D2 version of *C. davidii* for the subsequent analysis.

The genome size of D1 is 5,385 bp shorter than D2, with the difference mainly concentrated in the IR region, especially *ndh* genes, which encode the subunits of the nicotinamide adenine dinucleotide (NADH) dehydrogenase-like complex proteins (Yamori & Shikanai, 2016). The assembling and annotating of D1 used *Dendrobium nobile* as the reference (Dong et al., 2018), which has a discrepancy with *C. triplicata* we used as reference in this study, especially *ndh*, e.g., the *ndhC*, *ndhI*, *ndhK*, and *nad6* genes were lost in *D. nobile* but exist in *C. triplicata*. Moreover, *D. nobile* showed a distant relationship with *C. davidii*, *Dendrobium* belong to Malaxideae tribe were far away from the Collabieae tribe, to which *C. davidii* belongs (Chase et al., 2015). Considering the conservatism of the plastome, and the reasons mentioned above, we may infer that the results of this study are more accurate.

The plastome of the seven species of *Calanthe* s.l. ranged from 150,181 bp (*C. delavayi*) to 159,014 bp (*C. davidii*, D2), with a GC content of 36.60%–37.00% (Table 1). They all shared the common structure: a pair of IRs (IRa and IRb) (25,216–26,617 bp), separated by LSC region (83,411–87,857 bp) and SSC region (16,338–18,589 bp) (Table 1 and Fig. 1). The gene numbers, orders, and names were very similar among *Calanthe* s.l., except for the three gene losses of *C. delavayi*, namely, *ndhC*, *ndhF*, and *ndhK* (Figs. 2–4).

The genomes encoded 133–136 genes, of which 112–115 were unique genes, containing 77–80 protein-coding genes, 30 tRNA genes, and four rRNA genes. Two pseudogenes were found in the seven plastomes; the *ycf15* was the only pseudogenization in *S. lyroglossa* and *P. rubens* while the *ycf68* gene was a pseudogene in all seven species. They both contain many internal stop codons. The 18 genes were found duplicated in the IR regions, including three type of genes, namely, coding genes, tRNA genes and rRNA genes (Table 3). In the seven plastomes, 15 genes contained one intron (six tRNA and nine protein-coding genes) and three genes (*ycf3*, *clpP*, and *rps12*) contained two introns. The gene *rps12* was trans-spliced, crossing the two areas with the 5'-end exon lie in the LSC region and the intron, 3'-end exon located in the IR region. Overlapping sequences were found in three pair of genes: *trnK*-UUU/*matK*, *atpE*/*atpB*, and *psbD*/*psbC*.

Comparative genomic analysis

The seven plastome sequence identities were plotted using mVISTA and with *C. triplicata* (KF753635) as reference. The alignment revealed high sequence similarity across the seven plastomes and no rearrangement occurred (Figs. 2 and 3) which suggested that they are highly conserved. The SSC region has the highest divergence ($\pi = 0.0187$) while IR region is the most conservative ($\pi = 0.0036$) (Table S3). We chose the nine mutational hotspots according to the sequence divergence analysis (nucleotide diversity ($\pi > 0.03$)), including 12 regions (*atpB-rbcL ccsA-ndhD*, *clpP-psbB*, *nad6-ndhI*, *psbA-matK*, *psbB-psbT*, *rpl32-trnL-UAG*, *rps11-rpl36-infA*, *rps16-trnQ-UUG*, *trnE-UUC-trnM-CAU*, *trnF-GAA-ndhJ*, *trnS-UGA-trnG-GC*), and two gene regions (*trnK*-UUU, *clpP*) (Figs. 3 and 4).

We examined the expansion and contraction of the IR area between the single-copy regions and the pair of IR regions for the seven *Calanthe* s.l. species (Fig. 5). Among the seven species, the gene positions of four borders (LSC/IRb, IRb/SSC, SSC/IRa, and IRa/LSC) had different types. The LSC/IRb border had three situations. First, in *C. triplicata*,

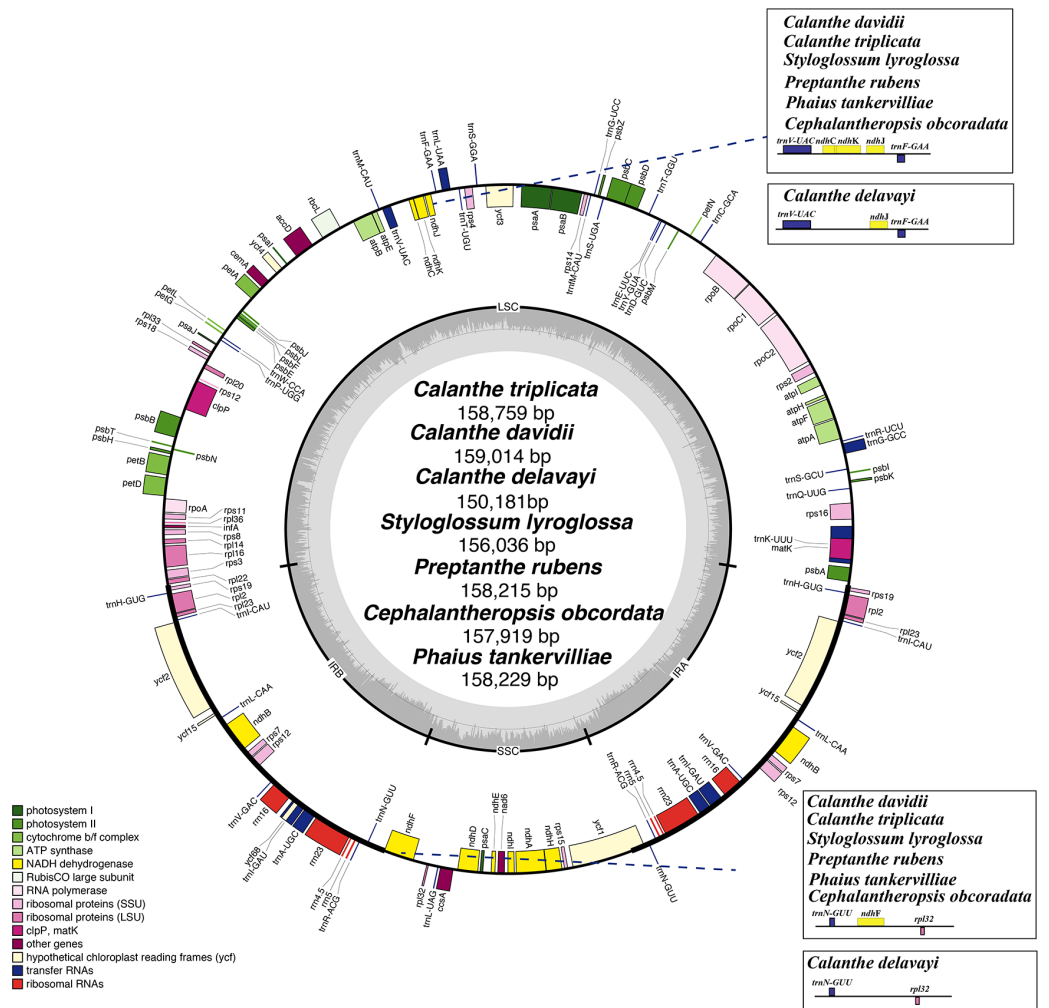


Figure 1 Plastome map of *Calanthe* s.l. Genes inside the circle are transcribed clockwise, and those outside are transcribed counterclockwise. Genes of different functions are color-coded. The darker gray in the inner circle shows the GC content, while the lighter gray shows the AT content.

Full-size [DOI: 10.7717/peerj.10051/fig-1](https://doi.org/10.7717/peerj.10051/fig-1)

C. delavayi, *P. tankervilleae*, and *C. obcordata*, the *rpl22* gene overlapped in the LSC/IRb region; second, in *C. davidii* and *S. lyroglossa* the *rpl22* gene was located in the LSC region, 9–47 bp away from the IRb region; the third situation is in *P. rubens*, where the *rps19* gene overlapped the border instead of the *rpl22* gene. While the IRb/SSC junction regions were relatively stable in the seven species, the *ndhF* gene crossed the border of six of the seven species, except for *C. delavayi*, due to its *ndhF* gene loss, and the nearest gene (*trnN* (in IRb)) is 361 bp away from the SSC region. SSC/IRa and IRa/LSC are both very conserved among the seven plastomes. The *ycf1* gene strode the SSC/IRa boundary, having 35–1,042 bp into the IRa region. The distance between *psbA* and the IRa/LSC junction ranged from 91 to 239 bp.

