Independent degradation in genes of the plastid ndh gene family in species of the orchid genus *Cymbidium* (Orchidaceae; Epidendroideae)

Hyoung Tae Kim¹, Mark W. Chase²*

1 College of Agriculture and Life Sciences, Kyungpook University, Daegu, Korea, 2 Jodrell Laboratory, Royal Botanic Gardens, Kew, Richmond, Surrey, United Kingdom

* m.chase@kew.org

**Abstract**

In this paper, we compare ndh genes in the plastid genome of many *Cymbidium* species and three closely related taxa in Orchidaceae looking for evidence of ndh gene degradation. Among the 11 ndh genes, there were frequently large deletions in directly repeated or AT-rich regions. Variation in these degraded ndh genes occurs between individual plants, apparently at population levels in these *Cymbidium* species. It is likely that ndh gene transfers from the plastome to mitochondrial genome (chondriome) occurred independently in Orchidaceae and that ndh genes in the chondriome were also relatively recently transferred between distantly related species in Orchidaceae. Four variants of the *ycf1-rpl32* region, which normally includes the ndhF genes in the plastome, were identified, and some *Cymbidium* species contained at least two copies of that region in their organellar genomes. The four *ycf1-rpl32* variants seem to have a clear pattern of close relationships. Patterns of ndh degradation between closely related taxa and translocation of ndh genes to the chondriome in *Cymbidium* suggest that there have been multiple bidirectional intracellular gene transfers between two organellar genomes, which have produced different levels of ndh gene degradation among even closely related species.

**Introduction**

The first two plastid genomes (plastomes) sequenced included the entire ndh 11-gene family, which is analogous to complex I in the mitochondrial genome (chondriome) [1, 2]. Subsequently, the function of the ndh plastome genes has been described in many studies. The Ndh complex codes for an NADH-specific dehydrogenase with low levels of expression [3, 4], and the family is involved in cyclic electron flow and chlororespiration [4, 5]. Recently, Yamori et al. [6] investigated the function of Ndh complex in low light. However, in spite of this role, the Ndh complex is dispensable for plant growth under optimal conditions [4], and an alternative cyclic electron transport pathway has been reported [7, 8]. Therefore, it has been suggested
that ndh-lacking species in which at least one of ndh genes is non-functional may be able to use the alternative pathway for cyclic electron transport [9].

When the loss of the 11 ndh genes in Pinus thunbergii was reported [10], this striking feature was considered unique because ndhF had been found to be present in all other major sequenced vascular plant clades [11]. However, losses of ndh gene function have subsequently been reported in various clades of land plants. In bryophytes, the 11 ndh genes in the parasitic liverwort, Aneura mirabilis (synonym, Cryptothallis mirabilis), were partially or completely deleted [12], and ndhF of the leafy liverwort, Ptilidium pulcherrimum, was found to be a pseudogene [13]. In the fern clade, some leptosporangiate ferns had internal stop codons in ndh genes, but this seemed to be related RNA editing [14–16]. In gymnosperms, ndh gene losses have been reported in Pinaceae [10, 17–19] and Gnetales [20, 21]. Parasitic angiosperms have lost the function of ndh genes as well as other photosynthesis-related genes [22–25], but some autotrophs also lack the ndh gene [26–29].

Degradation of ndh in Orchidaceae is noteworthy from the perspective of the 11 ndh genes found in 743 angiosperm plastomes (Fig 1) (S1 Table). All 11 ndh genes had been coded into four classes [30], and different coding ndh gene patterns have been in each order based on the extent to which ndh genes were variously degraded. Reported plastome sequences of rosids comprise 32.5% of the 743 plastid genomes, but only the rosid order Geraniales have degraded ndh genes [28, 31]. With the exception of internal stop codons caused by 1-bp insertions or deletions (indels) in Asterales [32, 33], ndh gene degradation in the asterids is restricted to parasitic taxa in Lamiales and Solanales [23, 24, 34–37]. In monocots, the number of sequenced Poales is 21.4% of angiosperms, but only ndhA in some species seems to be a pseudogene caused by short indels.

In contrast, among Asparagales, in which most of the sequenced species are orchids, ndh degradation patterns vary considerably. Even though many orchids have all 11 ndh genes intact in their plastomes [9, 30, 38], a number of degraded ndh genes in photosynthetic orchids have been reported [9, 30, 39–44] in addition to those in non-photosynthetic Orchidaceae [45–49]. This result demonstrates that more ndh genes in Orchidaceae have been independently modified than in any other family of angiosperms. Therefore, to understand better ndh gene degradation, we focus here on orchid plastomes.

Degradation of ndh genes among genera in Orchidaceae seems to be independent [9, 30], but the scale of variation among closely related species level has yet to be investigated. The plastomes of the two Phalaenopsis species sequenced had similar ndh gene degradation patterns [41], which was observed as well as in the plastome of Phalaenopsis hybrids [30]. Most ndh genes in the eight species of Cymbidium sequenced were full-length, although some of them had frame-shift mutations that render them functionless [43]. Degradation of ndh in subtribe Oncidiinae varied slightly among genera [40]. However, 15 of the reported Oncidinae were complex hybrids, and it was difficult to determine the ancestral character status of ndh gene degradation among these. Comparative analysis of ten species of coralroot orchids [48] and two species of a distantly related genus, Epipogium [49], all of which are holomycoheterotrophic, indicated ndh genes had become pseudogenes or were completely deleted in each of their common ancestors. However, recently submitted plastome sequences of Cymbidium in GenBank showed different ndh gene deletions among individuals within species. Therefore, it seems that ndh genes in Cymbidium may be being actively degraded and that an investigation of ndh gene status will help us understand broader patterns of ndh gene degradation in Orchidaceae.

In this paper, 11 ndh loci among 23 Cymbidium species including hybrids and three closely related taxa are analyzed for ndh gene degradation. Except for ndhF, we tried to investigate all ndh genes. The ndhF gene was completely deleted in some species in Cymbidium or contained
a number of internal homopolymer regions, which we assume indicates non-functional genes. Therefore, we confirmed only the presence of \textit{ndhF} in each plastome. Additionally, we analyzed NGS data to determine if \textit{ndh} genes had been translocated to the chondriome [9] because we found multiple copies of some \textit{ndh} genes in \textit{Cymbidium} species in our investigations.

**Results**

**Ten \textit{ndh} loci among 23 \textit{Cymbidium} species and three closely related taxa**

Four regions (\textit{ndhB, ndhJ-K-C, ndhD, ndhE-G-I-A-H}) that included ten \textit{ndh} genes from 23 \textit{Cymbidium} species and three outgroups were amplified by PCR and sequenced (Table 1). However, some intergenic or coding regions could not be sequenced because they contained homopolymers and polyA/T-polyG/C or problematic secondary structure (inverted repeats). To identify indels in ten \textit{ndh} genes among 23 \textit{Cymbidium} species and three closely related taxa, the fully intact (functional) \textit{ndh} genes of \textit{Masdevallia coccinea} were used as reference sequence.