Repeats in the plastome were detected. *P. rubens* had the greatest number of long repeats and *C. triplicata* had the greatest number of tandem repeat regions (Tables S4 and

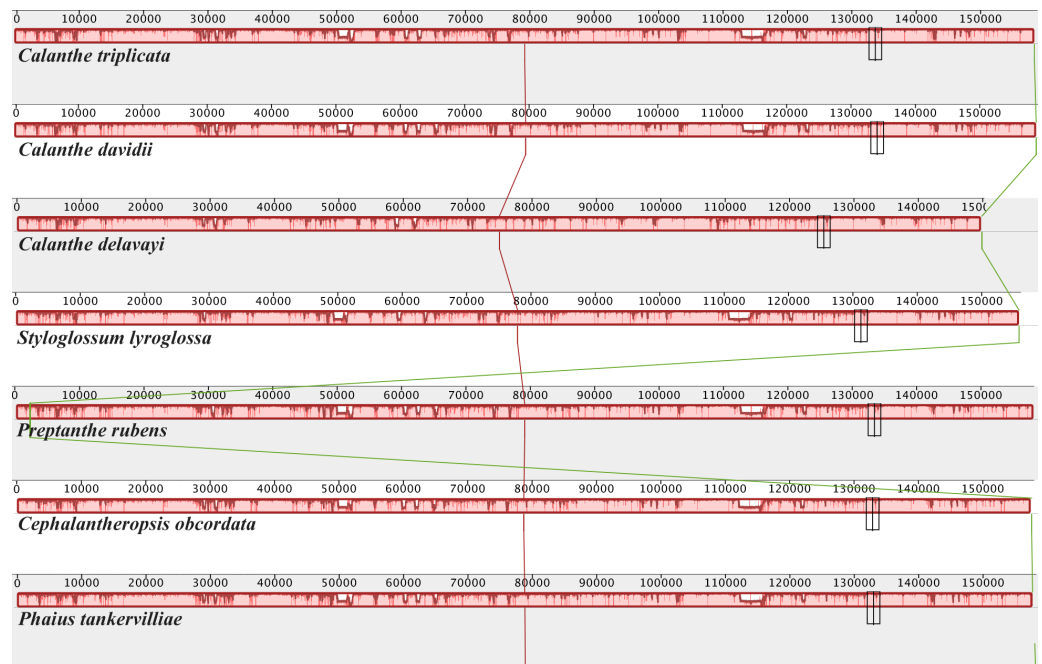


Figure 2 Mauve (Multiple Alignment of Conserved Genomic Sequence With Rearrangements) alignment of plastome of seven species of *Calanthe* s.l. The *Calanthe triplicata* genome is shown at the top as the reference genome.

Full-size DOI: 10.7717/peerj.10051/fig-2

S5). A total number of 49–73 SSRs were found in the seven plastomes. Mononucleotide, dinucleotide, trinucleotide, tetranucleotide, and pentanucleotide SSRs were all discovered in seven *Calanthe* s.l. species (Fig. 6). Hexanucleotide SSRs were found except in *C. davidii* and *P. rubens*. In all seven species, mononucleotide repetitions accounted for more than half of all (57.89%, 56.67%, 54.79%, 61.22%, 54.68%, 53.73%, and 60%). The IGS region contained the largest number of SSRs (with 231 identified), while 175 were identified in the CDS and 64 in the coding sequence introns. In particular, all mononucleotide SSRs belonged to A or T types, and the richness of A or T were found in the major of dinucleotide, trinucleotide, tetranucleotide, pentanucleotide, and hexanucleotide SSRs (Table S6).

Gene selection pressure analysis of protein sequences

The average Ka/Ks ratio for 76 protein-coding genes analysed in the seven genomes was 0.2455. We found the 13 most conserved genes with average Ka/Ks values between 0 and 0.01, *petG*, *psaC*, *psbE*, *psbI*, *psbJ*, *psbL*, *psbN*, *psbT*, *rpl23*, *rpl36*, *rps12*, *rps19* and *rps7*. These genes are under very strong purifying selection pressure. Values of Ka/Ks in the range of 0.5 to 1.0 (indicating relaxed selection) were observed for the genes *matK*, *atpF*, *rpoC2*, *accD*, *rps18*, *rpoA*, *ndhD*, *ndhI* and *rps15*. We found six protein-coding genes with Ka/Ks >1, including *cemA* (1.3680) in *C. obcordata*, *infA* (2.9934), *ycf1* (1.2618) and *ycf2* (1.0886) in *S. lyroglossa*, *nad6* (3.0761) and *ndhB* (1.4169) in *C. delavayi* (Table S7, and Fig. 7).

The site-specific selective pressure on seven species of *Calanthe* s.l. were assessed using the site model in the PAML program. Three pairs of site model comparisons (M0 vs. M3,

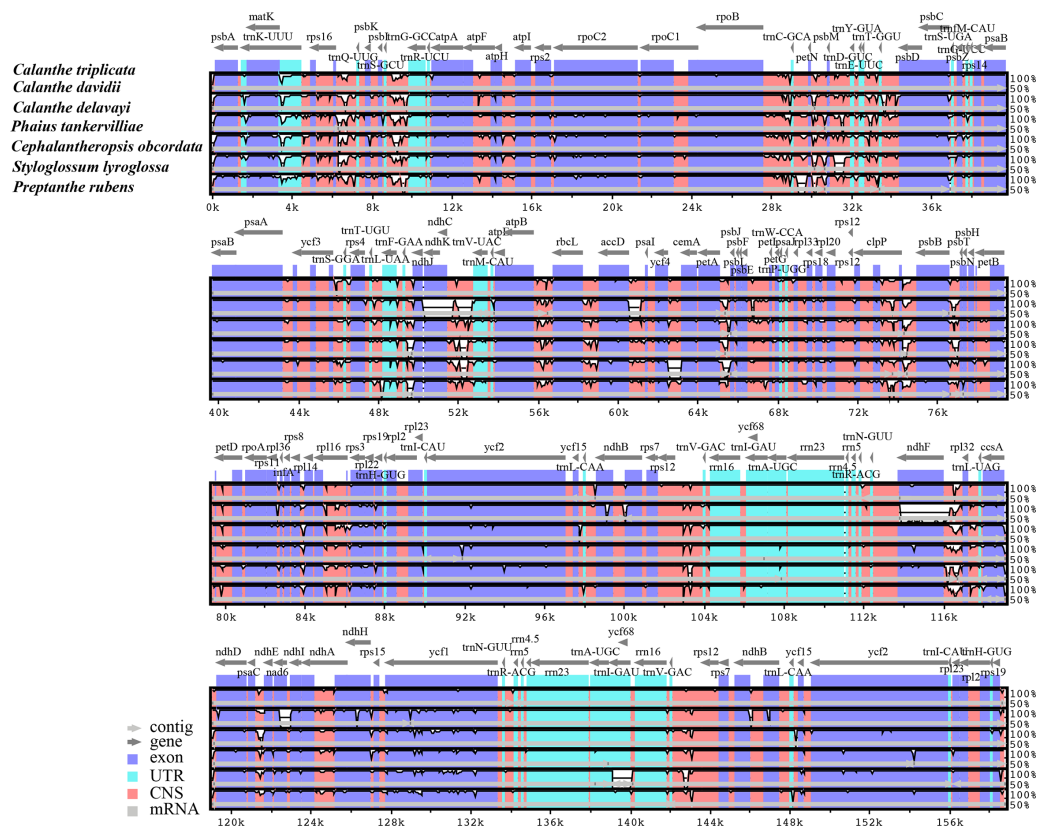


Figure 3 Comparison of seven *Calanthe* s.l. plastomes using mVISTA program, taking the annotation of *Calanthe triplicata* as a reference. The top line shows the genes in order. A cut-off of 70% identity was used for the plots and the Y-scale represents the percent identity between 50 and 100%. Genome regions are color-coded as exon and conserved non-coding sequences (CNS).

Full-size [DOI: 10.7717/peerj.10051/fig-3](https://doi.org/10.7717/peerj.10051/fig-3)

M1 vs. M2a, M7 vs. M8) showed that six genes were related to two photosynthetic electron transport (*ndhD* and *rbcl*), gene expression (*rpoC2*), and other functions (*accD*, *yef1* and *yef2*) had been subjected to positive selection (LRT of the three comparison all $P < 0.05$) (Tables S8 and S9).

Phylogenetic analysis

The phylogenetic relationship inferred by ML and BI analysis of the four data sets resulted in the same topology (Fig. 8). The seven *Calanthe* s.l. species were classified into three major clades, with all the *Calanthe* s.l. species composing a monophyly. *Preptanthe* is located at the basal position and is sister to the clade formed by the remaining other four genera. *Cephalantheropsis* and *Styloglossum* are clustered into a sister clade to *Phaius*, which forms the sister group to *Calanthe*.

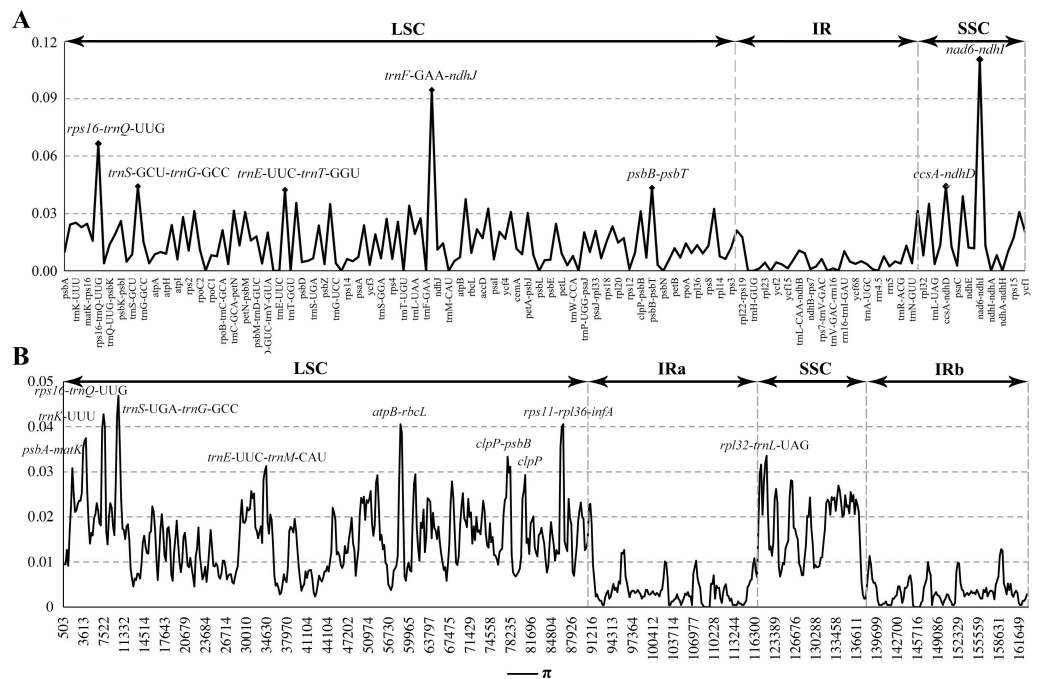


Figure 4 Comparison of nucleotide diversity (π) values among the seven *Calanthe* s.l. (A) The protein coding sequences (CDS) and intergenic spacers (IGS); (B) Sliding window analysis of the whole plastome of seven species.