Except for \textit{C. tigrinum} in which only half of exon1 is present and \textit{C. mastersii} in which the 5′ region failed to produce sequence, all \textit{Cymbidium} species were documented to contain a full-length \textit{ndhB} gene (S1A Fig). A 1-bp insertion at 37 bp downstream of the 5′ end of \textit{ndhB} results in a frame-shift mutation in \textit{ndhB} in reported plastome sequences of \textit{Cymbidium}, and this was also identified in all \textit{Cymbidium} species studied here and the closely related \textit{Acriopsis} and \textit{Thecostele} accessions (subtribe Cymbidiinae)[51]. A large deletion including exon1, intron and exon2 was detected in \textit{ndhB} of \textit{Acriopsis}.
The *ndhJ*-K-C region was more variable than that of *ndhB* (S1B Fig). A 12-bp direct repeat was distributed 63 bp downstream of the 5' end of *ndhC* and 69–82 bp downstream of 3' end of *ndhJ* in most *Cymbidium* species. However, the sequence between the direct repeats was only deleted in *C. goeringii*, a result that conflicts with the complete plastome sequence of same species in GenBank (NC_028524), but this was based on a different individual of that species. Deletions caused by direct repeat sequences were also found in the 5' region of *ndhJ* in three *Cymbidium* species (*C. floribundum*, *C. erythrostylum*, and *C. tigrinum*), *Acriopsis* and *Thecos tele*. Unexpectedly, two copies of *ndhJ*-K-C region were detected in *C. atropurpureum*. Type I was similar to other *Cymbidium* sequences, whereas type II contained a 87-bp insertion 39 bp downstream of the 5' end of *ndhK*. This 87 bp insertion is not present in any other of the 743

### Table 1. Taxa list for this study.

| Subgenus a | Section b | Species | DNA bank number or living collection number |
|------------|-----------|---------|------------------------------------------|
| **Cymbidium** | Austrocymbidium | Cymbidium madidum | O-1472 |
| | Bigibbarium | Cymbidium devonianum | * |
| | Cymbidium | Cymbidium atropurpureum | O-1465 |
| | Cymbidium | Cymbidium finlaysonianum | 1954–41302 BEAK |
| | Floribundum | Cymbidium floribundum | O-1461 |
| | Floribundum | Cymbidium pumilum | O-1469 |
| | Floribundum | Cymbidium suavissimum | O-1467 |
| | Himantophyllum | Cymbidium dayanum | O-1468 |
| **Cyperorchis** | Annamæa | Cymbidium erythrostylum | O-1471 |
| | Cyperorchis | Cymbidium eburneum | O-1505 |
| | Cyperorchis | Cymbidium elegans | O-1479 |
| | Cyperorchis | C. eburneum x C. hookerianum | O-1481 |
| | Cyperorchis | Cymbidium erythraeum | O-1463 |
| | Cyperorchis | Cymbidium giganteum | O-69 |
| | Cyperorchis | Cymbidium hookerianum | O-1466 |
| | Cyperorchis | Cymbidium insigne | O-1475 |
| | Cyperorchis | Cymbidium iridoides | O-1462 |
| | Cyperorchis | Cymbidium lowianum | O-1476 |
| | Cyperorchis | Cymbidium mastersii | O-1506 |
| | Cyperorchis | Cymbidium sanderae | O-1470 |
| | Cyperorchis | Cymbidium whiteae | O-1473 |
| | Parishiella | Cymbidium tigrinum | 17717 |
| **Jensoa** | Jensoa | Cymbidium ensifolium c | O-1478 |
| | Jensoa | Cymbidium goeringii | O-1477 |
| | Jensoa | Cymbidium kanran c | O-1499 |
| | Jensoa | Cymbidium lancifolium c | O-293 |
| | Jensoa | Cymbidium sinense c | O-1503 |
| Outgroup | Acriopsis sp. | 9060 |
| | Grammatophyllum speciosum | 1983–2947 BACR 450 |
| | Thecostele secunda | O-406 |

a: Subgeneric delimitation of *Cymbidium* is based on Du Puy and Cribb [50]
b: Sectional delimitation of *Cymbidium* is based on Du Puy and Cribb [50]
c: The plastome sequence of these species have been reported by Yang et al. [43] and directly submitted by Kim et al. in NCBI. Therefore, these four species are used for confirming the location of *ndh* genes to mitochondrial genome.

*: Only fresh leaves were collected from Ractliffe Orchids, Ltd. (Hampshire, UK)

https://doi.org/10.1371/journal.pone.0187318.t001
angiosperm plastomes in GenBank. Only *C. madidum*, *C. finlaysonianum* and the mt copy of *ndhK* in all *Cymbidium* species contained sequences of this same type.

The *ndhD* regions of *Cymbidium* were relatively conserved (S2A Fig). Large deletions were located in the 3′ region of the gene. Some of these occurred between direct repeat sequences.

The largest deletion of *ndh* genes in *Cymbidium* was identified in the *ndhE-G-I-A-H* region (S2B Fig), the end points of which were commonly located in an extremely AT-rich region. In particular, deletion of *ndhA* exon1 and *ndhH* in *C. goeringii* corresponded to those occurring in the plastomes of *C. ensifolium*, *C. kanran*, *C. lancifolium* and *C. macrorhizum* even though the plastome of different individuals of *C. ensifolium* (NC_028525) and *C. goeringii* (NC_028524) contained full length pt-*ndhA* and *ndhH*.

**Different types of the *ycf1-rpl32* region in *Cymbidium***

The *ycf1-rpl32* region of the sequenced plastomes of *Cymbidium* was subdivided into two different types in comparison with that of *M. coccinea* (Fig 2A). Type A *ycf1-rpl32* was similar to the reference, whereas 420 bp of 3′ region of *ndhF* was replaced with *ycf1* sequence in type B *ycf1-rpl32*.

*Cymbidium dayanum* in subg. *Cymbidium* and nine species of subg. *Cyperorchis* contained type A *ycf1-rpl32*, which was highly conserved (Fig 2B). In contrast to type A *ycf1-rpl32*, type B *ycf1-rpl32* of *Cymbidium* had number of indels in 3′ region of *ndhF* (Fig 2C). The type B *ycf1-rpl32* of *C. sinense* sequenced in this paper was only 87% similar to that of *C. sinense* plastome owing to many indels. Type B *ycf1-rpl32* was also found in three *Cymbidium* species in which plastid *ndhF* was completely deleted. In comparison to type B *ycf1-rpl32*, type C *ycf1-rpl32* had large deletion in the 3′ region of *ndhF*, and the end point of the deletion corresponded to the end point of the replaced *ycf1* region (Fig 2C and 2D).

Type D *ycf1-rpl32* in which *ndhF* was completely deleted was found in half of the *Cymbidium* species examined and the three closely related taxa with a high level of similarity among them (Fig 2E). In comparison with type A *ycf1-rpl32*, two large deletions occurred in type D *ycf1-rpl32*; one was the complete deletion of *ndhF* and the other was an intergenic deletion between *ndhF* and *rpl32*.

**Multiple copies of *ndh* genes in Orchidaceae***

The 38 *ndh* partial sequences were detected from 15 contigs using four sets of NGS data from Orchidaceae (Table 2). With the exception of one contig in *C. lancifolium*, the ratio of the depth of mt-*ndh* genes to the depth of plastome in 15 contigs was 5.5~14.5, and BLAST results confirmed that they were derived from the chondriome.

The contig that contained the *ndhJ-K-C* region in *C. lancifolium* was present in relatively lower depth and did not contain a mitochondrial region, but there were only two SNPs and one indel that differed among the mt-*ndhJ-K-C* region in *C. lancifolium* and *C. macrorhizum*. Consequently, we concluded all 16 contigs have been translocated from the plastome to the chondriome.

Two *Cymbidium* species in section *Pachyrhizanthe*. All 11 *ndh* genes have been found in the chondriome of two *Cymbidium* species, and most of them do not differ in these two species. The mt-*ndhB* gene lacked 44 bp of exon1 and contained a 132-bp deletion in exon2 (Fig 3A). Similarities of the *ndhB* genes in the same genome among different species were 99.0 and 99.5%. However, those in the genomes of two accessions of same species were only 91.1 and 91.9% similar. Mt-*ndhJ* and *ndhK* contained a large deletion and insertion, respectively (Fig 3B). The length variation of insertion in mt-*ndhK* between two *Cymbidium* species was due to tandem repeats of 28 bp sequence. Even though plastid *ndhF* was completely deleted, two
Fig 2. Four types of ycf1-rpl32 regions in organellar genomes of *Cymbidium* and closely related taxa. The red dotted line refers to identical position of C) the end of replaced *ycf1* and D) the end of deletion in *Acriopsis* and *Thecostele*. A) The *ndhF* genes of currently sequenced plastomes are divided into two groups. Type A is similar to *ndhF* of *Masdevallia coccinea* whereas type B has 420 bp *ycf1*-like region at 3' region of *ndhF*. B) Type A ycf1-rpl32 region is more conserved than the others. C) Type B ycf1-rpl32 regions have a number of deletions. D) The 3' region of *ndhF* is deleted in the type C ycf1-rpl32 region. E) Type D ycf1-rpl32 region completely lacks *ndhF*.