Full-size DOI: [10.7717/peerj.10051/fig-4](https://doi.org/10.7717/peerj.10051/fig-4)

DISCUSSION

Chloroplast sequence variation

The whole plastome of five *Calanthe* s.l. species were determined, which had never been done for *Phaius*, *Cephalantheropsis*, *Styloglossum* and *Preptanthe* and was a reconstruction for *C. davidii*. Comparative analyses of the plastome of seven *Calanthe* s.l. species showed highly conserved structures and genes. The published Orchidaceae plastome genome size ranges from 35,304 bp (*Gastrodia elata*) to 178,131 bp (*Cypripedium formosanum*). This discrepancy in the plastome genome size can be explained by the different life forms, as the heterotrophic plants that do not photosynthesize have lost the photosynthesis-related genes. In the present study, only *C. delavayi* lost the *ndh* genes (*ndhC*, *F*, and *K*) in *Calanthe* s.l. It is also common to see a lack of functional *ndh* genes in other autotrophic orchids, i.e., *Cymbidium*, *Dendrobium*, *Phalaenopsis* and *Ophrys* (Yang et al., 2013; Lin et al., 2017; Roma et al., 2018). Although the transfer of the *ndh* genes from the plastome genome to the mitochondrial (mt) genome was detected in some orchids, there is no direct evidence showing that these transfers were linked to the losses of *ndh* in the plastome (Lin et al., 2015). The mechanisms for the complex deletion and truncation of the genes encoding NADH dehydrogenase subunits in orchids remain unclear.

We found two pseudogenes of the seven species, two hypothetical genes *ycf15* and *ycf68*. The *ycf15* was first identified as ORF87 in the *Nicotiana* chloroplast genome (Shinozaki et al., 1986); however, the validity of the gene as protein-coding gene has been questioned

Table 3 Gene contents in seven *Calanthe* s.l. species plastomes.

Classification	Genes
Genetic apparatus	
Large ribosomal subunits	<i>rpl2^a</i> (×2), <i>rpl14</i> , <i>rpl16^a</i> , <i>rpl20</i> , <i>rpl22</i> , <i>rpl23</i> (×2), <i>rpl32</i> , <i>rpl33</i> , <i>rpl36</i>
Small ribosomal subunits	<i>rps2</i> , <i>rps3</i> , <i>rps4</i> , <i>rps7</i> (×2), <i>rps8</i> , <i>rps11</i> , <i>rps12^b</i> , <i>rps14</i> , <i>rps15</i> , <i>rps16^a</i> , <i>rps18</i> , <i>rps19</i> (×2)
RNA polymerase subunits	<i>rpoA</i> , <i>rpoB</i> , <i>rpoC1^a</i> , <i>rpoC2</i>
DNA dependent RNA polymerase	<i>clpP^b</i>
Protease	
Maturase	<i>matK</i>
Ribosomal RNAs	<i>rrn4.5</i> (×2), <i>rrn5</i> (×2), <i>rrn23</i> (×2), <i>rrn16</i> (×2)
Transfer RNAs	<i>trnA</i> -UGC(×2) ^a , <i>trnC</i> -GCA, <i>trnD</i> -GUC, <i>trnE</i> -UUC, <i>trnF</i> -GAA, <i>trnM</i> -CAU, <i>trnG</i> -GCC, <i>trnG</i> -UCC ^a , <i>trnH</i> -GUG(×2), <i>trnI</i> -CAU(×2), <i>trnH</i> -GUG, <i>trnI</i> -GAU(×2) ^a , <i>trnK</i> -UUU ^a , <i>trnL</i> -CAA(×2), <i>trnL</i> -UAG, <i>trnL</i> -UUA ^a , <i>trnM</i> -CAU, <i>trnN</i> -GUU(×2), <i>trnP</i> -UGG, <i>trnQ</i> -UUG, <i>trnR</i> -ACG(×2), <i>trnR</i> -UCU, <i>trnS</i> -GCU, <i>trnS</i> -GGA, <i>trnS</i> -UGA, <i>trnT</i> -GGU, <i>trnV</i> -GAC(×2), <i>trnV</i> -UAC ^a , <i>trnW</i> -CCA, <i>trnY</i> -GUA
Light dependent photosynthesis	
Photosystem I	<i>psaA</i> , <i>psaB</i> , <i>psaC</i> , <i>psaI</i> , <i>psaJ</i> , <i>ycf3^b</i> , <i>ycf4</i>
Photosystem II	<i>psbA</i> , <i>psbB</i> , <i>psbC</i> , <i>psbD</i> , <i>psbE</i> , <i>psbF</i> , <i>psbH</i> , <i>psbI</i> , <i>psbJ</i> , <i>psbK</i> , <i>psbL</i> , <i>psbM</i> , <i>psbN</i> , <i>psbT</i> , <i>psbZ</i>
NAD(P)H dehydrogenase complex	<i>ndhA^a</i> , <i>ndhB^a</i> (×2), <i>ndhC^c</i> , <i>ndhD</i> , <i>ndhE</i> , <i>ndhF^c</i> , <i>ndhG</i> , <i>ndhH</i> , <i>ndhI</i> , <i>ndhJ</i> , <i>ndhK^c</i>
F-type ATP synthase	<i>atpA</i> , <i>atpB</i> , <i>atpE</i> , <i>atpF^a</i> , <i>atpH</i> , <i>atpI</i>
Cytochrome b6/f complex	<i>petA</i> , <i>petB^a</i> , <i>petD^a</i> , <i>petG</i> , <i>petL</i> , <i>petN</i>
Light independent photosynthesis	
Inner membrane protein	<i>cemA</i>
Cytochrome C biogenesis protein	<i>ccsA</i>
Large subunit of Rubisco	<i>rbcL</i>
Subunit of acetyl-CoA-carboxylase	<i>accD</i>
Translation initiation factor	<i>infA</i>
Function uncertain	<i>ycf1</i> , <i>ycf2</i> (×2), <i>ycf15</i> (×2), <i>ycf68</i> (×2)

Notes.^aGene containing one intron.^bGene containing two introns, a trans-splinting gene, (2) shows genes have two copies.^cGene lost in *C. delavayi*.

(Chumley et al., 2006). It was disabled in *Amoborella*, *Nuphar* and most orchids (Goremykin et al., 2003; Raubeson et al., 2007). In some lineages it has been completely lost, for example *Illicium*, *Acorus*, *Ceratophyllum*, *Ranunculus* and some non-photosynthetic orchids (Shi et al., 2013; Yuan et al., 2018). The likelihood that the *ycf15* and *ycf68* are not protein-coding genes has been discussed in previous research (Raubeson et al., 2007). In the present study, *ycf15* was annotated as a protein-coding gene in five other species (*C. triplicata*, *C. davidii*, *C. delavayi*, *C. obcordata* and *P. tankervilleae*). Whether these sequences represent protein-coding regions and their possible functions in chloroplasts needs to be further studied.

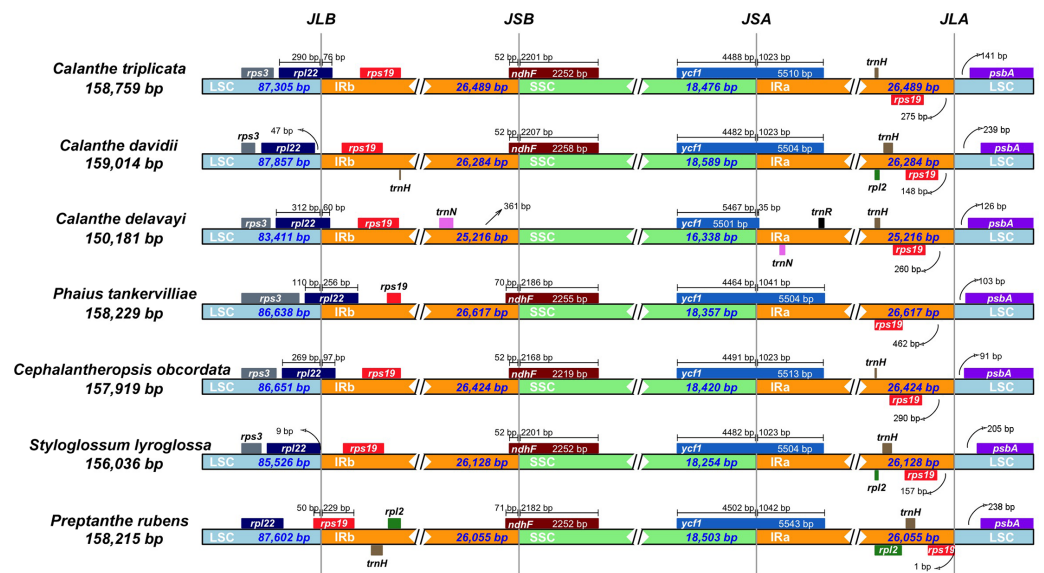


Figure 5 Comparison of the borders of LSC, SSC, and IR regions in seven *Calanthe* s.l. complete chloroplast genomes. JLB (IRb/LSC), JSB (IRb/SSC), JSA (SSC/IRa) and JLA (IRa/LSC) denote the Js between each corresponding region in the genome.

Full-size [DOI: 10.7717/peerj.10051/fig-5](https://doi.org/10.7717/peerj.10051/fig-5)

Although the overall genomic structures and gene orders of *Calanthe* s.l. are highly conserved, significant differences at the IR/SSC junction area were detected. The different expansion and contraction situation of the IR junction area will cause the plastome size differentiation (Bock & Knoop, 2012). In present study, among the four boundaries, only LSC/IRb shows three types in the seven species, while the remaining three (IRb/SSC, SSC/IRa, and IRa/LSC) are conservative and stable. Because of the *ndhF* gene loss in *C. delavayi*, IR contraction was detected. Previous research has pointed out that the deletion of the *ndh* genes has great influence on the instability of the IR/SSC boundary in orchids (Kim et al., 2015; Niu et al., 2017). The position of the boundary, especially the expansion and contraction of the region, could shed light on the evolution of the lineage. However, in *Calanthe* s.l., our observations would not provide the information required to elucidate the evolutionary relationships of the taxa, and whether it can benefit from adaptation require further investigation; thus, additional sampling of *Calanthe* spp. and related genera will allow for explicit tests.

Molecular markers and candidate SSRs

Coding regions and conserved sequences of the plastome are widely used for phylogenetic inferences at higher taxonomic levels (family or genus) (Givnish et al., 2015; Jansen et al., 2007). Plastomes are ideal resources for selecting the mutational hotspots of various lineages and are used for intraspecies discrimination and phylogenetic studies at the species level (Ahmed et al., 2013; Liu et al., 2018). At present, some plastid DNA fragments are used in the taxonomy of *Calanthe* s.l., for instance *matK*, *rbcL*, and intergenic spacer *trnL-trnF* (Zhai et al., 2014; Guo et al., 2017), but they could not provide sufficient phylogeny signals

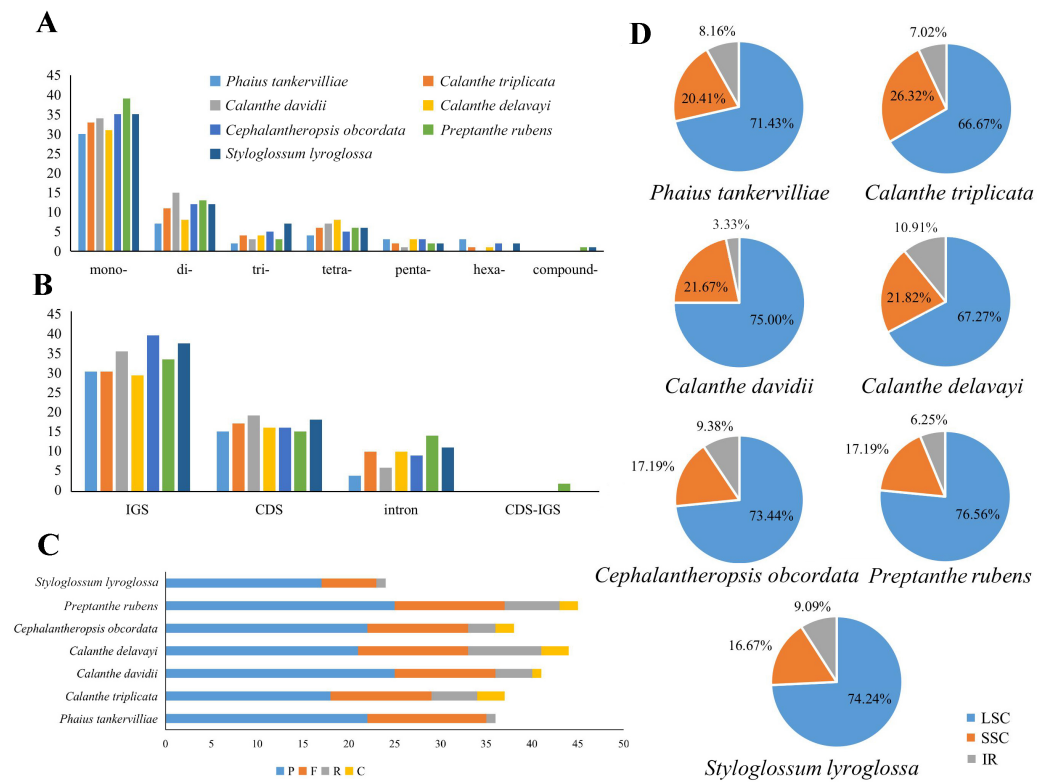


Figure 6 Maps of repeat sequence analyses in seven *Calanthe* s.l. species plastome genome. (A) Classification of SSRs by repeat type. mono-, mononucleotides; di-, dinucleotides; tri-, trinucleotides; tetra-, tetranucleotides; penta-, pentanucleotides; and hexa-, hexanucleotides. compound-, compound formation. (B) Classification of SSRs in seven species. IGS, intergenic spacer; CDS, coding sequence, CDS-IGS, part in CDS and part in IGS. (C) Number of the four repeat types, F, P, R, and C, indicate the long repeat type (F: forward, P: palindrome, R: reverse, and C: complement, respectively). (D) SSRs locus distribution among three different regions.

Full-size DOI: 10.7717/peerj.10051/fig-6

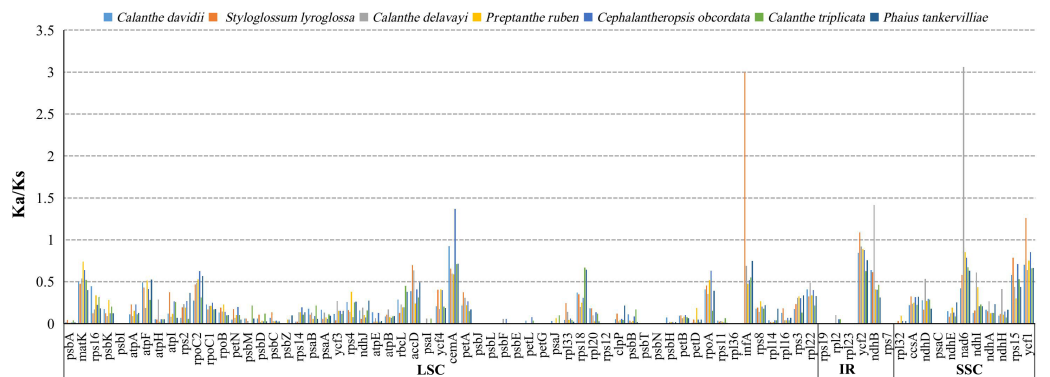


Figure 7 The Ka/Ks ratio of 76 protein-coding genes of seven *Calanthe* s.l. plastomes.

Full-size DOI: 10.7717/peerj.10051/fig-7

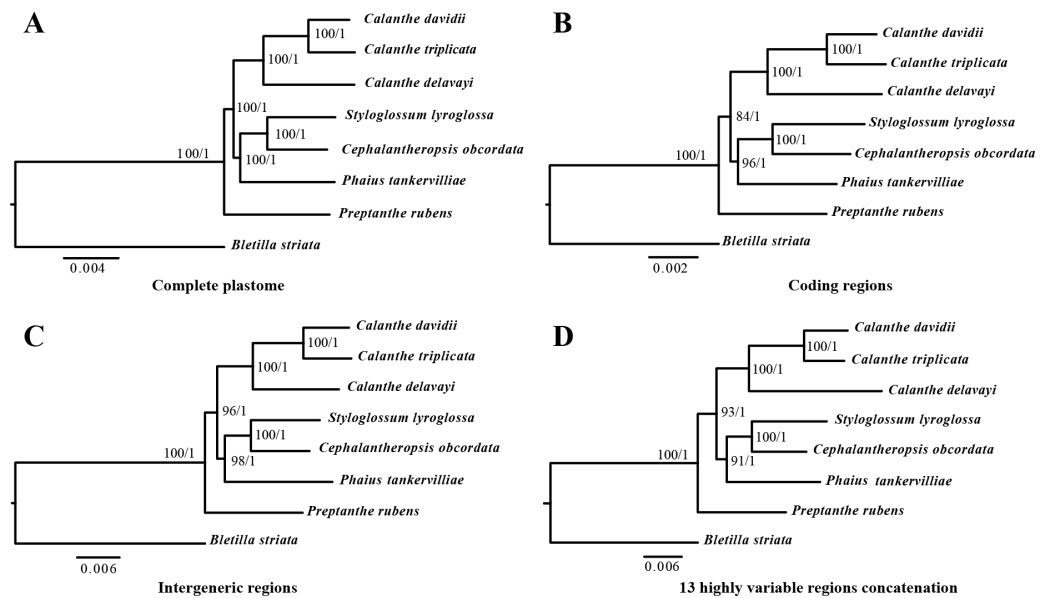


Figure 8 Phylogenetic relationships of the seven *Calanthe* s.l. taxon constructed by four DNA data sets. Including (A) whole plastome sequences, (B) coding regions, (C) intergenic regions, and the (D) 13 highly variable regions concatenation with Maximum likelihood (ML), and Bayesian inference (BI) methods. ML topology shown with ML bootstrap support value/Bayesian posterior probability listed at each node.

Full-size DOI: [10.7717/peerj.10051/fig-8](https://doi.org/10.7717/peerj.10051/fig-8)

to establish the high-resolution phylogeny relationship for classification of related taxa, especially some infrageneric taxa whose taxonomic classification status are unclear. Our alignment screened the top 14 loci that most likely contained the highest degrees of genetic variability in *Calanthe* s.l., which can be useful in species-level phylogenetic studies of *Calanthe* s.l..