https://doi.org/10.1371/journal.pone.0187318.g002
Table 2. The information of mt-ndh genes assembled from NGS data.

| Taxa                         | Region | Accession | Length | Average depth of mt-ndh gene / Average depth of plastome | Reference |
|------------------------------|--------|-----------|--------|----------------------------------------------------------|-----------|
| *Cymbidium macrorhizon*      | ndhA   | KX962303  | 2176   | 48.3 / 459.2                                             |           |
| *(4 contigs)*                | ndhB   | KX962302  | 2036   | 25.5 / 459.2                                             |           |
|                              | ndhC   | KX962305  | 356    | 39.9 / 459.2                                             |           |
|                              | ndhD   | KX962303  | 236    | 43.9 / 459.2                                             |           |
|                              | ndhE   | KX962303  | 284    | 57.4 / 459.2                                             |           |
|                              | ndhF   | KX962304  | 1910   | 41.4 / 459.2                                             |           |
|                              | ndhF   | KX962304  | 779    | 31.4 / 459.2                                             |           |
|                              | ndhF   | KX962303  | 497    | 48.4 / 459.2                                             |           |
|                              | ndhH   | KX962303  | 1127   | 39.5 / 459.2                                             |           |
|                              | ndhI   | KX962303  | 464    | 48.0 / 459.2                                             |           |
|                              | ndhJ   | KX962305  | 362    | 44.0 / 459.2                                             |           |
|                              | ndhK   | KX962305  | 867    | 46.9 / 459.2                                             |           |
| *Cymbidium lancifolium*      | ndhA   | KX962298  | 2199   | 35.3 / 318.7                                             |           |
| *(6 contigs)*                | ndhB   | KX962296  | 2047   | 25.8 / 318.7                                             |           |
|                              | ndhC   | KX962301  | 356    | 13.1 / 318.7                                             |           |
|                              | ndhD   | KX962297  | 773    | 23.0 / 318.7                                             |           |
|                              | ndhD   | KX962298  | 236    | 20.8 / 318.7                                             |           |
|                              | ndhE   | KX962298  | 284    | 40.2 / 318.7                                             |           |
|                              | ndhF   | KX962300  | 2100   | 39.8 / 318.7                                             |           |
|                              | ndhF   | KX962299  | 955    | 32.8 / 318.7                                             |           |
|                              | ndhG   | KX962298  | 497    | 24.8 / 318.7                                             |           |
|                              | ndhH   | KX962298  | 1127   | 29.4 / 318.7                                             |           |
|                              | ndhI   | KX962298  | 464    | 22.3 / 318.7                                             |           |
|                              | ndhJ   | KX962301  | 362    | 6.4 / 318.7                                              |           |
|                              | ndhK   | KX962301  | 811    | 6.6 / 318.7                                              |           |
| *Dendrobium catenatum*       | ndhA   | KX962306  | 1537   | 779.9 / 7687.6                                           | SRR2084072|
| *(4 contigs)*                | ndhA   | KX962306  | 575    | 739.8 / 7687.6                                           | SRR2084072|
|                              | ndhC   | KX962309  | 355    | 671.5 / 7687.6                                           | SRR2084072|
|                              | ndhD   | KX962307  | 1323   | 751.0 / 7687.6                                           | SRR2084072|
|                              | ndhE   | KX962307  | 306    | 780.6 / 7687.6                                           | SRR2084072|
|                              | ndhF   | KX962308  | 1600   | 738.6 / 7687.6                                           | SRR2084072|
|                              | ndhG   | KX962307  | 212    | 1118.4 / 7687.6                                          | SRR2084072|
|                              | ndhH   | KX962306  | 1155   | 664.9 / 7687.6                                           | SRR2084072|
|                              | ndhI   | KX962306  | 501    | 731.5 / 7687.6                                           | SRR2084072|
|                              | ndhJ   | KX962309  | 472    | 626.7 / 7687.6                                           | SRR2084072|
|                              | ndhK   | KX962309  | 610    | 636.7 / 7687.6                                           | SRR2084072|
| *Epipogium aphyllum*         | ndhA   | KX962310  | 215    | 28.7 / 216.8                                             | SRR1344939|
| *(1 contig)*                 | ndhA   | KX962310  | 425    | 29.4 / 216.8                                             | SRR1344939|
|                              | ndhI   | KX962310  | 684    | 15.1 / 216.8                                             | SRR1344939|

Evolution of ndh genes in Cymbidium

Copies of mt-ndhF were found in two Cymbidium species (Fig 3C). One copy of these was similar to ndhF in type B ycf1-rpl32, and the other was similar to ndhF in type C ycf1-rpl32. In comparison with their plastome sequence, mt-ndhD was truncated and mt-ndhA and ndhH genes were almost full length (Fig 3D). In addition, another mt-ndhD (773 bp) was found in C. lancifolium.
Dendrobium catenatum. The nine mt-ndh genes were found in four large contigs (Table 2). Among them, three contigs could form subgenomic circles [52]. Because a number of pt-ndh genes of D. catenatum have been deleted [53], we used a completely intact set of pt-ndh genes as a reference sequence, in this case Sobralia.

The region of mt-ndhJ-K-C was similar to the reference sequence in length with the exception of a large deletion in mt-ndhK, whereas pt-ndhK and ndhC were completely absent (Fig 4A). Mt-ndhF was longer than pt-ndhF, but both of them were highly truncated (Fig 4B). The regions between 194 bp downstream of rpl32 and 317 bp downstream of the 5' end of ndhG were relatively conserved between pt- and mt-ndh genes, but the 3' region of ndhD had a large deletion in both genomes (Fig 4C). The regions with pt-ndhI and ndhA exon2 were deleted [53], whereas these genes were found in chondriome but with a large inversion upstream of 5' end of ndhG and downstream of the 5' end of ndhA (Fig 4D).
Epipogium aphyllum. We found mt-ndhI and ndhA genes in achlorophyllous (holomy- cotrophic) *E. aphyllum*, but all pt-ndh genes in this species were completely deleted [49]. Unexpectedly, there was also an inversion mutation like that found in mt-ndhI-A of *D. catenatum* (Fig 4E).

Phylogenetic relationships between pt- and mt-ndh genes in Orchidaceae

In most ndh-gene trees (S3 Fig), the mt-ndh genes of *Cymbidium* formed a clade. It was noteworthy that the clustering of mt-ndhD, ndhE and ndhG from the NGS data and direct sequencing was strongly supported. However, the mt-ndhH genes of section *Pachyrhizanthus* formed a clade with the pt-ndhH genes of previously sequenced *Cymbidium* plastomes [43], whereas all pt-ndhH genes of *Cymbidium* sequenced in this study formed a strongly supported cluster. In addition, the ndhI, ndhK and ndhC genes of *C. madidum*, *C. finlaysonianum* and type II *C.*
Evolution of ndh genes in Cymbidium

Patterns of ndh degradation in Cymbidium

Function of ndh genes has been independently lost in some orchid clades [9, 30]. With the exception of the directly sequenced plastomes of Goodyera, ndh-missing/non-intact species and ndh-intact species have not been so far found in same genus of Orchidaceae [41, 43, 48], in contrast to the situation in Erodium [27, 28]. Therefore, loss of function in the ndh complex seems to have occurred in the common ancestor of the ndh-missing/non-intact species within those genera rather than independently at the species level. The situation for ndhA, ndhB, ndhF, and ndhH genes of Cymbidium was strongly supported, whereas another ndhB from C. ensifolium (KU179434) formed a group with ndhG in Cymbidium. Multiple copies of the mt-ndh genes from Erycina pusilla (subtribe Oncidiinae) formed a unique cluster with the exception of one copy of mt-ndhD (246 bp), which was relatively shorter than other mt-ndhD genes (480–1078 bp) in E. pusilla. Furthermore, these mt-ndh genes clustered with their pt-counterparts with the exception of pt-ndhA, ndhI and ndhE, which were truncated or missing from the plastome of E. pusilla.

The mt-ndhA, ndhD, ndhE, ndhG, ndhH, ndhI and ndhJ genes in Masdevallia picturata were most closely related to the pt-ndh genes of Masdevallia, and almost all mt-ndh genes in Paphiopedilum also formed clusters with the pt-ndh genes of these species.