Simple sequence repeats are short (1–5 bp) repeat motifs that are tandemly repeated for varying numbers of times (Kantety *et al.*, 2002). Because of its extensively dispersal in genomes, they are widely utilized in population genetics and molecular evolution studies (Guichoux *et al.*, 2011). SSRs can provide interspecific polymorphisms, which are effective markers in population genetic analysis. We identified 18 SSRs as polymeric SSRs between *Calanthe* s.l. species; however, further experiments are needed to verify its effectiveness (Table 4). Mutational hotspots and SSRs derived from the plastome can serve as valuable tools for elucidating the evolutionary relationships and plant identification of *Calanthes*.l..

Adaptive evolution analysis

Calanthe s.l. are pantropical in their distribution, with high geographic and ecological diversity, from obligate epiphytic to hemi-epiphytic and terrestrial, ranging from sea level to alpine mountain areas. However, the majority live under tropical woods and forests, very often in deep shade (Clayton & Cribb, 2013; Stone & Cribb, 2017). To better understand the evolutionary history of these groups, the analysis of the genetic diversity and adaptive evolution of *Calanthe* s.l. is essential. Positive selection genes played an important part in the adaption to various environments. Within the seven species, we detected the six

Table 4 The polymorphic simple sequence repeats in *Calanthe* s.l. plastomes.

Type	<i>C. triplicata</i> / <i>C. davidii</i> / <i>C. delavayi</i> / <i>S. lyroglossa</i> / <i>P. rubens</i> / <i>C. obcordata</i> / <i>P. tankervilleae</i>	Location	Region
AT	6/6/0/6/0/0/0	<i>rpoB-trnC</i> -GCA	LSC
AT	7/7/0/0/5/0/0	<i>trnE</i> -UUC- <i>trnT</i> -GGU	LSC
AT	0/0/7/6/5/7/0	<i>trnL</i> -UAA- <i>trnF</i> -GAA	LSC
AT	0/0/7/6/0/7/0	<i>trnL</i> -UAA- <i>trnF</i> -GAA	LSC
GA	5/5/5/0/5/5/5	<i>ycf2</i>	IR
TA	5/7/0/5/0/6/0	<i>psbB-psbT</i>	LSC
TA	5/5/0/0/6/0/0	<i>ndhF-rpl32</i>	SSC
TA	0/0/8/5/5/0/5	<i>clpP-psbB</i>	LSC
TC	5/5/5/0/5/5/5	<i>ycf2</i>	IR
TG	5/5/5/0/5/5/5	<i>rpl33-rps18</i>	LSC
AAT	4/4/4/4/0/4/0	<i>psaC-ndhE</i>	SSC
AGAA	0/0/3/0/3/3/0	<i>psbM-trnD</i> -GUC	LSC
AATG	0/3/3/3/0/3/0	<i>cemA</i>	LSC
ATTA	3/3/0/3/3/0/0	<i>psaJ-rpl33</i>	LSC
GTCT	3/3/3/3/0/3/3	<i>atpA</i>	LSC
TTGA	3/3/3/3/0/3/3	<i>ndhE</i>	SSC
ATCTT	3/3/0/0/3/3/3	<i>psbK-psbI</i>	LSC
ACAAA	3/0/0/0/0/3/3	<i>ndhC-trnV</i> -UAC	LSC

species-specific positive selection genes in only three species, namely, one in *C. obcordata* (*cemA*), three in *S. styloglossa* (*infA*, *ycf1* and *ycf2*), and two in *C. delavayi* (*nad6* and *ndhB*). Those identified selected genes may have undergone certain functional diversification during their evolutionary history.

cemA encodes a chloroplast envelope membrane protein (Sasaki et al., 1993) and is inferred to indirectly influence CO₂ uptake in plastid (Rolland et al., 1997). In some non-photosynthetic orchids, *cemA* exists as pseudogene or is lost (Feng et al., 2016; Kim et al., 2020). *infA* is crucial for genetic information transmission that affects transcription of DNA into RNA and translation of RNA to protein; it is seen as a housekeeping gene (Schelkunov et al., 2015). Two genes with large ORFs encoding proteins of the unknown function, *ycf1* and *ycf2*. *nad6* gene encoding NADH-plastoquinone oxidoreductase subunit 6, *ndhB* gene encoding NAD(P)H-quinone oxidoreductase subunit 2 can influence the cyclic electron flow around photosystem I (Shikanai et al., 1998). They are both NAD(P)H-plastoquinone-oxidoreductase (NDH complex) encoding-related genes. Chloroplast NAD(P)H dehydrogenase is sensitive to strong light stress and can protect plants from photoinhibition or photooxidation stress and alleviate decreases in the photosynthetic rate and growth delay caused by drought, suggesting that the *ndh* genes encoding the NAD(P)H dehydrogenase (NDH) may also be involved in stress acclimation through the optimization of photosynthesis (Horváth et al., 2000; Rumeau, Peltier & Cournac, 2007). *C. delavayi* is one of species with the highest elevation distribution in the *Calanthe* s.l. (up to 3500 m). The positive selection signal of *ndhB* and *nad6* in the *C. delavayi* might be the result of adaptation to different environment with other six species.

Six genes were under positive selection in seven *Calanthe* s.l. species. Of these genes, one functions as a subunit of acetyl-CoA-carboxylase (*accD*), two NADH dehydrogenase subunit genes (*ndhB* and *ndhD*), one RNA polymerase subunit (*rpoC2*), and two genes whose functions are uncertain (*ycf1* and *ycf2*). Except for the *ndh* genes, the remaining four positively selected genes were also detected in orchid species (Dong et al., 2018), and these four genes may have played significant roles in the adaptive evolution history of *Calanthe* s.l. Specific roles need to be further studied.

The function of the plastid *accD* gene has been reported as essential for plant leaf development (Kode et al., 2005). We detected the two positively selected sites in *accD* genes for *Calanthe* s.l.. The *ndh* gene family in the plastome is involved in photosynthesis. The *ndhB* and *ndhD* genes possessed 12 and 10 positively selected sites, respectively. Although there is the viewpoint that NDH activity may not be required in some plants (Yang et al., 2013), these two genes may play important roles in the adaptation of *Calanthe* s.l. species to deep shade environments. In addition, the *rpoC2* (RNA polymerase subunit C2) gene was crucial for gene transcription (Xie, Jäger & Potts, 1989), and only one site was positively detected in our study. Plastomes contain a number of uncertain genes. The *ycfs* (hypothetical open reading frame) gene has great application potential for elucidating plant phylogeny (Neubig et al., 2009; Huang, Sun & Zhang, 2010; Dong et al., 2015). As the largest genes in *ycfs*, *ycf1* and *ycf2* have shown positive selection in one and eight sites, respectively, and the phenomenon extends to the many plant lineages, including orchids (Greiner et al., 2008; Huang, Sun & Zhang, 2010; Carbonell-Caballero et al., 2015).

Phylogenetic analysis

Plastid sequences have been used in phylogenetic analyses based on their most nonrecombinant and uniparentally inherited and on their slower evolutionary rates than nuclear and mitochondrial genomes (Wolfe, Li & Sharp, 1987; Birky, 1995). The plastid region of *matK*, *rbcL* and *trnL-F* have been used as genetic markers with great success in Orchidaceae (Cameron et al., 1999; Salazar et al., 2017; Hu et al., 2020). However, the limited loci in phylogenetic inference are not powerful enough when closely related species are under consideration. Phylogenetic analyses based on plastome data sets have become a popular and practical approach and therefore comparative genomic studies of more plastome sequences have become necessary.

In the present study, the relationships among seven *Calanthe* s.l. species all had high support value (bootstrap support value > 90 and Bayesian posterior probability > 0.95), and they can be completely distinguished from one another (Fig. 8). The seven species were separated into three evolutionary branches. The branch including *C. davidii*, *C. triplicata* and *C. delavayi* was a sister to branch containing *S. lyroglossa*, *C. obcordata* and *P. tankervilleae*, and *P. rubens* was placed at the basal position. We obtained high resolution using the nine highly variable regions with lower cost.

CONCLUSIONS

In this study, we assembled and analysed five new complete plastome sequences of *Calanthe* s.l. and compared them with other *Calanthe* species for the first time. The annotation and

comparison within *Calanthe* s.l. species showed conservative in gene sequence, GC content and genomic composition. The repeated sequences, 18 microsatellites and 14 highly mutational hotspot regions were identified in the *Calanthe* s.l. plastome. Six site-specific positively selected genes were detected. These genes will lead to understanding of the adaptations of *Calanthe* s.l. species to deep shade environments. The study will help to resolve the phylogenetic relationships and understand the adaptive evolution of *Calanthe*. It will also provide genomic resources and potential markers suitable for future species identification and speciation studies of the genus.

ADDITIONAL INFORMATION AND DECLARATIONS

Funding

This research was funded by Government Business Entrustment Project of National Forestry and Grassland Administration, grant number 2019073046 and the Key Laboratory of National Forestry and Grassland Administration for Orchid Conservation and Utilization Construction Funds, grant number 115/118990050 & 115/KJG18016A. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Grant Disclosures

The following grant information was disclosed by the authors:

National Forestry and Grassland Administration: 2019073046.

Key Laboratory of National Forestry and Grassland Administration for Orchid Conservation and Utilization Construction Funds: 115/118990050, 115/KJG18016A.

Competing Interests

The authors declare there are no competing interests.