Discussion

Patterns of ndh degradation in Cymbidium

atropurpureum formed a cluster with mt-ndhJ, ndhK and ndhC of section Pachyrhizanthe. The second copy of mt-ndhD in C. lancifolium clustered with the mt-ndhD of Oncidium, and they formed a strongly supported group with other orchid mt-ndhD genes. The clustering of the pt-ndhG of C. ensifolium (NC_028525) and mt-ndhG from other species of Cymbidium was strongly supported, whereas another pt-ndhG from C. ensifolium (KU179434) formed a group with pt-ndhG in Cymbidium.

The first sequenced plastomes of Cymbidium [43] and directly uploaded sequences (NC_028525 and NC_028524) contained full-length ndh genes even though most of them were pseudogenes due to frameshift mutations. However, recently a sequenced plastome of Cymbidium lacked pt-ndhF, ndhH and ndhA exon1. As a result, there are two plastomes of C. ensifolium with different ndh gene content. With the exception of technical errors (misidentification at the time of collection or laboratory errors), which is difficult to determine in this study, our results support the hypothesis that Cymbidium species have undergone dynamic and recent ndh gene degradation. Because the common ancestor of all Cymbidium species seems to have lacked ndh function, many different substitutions and indels may have accumulated in the various species due to relaxed selection. The large deletions that caused ndh degradation should be shared between closely related taxa if ndh gene degradation had occurred in an ancestral pseudogene further in the past. However, most of the large deletions detected are unique in each accession.

In addition, one of the main factors involved in ndh gene degradation is likely to be intracellular recombination. A number of deletions have been found between direct repeat sequences or extremely AT-rich (homopolymer) regions. These patterns have been known to relate to intramolecular recombination [60, 61] and illegitimate recombination [62].
respectively. These results suggest that the plastomes in *Cymbidium* species have undergone independent *ndh* gene degradation, probably after they speciated. The different levels of plastid *ndh* gene degradations in different individuals of *C. ensifolium* and *C. goeringii* also support a hypothesis of recent *ndh* gene degradation in *Cymbidium*.

However, we cannot suggest a clear explanation for why there appears to be a recent burst in this activity in the extant species of *Cymbidium*. In contrast, the *ndh*-lacking genera of photosynthetic orchids, i.e. *Phalaenopsis* [41], *Oncidium*, *Paphiopedilum* [30], *Dendrobium* and *Bletilla*, have retained similar *ndh* gene degradation patterns among their species. In general, with the exception of extremely reduced mycoheterotrophic orchids [45, 49], a number of pseudogenes have been retained in the plastomes of Orchidaceae [46–48]. In particular, the closely related green and non-green coralroot orchids (*Corallorhiza*), which have lost some *ndh* genes, are similar in plastid genome size [48]. Therefore, the plastome of Orchidaceae may be prone to retain its size due to some selective constraints.

Barrett et al. [47] hypothesised that non-functional genes in mycoheterotrophic plants may have undergone point mutations and frame-shift mutations under relaxed selective pressure over time, and large deletions occur rarely after purifying selection on non-functional genes ceases. Unlike other genera in Orchidaceae, the most recent common ancestor (MRCA) of *Cymbidium* seems to have been under selective genome size constraint even though *ndh* function had been lost. However, structural mutations like bidirectional homologous recombination between the two organellar genomes or gene conversion in *ndhF* after splitting of populations or speciation might have led the plastome to be under relaxed selective constraints. As a result, it is likely that dynamic *ndh* gene degradation has occurred among *Cymbidium* species, perhaps even among populations.

**Diverse *ndhF* genes result from gene conversion and indels**

The first five *Cymbidium* species studied previously had full-length plastid *ndhF* genes [43], but *ndhF* deletions occurred in four recently submitted sequences. As we reported for the *ndhA-H* region, the deleted *pt-ndhF* genes of *C. lancifolium* and *C. macrorhizon* were transferred to chondriome (Fig 3C). As a result, *C. sinense* contains type B *ycf1-rpl32* in its plastome and type D *ycf1-rpl32* in its chondriome, whereas *C. kanran*, *C. ensifolium*, *C. macrorhizon* and *C. lancifolium* contain type D *ycf1-rpl32* in their plastomes and type B *ycf1-rpl32* in their chondriomes. Other *Cymbidium* species also contain different types of *ycf1-rpl32* in their organellar DNAs, but we do not know in which genomes these are located. Species that have the same type of the *ycf1-rpl32* region are not related to each other (i.e. they belong to different clades in the *Cymbidium* phylogenetic tree). Nevertheless, four types of the *ycf1-rpl32* region seem to be related each other.

Type A *ycf1-rpl32* is similar to that of other Orchidaceae, whereas 420 bp of the 3’ region of *ndhF* in type B *ycf1-rpl32* is similar to the *ycf1* region and contained a number of indels. The *ndhF* sequence near IR_\(\alpha\)/SSC was replaced with *ycf1* near SSC/IR_\(\alpha\). This replacement might result from IR expansion via gene conversion [63] (S4 Fig). First, recombination was initiated within the IR. Then, a Holliday junction on the IR was moved to SSC, creating heteroduplex DNAs. These heteroduplex DNAs were repaired using the complementary strand as the model. Finally, base substitutions and indels occurred in the *ycf1* like region in *ndhF*. Significantly, an end point for deletion of *ndhF* in *Acriopsis* and *Thecostele* was identical to that of a *ycf1*-like region in *ndhF* of *C. tortisepalum* (Fig 2C and 2D). Therefore, it is possible that type C *ycf1-rpl32* was derived from type B *ycf1-rpl32* due to deletion of a chimeric region.

Kim et al. [30] described the important role of *ndhF* in the instability of the IR/SSC junction in Orchidaceae. Retention of full-length *ndhF* seems to be related to the selective constraints
that maintain the IR/SSC boundary. The \textit{ndhF} of the type B \textit{ycf1-rpl32} region is similar to \textit{ndhF} in type A \textit{ycf1-rpl32} in length, but in its content is similar to the truncated version of \textit{ndhF} due to the replacement of 3' end region of \textit{ndhF}. As a result, it seems that gene conversion leads to relaxed selective constraint of the IR/SSC junction, after which truncated \textit{ndhF} versions in type B and type C \textit{ycf1-rpl32} may be followed by \textit{ndhF} deletion as in type D \textit{ycf1-rpl32}.

**Intracellular gene transfers between organellar DNA**

Chang et al. [39] confirmed the in-frame sequences of \textit{ndhA}, \textit{ndhF} and \textit{ndhH} that are completely deleted in the plastome of \textit{Phalaenopsis aphrodite} and suggested that they were transferred to nuclear genome. However, in the recently published whole genome of \textit{P. equestris} [64], it was shown that there was also no intact \textit{ndh} gene [30]. Subsequently, mt-\textit{ndh} genes were found in many unrelated clades of Orchidaceae [9], and we also found mt-\textit{ndh} genes in several distantly related species. Therefore, intact \textit{ndh} genes that are deleted from the plastome of \textit{Phalaenopsis} are likely to be found in its chondriome. However, this is not surprising because such transfers are known to occur widely in seed plants [65–68].

To evaluate relationships between plastid and mitochondrial copies of \textit{ndh} genes in Orchidaceae, we constructed gene trees (S3 Fig), which gave us information about \textit{ndh} gene transfer, although some nodes are not well resolved. First, it is likely that the transfers of \textit{ndh} genes from plastome to chondriome have usually occurred in the MRCA of the species in each genus. As there is limited \textit{ndh} gene information at the species level, especially for mt-\textit{ndh} genes, it is impossible to infer a time for these transfers. However, many of the pt- and mt-\textit{ndh} genes from a given genus cluster together. For instance, mt-\textit{ndhC}, \textit{ndhD}, \textit{ndhG}, \textit{ndhH} and \textit{ndhF} of \textit{Erycina pusilla} (subtribe Oncidiinae) were transferred after \textit{Erycina} diverged from its common ancestor with \textit{Oncidium} (subtribe Oncidiinae). The mt-\textit{ndh} genes in \textit{Masdevallia picturata} (subtribe Laeliinae, subfamily Epidendroideae) and \textit{Paphiopedium} (subfamily Cypripedioideae) were also sister to pt-\textit{ndh} genes of species within each genus, respectively.