Author Contributions

- Yanqiong Chen conceived and designed the experiments, performed the experiments, analyzed the data, prepared figures and/or tables, and approved the final draft.
- Hui Zhong performed the experiments, analyzed the data, prepared figures and/or tables, and approved the final draft.
- Yating Zhu and Yuanzhen Huang analyzed the data, prepared figures and/or tables, and approved the final draft.
- Shasha Wu, Zhongjian Liu, Siren Lan and Junwen Zhai conceived and designed the experiments, authored or reviewed drafts of the paper, and approved the final draft.

Data Availability

The following information was supplied regarding data availability:

The five chloroplast genome sequences are available at GenBank: [MN708349](#) to [MN708353](#).

Supplemental Information

Supplemental information for this article can be found online at <http://dx.doi.org/10.7717/peerj.10051#supplemental-information>.

REFERENCES

- Ahmed I, Matthews PJ, Biggs PJ, Naeem M, McLenachan PA, Lockhart PJ. 2013.** Identification of chloroplast genome loci suitable for high-resolution phylogeographic studies of *Colocasia esculenta* (L.) S chott (A raceae) and closely related taxa. *Molecular Ecology Resources* **13**(5):929–937 DOI [10.1111/1755-0998.12128](https://doi.org/10.1111/1755-0998.12128).
- Amiryousefi A, Hyvönen J, Poczai P. 2018.** IRscope: an online program to visualize the junction sites of chloroplast genomes. *Bioinformatics* **34**(17):3030–3031 DOI [10.1093/bioinformatics/bty220](https://doi.org/10.1093/bioinformatics/bty220).
- Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Prjibelski AD, Pyshkin AV, Sirotkin AV, Vyahhi N, Tesler G, Alekseyev MA, Pevzner PA. 2012.** SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *Journal of Computational Biology* **19**(5):455–477 DOI [10.1089/cmb.2012.0021](https://doi.org/10.1089/cmb.2012.0021).
- Barrett CF, Kennedy AH. 2018.** Plastid genome degradation in the endangered, myco-heterotrophic, North American orchid *Hexalectris warnockii*. *Genome Biology and Evolution* **10**(7):1657–1662 DOI [10.1093/gbe/evy107](https://doi.org/10.1093/gbe/evy107).
- Benson G. 1999.** Tandem repeats finder: a program to analyze DNA sequences. *Nucleic Acids Research* **27**(2):573–580 DOI [10.1093/nar/27.2.573](https://doi.org/10.1093/nar/27.2.573).
- Bentham G. 1881.** Notes on Orchidaceae. *Botanical journal of the Linnean Society* **18**:281–360 DOI [10.1111/j.1095-8339.1881.tb01258.x](https://doi.org/10.1111/j.1095-8339.1881.tb01258.x).
- Birky CW. 1995.** Uniparental inheritance of mitochondrial and chloroplast genes: mechanisms and evolution. *Proceedings of the National Academy of Sciences of the United States of America* **92**(25):11331–11338 DOI [10.1073/pnas.92.25.11331](https://doi.org/10.1073/pnas.92.25.11331).
- Bock R, Knoop V (eds.) 2012.** *Genomics of chloroplasts and mitochondria*. Vol. 35. Dordrecht: Springer Science & Business Media.
- Brudno M, Malde S, Poliakov A, Do CB, Couronne O, Dubchak I, Batzoglu S. 2003.** Glocal alignment: finding rearrangements during alignment. *Bioinformatics* **19**(suppl_1):i54–i62 DOI [10.1093/bioinformatics/btg1005](https://doi.org/10.1093/bioinformatics/btg1005).
- Cameron KM, Chase MW, Whitten WM, Kores PJ, Jarrell DC, Albert VA, Goldman DH. 1999.** A phylogenetic analysis of the Orchidaceae: evidence from *rbcl* nucleotide sequences. *American Journal of Botany* **86**(2):208–224 DOI [10.2307/2656938](https://doi.org/10.2307/2656938).
- Carbonell-Caballero J, Alonso R, Ibañez V, Terol J, Talon M, Dopazo J. 2015.** A phylogenetic analysis of 34 chloroplast genomes elucidates the relationships between wild and domestic species within the genus *Citrus*. *Molecular Biology and Evolution* **32**(8):2015–2035 DOI [10.1093/molbev/msv082](https://doi.org/10.1093/molbev/msv082).
- Chang CC, Lin HC, Lin IP, Chow TY, Chen HH, Chen WH, Cheng CH, Lin CY, Liu SM, Chang CC, Chaw SM. 2006.** The chloroplast genome of *Phalaenopsis aphrodite* (Orchidaceae): comparative analysis of evolutionary rate with that of grasses and

- its phylogenetic implications. *Molecular Biology and Evolution* **23**(2):279–291 DOI [10.1093/molbev/msj029](https://doi.org/10.1093/molbev/msj029).
- Chase MW, Cameron KM, Freudenstein JV, Pridgeon AM, Salazar G, Van den Berg C, Schuiteman A. 2015. An updated classification of Orchidaceae. *Botanical Journal of the Linnean Society* **177**(2):151–174 DOI [10.1111/boj.12234](https://doi.org/10.1111/boj.12234).
- Chen SC, Liu ZJ, Zhu GH, Lang KY, Ji ZH, Luo YB, Jin XH, Cribb PJ, Wood JJ, Gale SW, Ormerod P, Vermeulen JJ, Wood HP, Clayton D, Bell A. 2009. *Flora of China*, 25. St. Louis: Science Press, Beijing Missouri Botanical Garden Press, 1–9.
- Christenhusz MJ, Byng JW. 2016. The number of known plants species in the world and its annual increase. *Phytotaxa* **261**(3):201–217 DOI [10.11646/phytotaxa.261.3.1](https://doi.org/10.11646/phytotaxa.261.3.1).
- Chumley TW, Palmer JD, Mower JP, Fourcade HM, Calie PJ, Boore JL, Jansen RK. 2006. The complete chloroplast genome sequence of *Pelargonium* × *hortorum*: organization and evolution of the largest and most highly rearranged chloroplast genome of land plants. *Molecular Biology and Evolution* **23**(11):2175–2190 DOI [10.1093/molbev/msl089](https://doi.org/10.1093/molbev/msl089).
- Clayton D, Cribb P. 2013. *The Genus Calanthe*. Natural History Publications, Malaysia.
- Darling AC, Mau B, Blattner FR, Perna NT. 2004. Mauve: multiple alignment of conserved genomic sequence with rearrangements. *Genome Research* **14**(7):1394–1403 DOI [10.1101/gr.2289704](https://doi.org/10.1101/gr.2289704).
- Dong WL, Wang RN, Zhang NY, Fan WB, Fang MF, Li ZH. 2018. Molecular evolution of chloroplast genomes of orchid species: insights into phylogenetic relationship and adaptive evolution. *International Journal of Molecular Sciences* **19**(3):716 DOI [10.3390/ijms19030716](https://doi.org/10.3390/ijms19030716).
- Dong WP, Xu C, Li CH, Sun JH, Zuo YJ, Shi S, Cheng T, Guo JJ, Zhou SL. 2015. *ycf1*, the most promising plastid DNA barcode of land plants. *Scientific Reports* **5**:8348 DOI [10.1038/srep08348](https://doi.org/10.1038/srep08348).
- Doyle JJ, Doyle JL. 1987. A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochemical Bulletin* **19**:11–15 DOI [10.1016/j.bse.2009.07.003](https://doi.org/10.1016/j.bse.2009.07.003).
- Feng YL, Wicke S, Li JW, Han Y, Lin CS, Li DZ, Zhou TT, Huang WC, Huang LQ, Jin XH. 2016. Lineage-specific reductions of plastid genomes in an orchid tribe with partially and fully mycoheterotrophic species. *Genome Biology and Evolution* **8**(7):2164–2175 DOI [10.1093/gbe/evw144](https://doi.org/10.1093/gbe/evw144).
- Frazer KA, Pachter L, Poliakov A, Rubin EM, Dubchak I. 2004. VISTA: computational tools for comparative genomics. *Nucleic Acids Research* **32**(suppl_2):W273–W279 DOI [10.1093/nar/gkh458](https://doi.org/10.1093/nar/gkh458).
- Ker Gawler JB. 1821. Botanical register consisting of coloured figures of exotic plants cultivated in british gardens, with their history and mode of treatment. In: *sub t.* vol. 7. London, 573.
- Givnish TJ, Spalink D, Ames M, Lyon SP, Hunter SJ, Zuluaga A, Iles WJD, Clements MA, Arroyo MTK, Leebens-Mack J, Endara L, Kriebel R, Neubig KM, Whitten WM, Williams NH, Cameron KM. 2015. Orchid phylogenomics and multiple drivers of their extraordinary diversification. *Proceedings of the Royal Society B: Biological Sciences* **282**(1814):201515 DOI [10.1098/rspb.2015.1553](https://doi.org/10.1098/rspb.2015.1553).