In the \textit{ndh} tree of \textit{Cymbidium}, most mt-\textit{ndh} genes are distantly located from their pt-\textit{ndh} counterparts, and the entire mt-\textit{ndhD-E-G-I-A-H} region can be assembled from NGS data for two species, which we confirmed by PCR of the mt-\textit{ndhD-E-G} region in six \textit{Cymbidium} species. These mt-\textit{ndh} genes clustered uniquely with strong support. Although the combined \textit{ndh} gene tree for ten species of \textit{Cymbidium} had a different topology from that of combined ITS + \textit{matK} [69], it is clear that the transfer of the \textit{ndh} genes in the single-copy region dates back at least to the common ancestor of these \textit{Cymbidium} species.

Secondly, transfers between the chondriome of photosynthetic orchids have occurred more than once. The mt-\textit{ndhD} genes of \textit{Cymbidium} (Cymbidiinae) and \textit{Erycina} (Oncidiinae) were divided into two groups. The mt-\textit{ndhD} genes (from mt-\textit{ndhD-E-G} region) of \textit{Cymbidium} and \textit{Erycina} clustered with mt-\textit{ndhD} genes in same genus. However, another copy of mt-\textit{ndhD} gene in \textit{C. lancifolium} and \textit{Erycina} formed a strongly supported cluster with the mt-\textit{ndhD} genes from \textit{Oncidesa Gower Ramsey} (a complex hybrid between species in \textit{Oncidium} and \textit{Gomesa}, most likely with the plastid genome of the former) and a member of another subfamily \textit{Goodyera fumata} (tribe Cranichidae, subfamily Orchidoideae). These four mt-\textit{ndhD} genes clustered with mt-\textit{ndhD} gene of \textit{D. catenatum} (tribe Malaxidae, subfamily Epidendroideae), to which the plastid \textit{ndhD} of \textit{Dendrobium} was an outlier with moderate support. It is therefore likely that mt-\textit{ndhD} of \textit{Dendrobium} has been directly transferred independently to the other four species [70]. In addition, mt-\textit{ndhE} of \textit{Oncidesa Gower Ramsey} (subfamily Epidendroideae) and \textit{V. planifolia} (subfamily Vanilloideae) are identical. Although the substitution rate of the chondriome is slower than in plastid DNA [52], it is unlikely that mt-\textit{ndhE} of two species originated in their common ancestor because of the long time, before the end of the
Cymbidium species can have different types of the organellar genomes. For example, some species also have other types, e.g. type D. It is highly perplexing that Cymbidium species can have different types of the ycf1-rpl32 region in one genome (plastome or chondriome) and the same type of ycf1-rpl32 region in different genomes. We have two hypotheses that could explain this phenomenon: C. sinense and C. macrorhizon represent non-functional ndhF (type A, B and C) and completely ndhF-deleted species (type D), respectively.

The first hypothesis is unidirectional transfer (Fig 5A). The ycf1-rpl32 region containing ndhF (ancestral type) was transferred to its chondriome. Subsequently, the mt-ndhF (C. sinense) and pt-ndhF (C. macrorhizon) were independently deleted. The second hypothesis is bidirectional transfer (Fig 5B). In this scenario, the ycf1-rpl32 region containing plastid ndhF was transferred to chondriome in the ancestor of Cymbidium and closely related genera of subtribe Cymbidiiinae. After this transfer, the mt-ndhF copy was eliminated by gene rearrangements or gene deletion (as in C. sinense). Some species then underwent homologous recombination between the two ycf1-rpl32 copies in their plastomes and chondriomes (e.g. C. macrorhizon).

Type D ycf1-rpl32 among Cymbidium and three closely related taxa is highly conserved and shares two large deletions (Fig 2). The first hypothesis therefore must assume that two deletions in ycf1-rpl32 in both the plastome and chondriome have occurred at exactly the same position in all Cymbidium species and closely related taxa. However, the second hypothesis more easily explains this high level of similarity of the type D ycf1-rpl32 region among these genera because it originated in their common ancestor and mt-DNA has low substitution rate. Similarly, because the plastid ndhH genes of previously sequenced Cymbidium plastomes have been re-transferred from chondriome, it is likely that they should cluster with the mt-ndhH genes of Cymbidium section Pachyrhizanthe.

In relative terms, the plastid genome is ten times more abundant than the mitochondrial genome of D. catenatum. This means that plastid regions are easier to amplify than mt-region even if the mt-region had exactly the same primer binding sites as the plastid copy. With the exception of C. atropurpureum, only one PCR product of the plastid ndhJ-K-C region was produced from all Cymbidium species and three related species studied here, and the plastid copies of ndhJ, ndhK and ndhC all clustered as expected with the exception C. finlaysonianum and C. madidum, making it likely that the ndhJ-K-C region of these two species was from their plastome.

In contrast, the type II ndhK found in C. atropurpureum was in mitochondrial genome of C. lancifolium and C. macrorhizon, so it is likely that type II ndhJ-K-C region of C. atropurpureum was located in the chondriome. Considering the phylogenetic relationship between C. atropurpureum and C. macrorhizon [69, 72], the plastid ndhJ-K-C region might have been transferred to chondriome in the ancestor of Cymbidium. It also seems that the mt-ndhJ-K-C
Fig 5. Two hypotheses for multiple copies of ycf1-rpl32 region in Cymbidium species. C. sinense illustrates the ndhF-containing types (type A, B, C), and C. macrorhizon the ndhF-deleted type (type D) in plastome. Green and red boxes indicate plastome and chondriome, respectively. A) The ycf1-rpl32 region containing the ndhF (ancestral type) was transferred to the chondriome, and then mt-ndhF (C. sinense) and plastid ndhF (C. macrorhizon) were independently deleted. B) The ycf1-rpl32 region containing ndhF were
transferred to chondriome in the ancestor of the extant species of *Cymbidium* and closely related genera. Then, the mt-ndhF was removed from ycf1-rpl32 via gene rearrangements or gene deletion (*C. sinense*). In addition, homologous recombination between two ycf1-rpl32 regions of the plastome and chondriome occurred in some taxa or populations. As a result, ndhF was found not in the plastome but in the chondriome (e.g. in *C. macrorhizon*).

https://doi.org/10.1371/journal.pone.0187318.g005

region of *C. finlaysonianum* and *C. madidum* was replaced with its plastid counterpart via recent homologous recombination. As a result, reimported plastid ndh genes are derived from the mt-ndh copies. The clustering of ndhG and ndhH among the two organellar genomes in some *Cymbidium* species also supports the hypothesis that their plastid ndh genes were relatively recently reimported from chondriome, probably via homologous recombination.

Materials and methods

DNA extraction, sequencing, annotation

Fresh leaves of *C. finlaysonianum*, *C. devonianum* and *Grammatophyllum speciosum* were collected from the orchid collection at the Royal Botanic Gardens, Kew, and Ratcliffe Orchids, Ltd. (Hampshire, UK). Total DNA was extracted by the CTAB method [73]. Except for these three, all other genomic DNAs were taken from DNA Bank at the Royal Botanic Gardens, Kew (Table 3; http://apps.kew.org/dnabank/introduction.html). Vouchers are deposited in the spirit collection at the Royal Botanic Gardens, Kew.

Four regions including all 11 ndh genes (*ndhB, ndhJ-K-C, ndhF*, and *ndhD-E-G-I-A-H*) were assembled from the plastomes of *Cymbidium* [43]. Except for the *ndhF* region, primers were designed for three regions to sequence the full length of each region. In the *ndhF* region, there were a number of homopolymers near both ends. According to previous studies [43] and submitted sequences, this gene was completely deleted in some accessions of *Cymbidium*. Therefore, primers were designed just to confirm absence/presence of *ndhF* in each accession.

The four regions in each species sampled were amplified as follows: 95˚C 5min, (95˚C 30 sec—50~55˚C 30sec—65~72˚C 2min) × 31 cycles, 65~72˚C 2min using TaKaRa Premix Taq. PCR products were purified with Qiagen kits using the protocol of the manufacturer and were sequenced using Big-Dye chemistry on an ABI3730XL sequencer following the protocols of the manufacturer. All sequences were assembled by taxon and region using Geneious [74]. We annotated 11 ndh genes in each *Cymbidium* and three closely related taxa using complete sequenced plastome sequences in Orchidaceae.

Detecting ndh genes in chondriome

We used the data set from the Sequence Read Archive [75] and *Cymbidium* data generated by Kim (not published) to confirm if ndh genes had been translocated to the chondriome (Table 2). We slightly modified the assembly method of Kim et al. [30] (Fig 6). Read ends were trimmed with an error probability limit of 0.01, and then reads under 40 bp and their counterpart reads were removed from data set. Each data set was aligned to the chondriome sequence of *Phoenix dactylifera* [65] under the medium sensitivity option in Geneious [74]. Then, the reads assembled with the reference were extracted and re-assembled using *de novo* assembly in Geneious with zero mismatch and gaps [74]. Several contigs were generated, and reads were re-aligned to them with zero mismatch and gaps with 25 iterations. We generated consensus contigs and aligned them by *de novo* assembly. The resulting contigs were re-used as reference sequences.

Whenever this process was repeated, the number of contigs was reduced, and lengths of resulting contigs extended, and this cycle was repeated until the contigs produced were not
Table 3. PCR amplified ndh genes among 23 Cymbidium species including hybrids and three closely related taxa.

| species                  | region | accession   | species                  | region | accession   |
|--------------------------|--------|-------------|--------------------------|--------|-------------|
| Acriopsis sp.             | ndhB   | KX962181    | Cymbidium whiteae        | ndhB   | KX962256    |
| Cymbidium atropurpureum  | ndhB   | KX962182    | Grammatophyllum speciosum| ndhB   | KX962257    |
| Cymbidium bicolor         | ndhB   | KX962183    | Thecostele secunda       | ndhB   | KX962258    |
| Cymbidium dayanum         | ndhB   | KX962184    | Cymbidium atropurpureum  | ndhJKC TYEP I | KX962234 |
| Cymbidium devonianum      | ndhB   | KX962185    | Cymbidium atropurpureum  | ndhJKC TYEP II | KX962233 |
| Cymbidium eburneum        | ndhB   | KX962186    | Acriopsis sp.            | ndh genes in SSC region | KX962259 |
| Cymbidium elegans         | ndhB   | KX962187    | Cymbidium atropurpureum  | ndhB   | KX962260    |
| Cymbidium erythraeum      | ndhB   | KX962188    | Cymbidium bicolor        | ndh genes in SSC region | KX962261 |
| Cymbidium erythrostylum   | ndhB   | KX962189    | Cymbidium dayanum        | ndh genes in SSC region | KX962262 |
| Cymbidium finlaysonianum  | ndhB   | KX962190    | Cymbidium eburneum       | ndh genes in SSC region | KX962263 |
| Cymbidium floribundum     | ndhB   | KX962191    | Cymbidium elegans        | ndh genes in SSC region | KX962264 |
| Cymbidium giganteum       | ndhB   | KX962192    | Cymbidium erythraeum     | ndh genes in SSC region | KX962265 |
| Cymbidium goeringii       | ndhB   | KX962193    | Cymbidium erythrostylum  | ndh genes in SSC region | KX962266 |
| Cymbidium hookerianum     | ndhB   | KX962194    | Cymbidium finlaysonianum | ndh genes in SSC region | KX962267 |
| Cymbidium insigne         | ndhB   | KX962195    | Cymbidium floribundum    | ndh genes in SSC region | KX962268 |
| Cymbidium iridioides      | ndhB   | KX962196    | Cymbidium giganteum      | ndh genes in SSC region | KX962269 |
| Cymbidium lowianum        | ndhB   | KX962197    | Cymbidium goeringii      | ndh genes in SSC region | KX962270 |
| Cymbidium madidum         | ndhB   | KX962198    | Cymbidium hookerianum    | ndh genes in SSC region | KX962271 |
| Cymbidium mastersii       | ndhB   | KX962199    | Cymbidium insigne        | ndh genes in SSC region | KX962272 |
| Cymbidium pumilum         | ndhB   | KX962200    | Cymbidium iridioides     | ndh genes in SSC region | KX962273 |
| Cymbidium sanderae        | ndhB   | KX962201    | Cymbidium lowianum       | ndh genes in SSC region | KX962274 |
| Cymbidium suavissimum     | ndhB   | KX962202    | Cymbidium madidum        | ndh genes in SSC region | KX962275 |
| Cymbidium tigrinum        | ndhB   | KX962203    | Cymbidium mastersii      | ndh genes in SSC region | KX962276 |
| Cymbidium whiteae         | ndhB   | KX962204    | Cymbidium pumilum        | ndh genes in SSC region | KX962277 |
| Grammatophyllum speciosum | ndhB   | KX962205    | Cymbidium sanderae       | ndh genes in SSC region | KX962278 |
| Thecostele secunda        | ndhB   | KX962206    | Cymbidium suavissimum    | ndh genes in SSC region | KX962279 |
| Acrospis sp.              | ndhD   | KX962207    | Cymbidium tigrinum       | ndh genes in SSC region | KX962280 |
| Cymbidium atropurpureum   | ndhD   | KX962208    | Cymbidium whiteae        | ndh genes in SSC region | KX962281 |
| Cymbidium bicolor         | ndhD   | KX962209    | Grammatophyllum speciosum| ndh genes in SSC region | KX962282 |
| Cymbidium dayanum         | ndhD   | KX962210    | Thecostele secunda       | ndh genes in SSC region | KX962283 |
| Cymbidium eburneum        | ndhD   | KX962211    | Cymbidium bicolor        | ycf1-ripl32_type I | KY006886 |
| Cymbidium elegans         | ndhD   | KX962212    | Cymbidium dayanum        | ycf1-ripl32_type I | KY006885 |
| Cymbidium erythraeum      | ndhD   | KX962213    | Cymbidium elegans        | ycf1-ripl32_type I | KY006884 |
| Cymbidium erythrostylum   | ndhD   | KX962214    | Cymbidium erythraeum     | ycf1-ripl32_type I | KY006878 |
| Cymbidium finlaysonianum  | ndhD   | KX962215    | Cymbidium giganteum      | ycf1-ripl32_type I | KY006880 |
| Cymbidium floribundum     | ndhD   | KX962216    | Cymbidium insigne        | ycf1-ripl32_type I | KY006881 |
| Cymbidium giganteum       | ndhD   | KX962217    | Cymbidium iridioides     | ycf1-ripl32_type I | KY006883 |
| Cymbidium goeringii       | ndhD   | KX962218    | Cymbidium lowianum       | ycf1-ripl32_type I | KY006882 |
| Cymbidium hookerianum     | ndhD   | KX962219    | Cymbidium mastersii      | ycf1-ripl32_type I | KY006888 |
| Cymbidium insigne         | ndhD   | KX962220    | Cymbidium sanderae       | ycf1-ripl32_type I | KY006879 |
| Cymbidium iridioides      | ndhD   | KX962221    | Grammatophyllum speciosum| ycf1-ripl32_type I | KY006887 |
| Cymbidium lowianum        | ndhD   | KX962222    | Cymbidium atropurpureum  | ycf1-ripl32_type II | KY006898 |
| Cymbidium madidum         | ndhD   | KX962223    | Cymbidium ensifolium     | ycf1-ripl32_type II | KY006890 |
| Cymbidium mastersii       | ndhD   | KX962224    | Cymbidium finlaysonianum | ycf1-ripl32_type II | KY006896 |
| Cymbidium pumilum         | ndhD   | KX962225    | Cymbidium floribundum    | ycf1-ripl32_type II | KY006899 |
| Cymbidium sanderae        | ndhD   | KX962226    | Cymbidium hookerianum    | ycf1-ripl32_type II | KY006892 |
| Cymbidium suavissimum     | ndhD   | KX962227    | Cymbidium kanran         | ycf1-ripl32_type II | KY006897 |

(Continued)
extended. To prevent misassembled contigs, only paired reads that matched and upstream or downstream sequence were used throughout the assembly process.