- Goremykin VV, Hirsch-Ernst KI, Wölfl S, Hellwig FH. 2003.** Analysis of the *Amborella trichopoda* chloroplast genome sequence suggests that *Amborella* is not a basal angiosperm. *Molecular Biology and Evolution* **20**(9):1499–1505 DOI [10.1093/molbev/msg159](https://doi.org/10.1093/molbev/msg159).
- Greiner S, Lehwark P, Bock R. 2019.** OrganellarGenomeDRAW (OGDRAW) version 1.3.1. *Nucleic Acids Research* **47**(W1):W59–W64 DOI [10.1093/nar/gkz238](https://doi.org/10.1093/nar/gkz238).
- Greiner S, Wang X, Rauwolf U, Silber MV, Mayer K, Meurer J, Haberer G, Herrmann RG. 2008.** The complete nucleotide sequences of the five genetically distinct plastid genomes of *Oenothera*, subsection *Oenothera*: I. Sequence evaluation and plastome evolution. *Nucleic Acids Research* **36**(7):2366–2378 DOI [10.1093/nar/gkn081](https://doi.org/10.1093/nar/gkn081).
- Guichoux E, Lagache L, Wagner S, Chaumeil P, Léger P, Lepais O, Lepais O, Lepoittevin C, Malausa T, Revardel E, Salin F, Petit RJ. 2011.** Current trends in microsatellite genotyping. *Molecular Ecology Resources* **11**(4):591–611 DOI [10.1111/j.1755-0998.2011.03014.x](https://doi.org/10.1111/j.1755-0998.2011.03014.x).
- Guo M, Zhai JW, Wu XY, Li SX, Wang M, Li ZJ, Chen LJ. 2017.** *Calanthe taibaishanensis*, a new orchid species from China: evidence from morphological and molecular analyses. *Phytotaxa* **327**(2):184–190 DOI [10.11646/phytotaxa.327.2.7](https://doi.org/10.11646/phytotaxa.327.2.7).
- Horváth EM, Peter SO, Joët T, Rumeau D, Cournac L, Horváth GV, Kavanagh TA, Schäfer C, Peltier G, Medgyesy P. 2000.** Targeted inactivation of the plastid *ndhB* gene in tobacco results in an enhanced sensitivity of photosynthesis to moderate stomatal closure. *Plant Physiology* **123**(4):1337–1350 DOI [10.1104/pp.123.4.1337](https://doi.org/10.1104/pp.123.4.1337).
- Hu AQ, Gale SW, Liu ZJ, Suddee S, Hsu TC, Fischer GA, Saunders RM. 2020.** Molecular phylogenetics and floral evolution of the *Cirrhopetalum* alliance (*Bulbophyllum* Orchidaceae): evolutionary transitions and phylogenetic signal variation. *Molecular Phylogenetics and Evolution* **143**:106689 DOI [10.1016/j.ympev.2019.106689](https://doi.org/10.1016/j.ympev.2019.106689).
- Huang JL, Sun GL, Zhang DM. 2010.** Molecular evolution and phylogeny of the angiosperm *ycf2* gene. *Journal of Systematics and Evolution* **48**(4):240–248 DOI [10.1111/j.1759-6831.2010.00080.x](https://doi.org/10.1111/j.1759-6831.2010.00080.x).
- Jansen RK, Cai Z, Raubeson LA, Daniell H, dePamphilis CW, Leebens-Mack J, Müller KF, Guisinger-Bellian M, Haberle RC, Hansen AK, Chumley TW, Lee S, Peery R, McNeal JR, Kuehl JV, Boore JL. 2007.** Analysis of 81 genes from 64 plastid genomes resolves relationships in angiosperms and identifies genome-scale evolutionary patterns. *Proceedings of the National Academy of Sciences of the United States of America* **104**(49):19369–19374 DOI [10.1073/pnas.0709121104](https://doi.org/10.1073/pnas.0709121104).
- Jin JJ, Yu WB, Yang JB, Song Y, de Pamphilis CW, Yi TS, Li DZ. 2019.** GetOrganelle: a fast and versatile toolkit for accurate de novo assembly of organelle genomes. *BioRxiv* 256479 DOI [10.1101/256479](https://doi.org/10.1101/256479).
- Kantety RV, La Rota M, Matthews DE, Sorrells ME. 2002.** Data mining for simple sequence repeats in expressed sequence tags from barley, maize, rice, sorghum and wheat. *Plant Molecular Biology* **48**(5–6):501–510 DOI [10.1023/A:1014875206165](https://doi.org/10.1023/A:1014875206165).
- Katoh K, Standley DM. 2013.** MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Molecular Biology and Evolution* **30**(4):772–780 DOI [10.1093/molbev/mst010](https://doi.org/10.1093/molbev/mst010).

- Kearse M, Moir R, Wilson A, Stones-Havas S, Cheung M, Sturrock S, Buxton S, Cooper A, Markowitz S, Duran C, Thierer T, Ashton B, Meintjes P, Drummond A. 2012.** Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics* **28**(12):1647–1649 DOI [10.1093/bioinformatics/bts199](https://doi.org/10.1093/bioinformatics/bts199).
- Kim HT, Kim JS, Moore MJ, Neubig KM, Williams NH, Whitten WM, Kim JH. 2015.** Seven new complete plastome sequences reveal rampant independent loss of the *ndh* gene family across orchids and associated instability of the inverted repeat/small single-copy region boundaries. *PLOS ONE* **10**(11):e014221 DOI [10.1371/journal.pone.0142215](https://doi.org/10.1371/journal.pone.0142215).
- Kim YK, Jo S, Cheon SH, Joo MJ, Hong JR, Kwak M, Kim KJ. 2020.** Plastome evolution and phylogeny of Orchidaceae, With 24 New Sequences. *Frontiers in Plant Science* **11**:22 DOI [10.3389/fpls.2020.00022](https://doi.org/10.3389/fpls.2020.00022).
- Kode V, Mudd EA, Iamtham S, Day A. 2005.** The tobacco plastid *accD* gene is essential and is required for leaf development. *The Plant Journal* **44**(2):237–244 DOI [10.1111/j.1365-313X.2005.02533](https://doi.org/10.1111/j.1365-313X.2005.02533).
- Kurtz S, Choudhuri JV, Ohlebusch E, Schleiermacher C, Stoye J, Giegerich R. 2001.** REPuter: the manifold applications of repeat analysis on a genomic scale. *Nucleic Acids Research* **29**(22):4633–4642 DOI [10.1093/nar/29.22.4633](https://doi.org/10.1093/nar/29.22.4633).
- Lin CS, Chen JJW, Chiu CC, Hsiao HCW, Yang CJ, Jin XH, Leebens-Mack J, Pamphilis CWde, Huang YT, Yang LH, Chang W J, Kui L, Wong GKS, Hu JM, Wang W, Shih MC. 2017.** Concomitant loss of NDH complex-related genes within chloroplast and nuclear genomes in some orchids. *The Plant Journal* **90**(5):994–1006 DOI [10.1111/tpj.13525](https://doi.org/10.1111/tpj.13525).
- Lin CS, Chen JJW, Huang YT, Chan MT, Daniell H, Chang WJ, Hsu CT, Liao DC, Wu FH, Lin SY, Liao CF, Deyholos MK, Wong GKS, Albert VA, Chou ML, Chen CY, Shih MC. 2015.** The location and translocation of *ndh* genes of chloroplast origin in the Orchidaceae family. *Scientific Reports* **5**:9040 DOI [10.1038/srep09040](https://doi.org/10.1038/srep09040).
- Lindley J. 1855.** *Folia orchidacea, an enumeration of the known species of orchids*, vol. 1. London:.
- Liu L, Wang Y, He P, Li P, Lee J, Soltis DE, Fu C. 2018.** Chloroplast genome analyses and genomic resource development for epilithic sister genera *Oresitrophe* and *Mukdenia* (Saxifragaceae), using genome skimming data. *BMC Genomics* **19**(1):235 DOI [10.1186/s12864-018-4633](https://doi.org/10.1186/s12864-018-4633).
- Neubig KM, Whitten WM, Carlswald BS, Blanco MA, Endara L, Williams NH, 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 DOI [10.1007/s00606-008-0105-0](https://doi.org/10.1007/s00606-008-0105-0).
- Nguyen LT, Schmidt HA, Haeseler AVon, Minh BQ. 2015.** IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Molecular Biology and Evolution* **32**(1):268–274 DOI [10.1093/molbev/msu300](https://doi.org/10.1093/molbev/msu300).
- 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. *Frontiers in Plant Science* **8**:715 DOI [10.3389/fpls.2017.00715](https://doi.org/10.3389/fpls.2017.00715).
- Pridgeon AM, Cribb PJ, Chase MW, Rasmussen FN. 2005.** *Genera Orchidacearum Volume 4. Epidendroideae*: Oxford University Press.
- Raubeson LA, Peery R, Chumley TW, Dziubek C, Fourcade HM, Boore JL, Jansen RK. 2007.** Comparative chloroplast genomics: analyses including new sequences from the angiosperms *Nuphar advena* and *Ranunculus macranthus*. *BMC Genomics* **8**(1):174 DOI [10.1186/1471-2164-8-174](https://doi.org/10.1186/1471-2164-8-174).
- Rolland N, Dorne AJ, Amoroso G, Sültemeyer DF, Joyard J, Rochaix JD. 1997.** Disruption of the plastid *ycf10* open reading frame affects uptake of inorganic carbon in the chloroplast of *Chlamydomonas*. *The EMBO Journal* **16**(22):6713–6726 DOI [10.1093/emboj/16.22.6713](https://doi.org/10.1093/emboj/16.22.6713).
- Roma L, Cozzolino S, Schlüter PM, Scopece G, Cafasso D. 2018.** The complete plastid genomes of *Ophrys iricolor* and *O. sphegodes* (Orchidaceae) and comparative analyses with other orchids. *PLOS ONE* **13**:e0204174 DOI [10.1371/journal.pone.0204174](https://doi.org/10.1371/journal.pone.0204174).
- Ronquist F, Teslenko M, Van Der Mark P, Ayres DL, Darling A, Höhna S, Larget B, Liu L, Suchard MA, Huelsenbeck JP. 2012.** MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. *Systematic Biology* **61**(3):539–542 DOI [10.1093/sysbio/sys029](https://doi.org/10.1093/sysbio/sys029).
- Rozas J, Ferrer-Mata A, Sánchez-DelBarrio JC, Guirao-Rico S, Librado P, Ramos-Onsins SE, Sánchez-Gracia A. 2017.** DnaSP 6: DNA sequence polymorphism analysis of large data sets. *Molecular Biology and Evolution* **34**(12):3299–3302 DOI [10.1093/molbev/msx248](https://doi.org/10.1093/molbev/msx248).
- Rumeau D, Peltier G, Cournac L. 2007.** Chlororespiration and cyclic electron flow around PSI during photosynthesis and plant stress response. *Plant, Cell & Environment* **30**(9):1041–1051 DOI [10.1111/j.1365-3040.2007.01675.x](https://doi.org/10.1111/j.1365-3040.2007.01675.x).
- Salazar GA, Cabrera LI, Madrinán S, Chase MW. 2009.** Phylogenetic relationships of Cranichidinae and Prescottiinae (Orchidaceae, Cranichideae) inferred from plastid and nuclear DNA sequences. *Annals of Botany* **104**:403–416 DOI [10.1093/aob/mcn257](https://doi.org/10.1093/aob/mcn257).
- Sasaki Y, Sekiguchi K, Nagano Y, Matsuno R. 1993.** Chloroplast envelope protein encoded by chloroplast genome. *FEBS Letters* **316**(1):93–98 DOI [10.1016/0014-5793\(93\)81743-J](https://doi.org/10.1016/0014-5793(93)81743-J).
- Schelkunov MI, Shtratnikova VY, Nuraliev MS, Selosse MA, Penin AA, Logacheva MD. 2015.** Exploring the limits for reduction of plastid genomes: a case study of the mycoheterotrophic orchids *Epipogium aphyllum* and *Epipogium roseum*. *Genome Biology and Evolution* **7**(4):1179–1191 DOI [10.1093/gbe/evv019](https://doi.org/10.1093/gbe/evv019).
- Shi C, Liu Y, Huang H, Xia EH, Zhang HB, Gao LZ. 2013.** Contradiction between plastid gene transcription and function due to complex posttranscriptional splicing: an exemplary study of *ycf15* function and evolution in angiosperms. *PLOS ONE* **8**(3):e59620 DOI [10.1093/gbe/evv019](https://doi.org/10.1093/gbe/evv019).
- Shikanai T, Endo T, Hashimoto T, Yamada Y, Asada K, Yokota A. 1998.** Directed disruption of the tobacco *ndhB* gene impairs cyclic electron flow around photosystem