All contigs were investigated for similarity to chondriome sequences using BLAST [76]. Thereafter, mitochondrial contigs were annotated in comparison with their own plastomes. To distinguish the location of genes, genes in the plastome are prefixed with pt- and those in chondriome are prefixed with mt-. Information on mt-ndh genes is described in Table 2.

Phylogenetic analysis of ndh genes in both organellar genomes in Orchidaceae

The pt- and mt-ndh genes in Cymbidium and three closely related taxa were sequenced in this paper. In addition, 55 plastomes (S2 Table) and 38 chondriome sequences (S3 Table) were downloaded from NCBI. The three Phalaenopsis plastomes and Vanilla planifolia have a 76 ~ 83 bp inversion upstream of the 3’ end of ndhB. Each ndh gene set was aligned via MAFFT alignment [77].

The ndhF gene was excluded from phylogenetic analysis because many species contained two types of ndhF genes, and it was difficult to determine where they were located in the organellar genomes. Introns in ndhA and ndhB were also removed from data set. The best-fit substitution model for each data set was determined using jModeltest2 [78].

### Table 3. (Continued)

| species                  | region | accession  | species                | region  | accession |
|--------------------------|--------|------------|------------------------|---------|-----------|
| Cymbidium tigrinum       | ndhD   | KX962228   | Cymbidium lancifolium  | ycf1-pl32_type II | KY006889  |
| Cymbidium whiteae        | ndhD   | KX962229   | Cymbidium pumilum      | ycf1-pl32_type II | KY006894  |
| Grammatophyllum speciosum| ndhD   | KX962230   | Cymbidium sinense      | ycf1-pl32_type II | KY006891  |
| Thecostele secunda       | ndhD   | KX962231   | Cymbidium suavissimum  | ycf1-pl32_type II | KY006893  |
| Acriopsis sp.            | ndhJKC | KX962232   | Cymbidium tigrinum     | ycf1-pl32_type II | KY006895  |
| Cymbidium bicolor        | ndhJKC | KX962235   | Acriopsis sp.          | ycf1-pl32_type III| KY006900 |
| Cymbidium dayanum        | ndhJKC | KX962236   | Cymbidium madidum      | ycf1-pl32_type III| KY006901 |
| Cymbidium devonianum     | ndhJKC | KX962237   | Thecostele secunda     | ycf1-pl32_type III| KY006902 |
| Cymbidium eburneum       | ndhJKC | KX962238   | Acriopsis sp.          | ycf1-pl32_type IV | KY006918  |
| Cymbidium elegans        | ndhJKC | KX962239   | Cymbidium bicolor      | ycf1-pl32_type IV | KY006905  |
| Cymbidium erythraeum     | ndhJKC | KX962240   | Cymbidium devonianum   | ycf1-pl32_type IV | KY006913  |
| Cymbidium erythrostylum  | ndhJKC | KX962241   | Cymbidium eburneum     | ycf1-pl32_type IV | KY006915  |
| Cymbidium finlaysonianum | ndhJKC | KX962242   | Cymbidium ensifolium   | ycf1-pl32_type IV | KY006904  |
| Cymbidium floribundum    | ndhJKC | KX962243   | Cymbidium erythrostylum| ycf1-pl32_type IV | KY006911  |
| Cymbidium giganteum      | ndhJKC | KX962244   | Cymbidium floribundum  | ycf1-pl32_type IV | KY006908  |
| Cymbidium goeringii      | ndhJKC | KX962245   | Cymbidium goeringii    | ycf1-pl32_type IV | KY006920  |
| Cymbidium hookerianum    | ndhJKC | KX962246   | Cymbidium kanran       | ycf1-pl32_type IV | KY006907  |
| Cymbidium insigne        | ndhJKC | KX962247   | Cymbidium lancifolium  | ycf1-pl32_type IV | KY006919  |
| Cymbidium indioides      | ndhJKC | KX962248   | Cymbidium lowianum     | ycf1-pl32_type IV | KY006909  |
| Cymbidium lowianum       | ndhJKC | KX962249   | Cymbidium mastersii    | ycf1-pl32_type IV | KY006914  |
| Cymbidium madidum        | ndhJKC | KX962250   | Cymbidium pumilum      | ycf1-pl32_type IV | KY006912  |
| Cymbidium mastersii      | ndhJKC | KX962251   | Cymbidium sinense      | ycf1-pl32_type IV | KY006906  |
| Cymbidium pumilum        | ndhJKC | KX962252   | Cymbidium tigrinum     | ycf1-pl32_type IV | KY006903  |
| Cymbidium sanderae       | ndhJKC | KX962253   | Cymbidium whiteae      | ycf1-pl32_type IV | KY006917  |
| Cymbidium suavissimum    | ndhJKC | KX962254   | Grammatophyllum speciosum| ycf1-pl32_type IV | KY006910  |
| Cymbidium tigrinum       | ndhJKC | KX962255   | Thecostele secunda     | ycf1-pl32_type IV | KY006916  |

https://doi.org/10.1371/journal.pone.0187318.t003
Supporting information

S1 Fig. Alignment of ndh genes of 23 Cymbidium species and three closely related genera. Masdevallia coccinea ndh genes were used as reference. A) ndhB region. B) ndhJ-K-C region. Grey and black in the alignment indicate agreement and disagreement with the consensus sequence, respectively. Red in the alignment indicates ambiguous sites. Black bars at
the bottom of the alignment indicate coding regions. Blue arrows and numbers at the bottom of the alignment indicate direct repeat sequences and length of repeat sequence, respectively.

(SPS)

S2 Fig. Alignment of ndh genes of 23 Cymbidium species and three closely related genera. Masdevallia coccinea ndh genes were used as reference. A) ndhD region. B) ndhE-G-I-A-H region Grey and black in the alignment indicate agreement and disagreement to consensus sequence, respectively. Red in the alignment indicates ambiguous sites. Black bars at the bottom of the alignment indicate coding regions. Blue arrows and numbers at the bottom of the alignment indicate direct repeat sequences and length of repeat sequence, respectively. Vertical red dotted lines indicate the end point of deletions. Green and blue lines at the bottom indicate AT- and GC-content of C. elegans.

(SPS)

S3 Fig. Ten gene trees produced by the Bayesian analysis.

(SPS)

S4 Fig. Gene conversion in the plastid ndhF gene.

(SPS)

S1 Table. The ndh status of 743 plastomes.

(DOCX)

S2 Table. The 55 plastome sequences for phylogenetic study of ndh genes.

(DOCX)

S3 Table. The mt-ndh genes for phylogenetic study of ndh genes.

(DOCX)

Acknowledgments

This work was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (2015R1A6A3A03020621). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Author Contributions

Conceptualization: Hyoung Tae Kim.

Data curation: Hyoung Tae Kim.

Investigation: Hyoung Tae Kim.

Methodology: Hyoung Tae Kim.

Software: Hyoung Tae Kim.

Supervision: Mark W. Chase.

Visualization: Hyoung Tae Kim.

Writing – original draft: Hyoung Tae Kim.