- I. *Proceedings of the National Academy of Sciences of the United States of America* **95**(16):9705–9709 DOI [10.1073/pnas.95.16.9705](https://doi.org/10.1073/pnas.95.16.9705).
- Shinozaki K, Ohme M, Tanaka M, Wakasugi T, Hayashida N, Matsubayashi T, Zaita N, Chunwongse J, Obokata J, Yamaguchi-Shinozaki K, Ohto C, Torazawa K, Meng BY, Sugita M, Deno H, Kamogashira T, Yamada K, Kusuda J, Takaiwa F, Kato A, Tohdoh N, Shimada H, Sugiura M. 1986.** The complete nucleotide sequence of the tobacco chloroplast genome: its gene organization and expression. *The EMBO Journal* **5**(9):2043–2049 DOI [10.1002/j.1460-2075.1986.tb04464.x](https://doi.org/10.1002/j.1460-2075.1986.tb04464.x).
- Sosa V, Cameron KM, Angulo DF, Hernández-Hernández T. 2016.** Life form evolution in epidendroid orchids: ecological consequences of the shift from epiphytism to terrestrial habit in *Hexalectris*. *Taxon* **65**(2):235–248 DOI [10.12705/652.2](https://doi.org/10.12705/652.2).
- Stone J, Cribb P. 2017.** Lady tankerville's legacy. In: *A history and monographic review of phaius and gastrorchis*. KEW: Royal Botanic Gardens.
- Thiel T, Michalek W, Varshney R, Graner A. 2003.** Exploiting EST databases for the development and characterization of gene-derived SSR-markers in barley (*Hordeum vulgare* L.). *Theoretical and Applied Genetics* **106**(3):411–422 DOI [10.1007/s00122-002-1031-0](https://doi.org/10.1007/s00122-002-1031-0).
- Unruh SA, McKain MR, Lee YI, Yukawa T, McCormick MK, Shefferson RP, Smithson A, Leebens-Mack JH, Pires J C. 2018.** Phylotranscriptomic analysis and genome evolution of the Cypripedioideae (Orchidaceae). *American Journal of Botany* **105**(4):631–640 DOI [10.1002/ajb2.1047](https://doi.org/10.1002/ajb2.1047).
- Wang D, Zhang Y, Zhang Z, Zhu J, Yu J. 2010.** KaKs_Calculator 2.0: a toolkit incorporating gamma-series methods and sliding window strategies. *Genomics, Proteomics & Bioinformatics* **8**(1):77–80 DOI [10.1016/S1672-0229\(10\)60008-3](https://doi.org/10.1016/S1672-0229(10)60008-3).
- Wick RR, Schultz MB, Zobel J, Holt KE. 2015.** Bandage: interactive visualization of *de novo* genome assemblies. *Bioinformatics* **31**:3350–3352 DOI [10.1093/bioinformatics/btv383](https://doi.org/10.1093/bioinformatics/btv383).
- Wolfe KH, Li WH, Sharp PM. 1987.** Rates of nucleotide substitution vary greatly among plant mitochondrial, chloroplast, and nuclear DNAs. *Proceedings of the National Academy of Sciences of the United States of America* **84**(24):9054–9058 DOI [10.1073/pnas.84.24.9054](https://doi.org/10.1073/pnas.84.24.9054).
- Wyman SK, Jansen RK, Boore JL. 2004.** Automatic annotation of organellar genomes with DOGMA. *Bioinformatics* **20**(17):3252–3255 DOI [10.1093/bioinformatics/bth352](https://doi.org/10.1093/bioinformatics/bth352).
- Xiang XG, Jin WT, Li DZ, Schuiteman A, Huang WC, Li JW, Jin XH, Li ZY. 2014.** Phylogenetics of tribe Collabieae (Orchidaceae, Epidendroideae) based on four chloroplast genes with morphological appraisal. *PLOS ONE* **9**(1):e87625 DOI [10.1371/journal.pone.0087625](https://doi.org/10.1371/journal.pone.0087625).
- Xie WQ, Jäger K, Potts M. 1989.** Cyanobacterial RNA polymerase genes *rpoC1* and *rpoC2* correspond to *rpoC* of *Escherichia coli*. *Journal of Bacteriology* **171**(4):1967–1973 DOI [10.1128/jb.171.4.1967-1973.1989](https://doi.org/10.1128/jb.171.4.1967-1973.1989).
- Yamori W, Shikanai T. 2016.** Physiological functions of cyclic electron transport around photosystem I in sustaining photosynthesis and plant growth. *Annual Review of Plant Biology* **67**:81–106 DOI [10.1146/annurev-arplant-043015-112002](https://doi.org/10.1146/annurev-arplant-043015-112002).

- Yang JB, Tang M, Li HT, Zhang ZR, Li DZ. 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 DOI [10.1186/1471-2148-13-84](https://doi.org/10.1186/1471-2148-13-84).
- Yang Z. 2007.** PAML 4: phylogenetic analysis by maximum likelihood. *Molecular Biology and Evolution* **24**(8):1586–1591 DOI [10.1093/molbev/msm088](https://doi.org/10.1093/molbev/msm088).
- Yang Z, Nielsen R. 2002.** Codon-substitution models for detecting molecular adaptation at individual sites along specific lineages. *Molecular Biology and Evolution* **19**(6):908–917 DOI [10.1093/oxfordjournals.molbev.a004148](https://doi.org/10.1093/oxfordjournals.molbev.a004148).
- Yuan Y, Jin X, Liu J, Zhao X, Zhou J, Wang X, Wang D, Lai CJS, Xu W, Huang JW, HuangL P, Liu DH, Ma X, Wang L, Zhou MY, Jiang Z, Meng HB, Peng HS, Liang YT, Li RQ, Jiang C, Zhao YY, Nan TG, Jin Y, Zhan ZL, Yang J, Jiang WK, Huang LQ. 2018.** The *Gastrodia elata* genome provides insights into plant adaptation to heterotrophy. *Nature Communications* **9**(1):1–11 DOI [10.1038/s41467-018-03423-5](https://doi.org/10.1038/s41467-018-03423-5).
- Yukawa T, Cribb P. 2014.** Nomenclatural changes in the genus *Calanthe* (Orchidaceae). *Bulletin of the National Museum of Nature and Science, Series B* **40**:145–151.
- Yukawa T, Ishida G. 2008.** Phylogeny and *systematics* of *Calanthe* and allied genera. Abstracts retrieved from 7th annual meeting of japanese society for plant systematics. Abstract, 32.
- Zhai JW, Zhang GQ, Li L, Wang MN, Chen LJ, Chung SW, Rodríguez FJ, Francisco-Ortega J, Lan SR, Xing FW, Liu ZJ. 2014.** *Molecular Phylogenetics and Evolution* **77**:216–222 DOI [10.1016/j.ympev.2014.04.005](https://doi.org/10.1016/j.ympev.2014.04.005).
- Zhang D, Gao F, Jakovlić I, Zou H, Zhang J, Li WX, Wang GT. 2020.** PhyloSuite: an integrated and scalable desktop platform for streamlined molecular sequence data management and evolutionary phylogenetics studies. *Molecular Ecology Resources* **20**(1):348–355 DOI [10.1111/1755-0998.13096](https://doi.org/10.1111/1755-0998.13096).