Writing – review & editing: Mark W. Chase.
References

1. Shinozaki K, Ohme M, Tanaka M, Wakasugi T, Hayashida N, Matsubayashi T, et al. The complete nucleotide sequence of the tobacco chloroplast genome: its gene organization and expression. EMBO J. 1986; 5(9):2043–2049. PMID: 16453699

2. Ohyama K, Fukuzawa H, Kohchi T, Shirai H, Sano T, Sano S, et al. Chloroplast gene organization deduced from complete sequence of liverwort Marchantia polymorpha chloroplast DNA. Nature. 1986; 322(6079):572–574. https://doi.org/10.1038/322572a0

3. Sazanov LA, Burrows PA, Nixon PJ. The plastid ndh genes code for an NADH-specific dehydrogenase: isolation of a complex I analogue from pea thylakoid membranes. Proc Natl Acad Sci U S A. 1998; 95(3):1319–1324. https://doi.org/10.1073/pnas.95.3.1319 PMID: 9448329

4. Burrows PA, Sazanov LA, Svab Z, Maliga P, Nixon PJ. Identification of a functional respiratory complex in chloroplasts through analysis of tobacco mutants containing disrupted plastid ndh genes. EMBO J. 1998; 17(4):868–876. https://doi.org/10.1093/emboj/17.4.868 PMID: 9463365

5. Bendall DS, Manasse RS. Cyclic photophosphorylation and electron transport. Biochim Biophys Acta. 1995; 1229(1):23–38. https://doi.org/10.1016/0005-2728(94)00195-b

6. Yamori W, Shikanai T, Makino A. Photosystem I cyclic electron flow via chloroplast NADH dehydrogenase-like complex performs a physiological role for photosynthesis at low light. Sci Rep. 2015; 5:13908. https://doi.org/10.1038/srep13908 PMID: 26389849

7. Ueda M, Kuniyoshi T, Yamamoto H, Sugimoto K, Ishizaki K, Kohchi T, et al. Composition and physiological function of the chloroplast NADH dehydrogenase-like complex in Marchantia polymorpha. Plant J. 2012; 72(4):683–693. https://doi.org/10.1111/j.1365-313X.2012.05115.x PMID: 22862786

8. Munekage Y, Hojo M, Meurer J, Endo T, Tasaka M, Shikanai T. PGR5 is involved in cyclic electron flow around photosystem I and is essential for photoprotection in Arabidopsis. Cell. 2002; 110(3):361–371. https://doi.org/10.1016/s0092-8674(02)00867-x PMID: 12176323

9. Lin CS, Chen JJ, Huang YT, Chan MT, Daniell H, Chang WJ, et al. The location and translocation of ndh genes of chloroplast origin in the Orchidaceae family. Sci Rep. 2015; 5:9040. https://doi.org/10.1038/srep09040 PMID: 25761566

10. Wakasugi T, Tsuzuki J, Ito S, Nakashima K, Tsuzuki T, Sugiuira M. Loss of all ndh genes code for an NADH-specific dehydrogenase: isolation of a complex I analogue from pea thylakoid membranes. Proc Natl Acad Sci U S A. 1998; 95(3):1319–1324. https://doi.org/10.1073/pnas.95.3.1319 PMID: 9448329

11. Nystedt B, Street NR, Wetterbom A, Zuccolo A, Lin YC, Scofield DG, et al. The Norway spruce genome sequence and conifer genome evolution. Nature. 2013; 497(7451):579–584. https://doi.org/10.1038/nature12211 PMID: 23698360

12. McCoy SR, Kuehl JV, Boore JL, Raubeson LA. The complete plastid genome sequence of Welwitschia mirabilis: an unusually compact plastome with accelerated divergence rates. BMC Evol Biol. 2008; 8:130. https://doi.org/10.1186/1471-2148-8-130 PMID: 18452621

13. Forrest LL. Deep sequencing of Ptilidium (Ptilidiaceae) suggests evolutionary stasis in liverwort plastid genome structure. Plant Ecol Evol. 2011; 144(1):29–43. https://doi.org/10.1051/plecoev.2011.535

14. Kim HT, Chung MG, Kim KJ. Chloroplast genome evolution in early diverged leptosporangiate ferns. Mol Cells. 2014; 37(5):372–382. https://doi.org/10.14348/molcells.2014.2296 PMID: 24823358

15. Gao L, Yi X, Yang YX, Su YJ, Wang T. Complete chloroplast genome sequence of a tree fern Marchantia polymorpha (Ptilidiaceae) suggests evolution ary stasis in liverwort plastid DNA. Nature. 1986; 322(6079):572–574. https://doi.org/10.1038/322572a0

16. Bendall DS, Manasse RS. Cyclic photophosphorylation and electron transport. Biochim Biophys Acta. 1995; 1229(1):23–38. https://doi.org/10.1016/0005-2728(94)00195-b

17. Sazanov LA, Burrows PA, Nixon PJ. The plastid ndh genes code for an NADH-specific dehydrogenase: isolation of a complex I analogue from pea thylakoid membranes. Proc Natl Acad Sci U S A. 1998; 95(3):1319–1324. https://doi.org/10.1073/pnas.95.3.1319 PMID: 9448329

18. Nystedt B, Street NR, Wetterbom A, Zuccolo A, Lin YC, Scofield DG, et al. The Norway spruce genome sequence and conifer genome evolution. Nature. 2013; 497(7451):579–584. https://doi.org/10.1038/nature12211 PMID: 23698360

19. McCoy SR, Kuehl JV, Boore JL, Raubeson LA. The complete plastid genome sequence of Welwitschia mirabilis: an unusually compact plastome with accelerated divergence rates. BMC Evol Biol. 2008; 8:130. https://doi.org/10.1186/1471-2148-8-130 PMID: 18452621

20. McColl MA, Sazanov LA, et al. The plastid genome of the parasitic liverwort Marchanti a polymorpha. Nature. 1986; 322(6079):572–574. https://doi.org/10.1038/322572a0

21. Sazanov LA, Burrows PA, Nixon PJ. The plastid ndh genes code for an NADH-specific dehydrogenase: isolation of a complex I analogue from pea thylakoid membranes. Proc Natl Acad Sci U S A. 1998; 95(3):1319–1324. https://doi.org/10.1073/pnas.95.3.1319 PMID: 9448329

22. Nystedt B, Street NR, Wetterbom A, Zuccolo A, Lin YC, Scofield DG, et al. The Norway spruce genome sequence and conifer genome evolution. Nature. 2013; 497(7451):579–584. https://doi.org/10.1038/nature12211 PMID: 23698360

23. McCoy SR, Kuehl JV, Boore JL, Raubeson LA. The complete plastid genome sequence of Welwitschia mirabilis: an unusually compact plastome with accelerated divergence rates. BMC Evol Biol. 2008; 8:130. https://doi.org/10.1186/1471-2148-8-130 PMID: 18452621
21. Wu CS, Lai YT, Lin CP, Wang YN, Chaw SM. Evolution of reduced and compact chloroplast genomes (cpDNAs) in green algae: selection toward a lower-cost strategy. Mol Phylogenet Evol. 2009; 52(1):115–124. https://doi.org/10.1016/j.ympev.2008.12.026 PMID: 19166950

22. dePamphilis CW, Palmer JD. Loss of photosynthetic and chlororespiratory genes from the plastid genome of a parasitic flowering plant. Nature. 1990; 348(6299):337–339. https://doi.org/10.1038/348337a0 PMID: 2250706

23. Haberhausen G, Zetsche K. Functional loss of all ndh genes in an otherwise relatively unaltered plastid genome of the holoparasitic plant species Cuscuta reflexa. Plant Mol Biol. 1994; 24(1):217–222. https://doi.org/10.1007/bf00040588 PMID: 8111019

24. McNeal JR, Kuehl JV, Boore JL, de Pamphilis CW. Complete plastid genome sequences suggest strong selection for retention of photosynthetic genes in the plastid genome of Cuscuta. BMC Plant Biol. 2007; 7:57. https://doi.org/10.1186/1471-2229-7-57 PMID: 17956636

25. Logacheva MD, Scheikunov MI, Nuraliev MS, Samigullin TH, Penin AA. The plastid genome of mycotrophic monocot Petrosavia stellaris exhibits both gene losses and multiple rearrangements. Genome Biol Evol. 2014; 6(1):238–246. https://doi.org/10.1093/gbe/evu001 PMID: 24398375

26. Peredo EL, King UM, Les DH. The plastid genome of Cymbidium. Plant Mol Biol. 2011; 76(3–5):263–272. https://doi.org/10.1007/s11103-011-9753-5

27. Guisinger MM, Kuehl JV, Boore JL, de Pamphilis CW. Complete plastid genome sequences suggest extreme reconfiguration of plastid genomes in the angiosperm family Geraniaceae: rearrangements, repeats, and codon usage. Mol Biol Evol. 2011; 28(1):583–600. https://doi.org/10.1093/molbev/mqs229 PMID: 20805190

28. Blazier J, Guisinger MM, Jansen RK. Recent loss of plastid-encoded ndh genes within Erodium (Geraniaceae). Plant Mol Biol. 2011; 76(3–5):263–272. https://doi.org/10.1007/s11103-011-9753-5 PMID: 21327834

29. Sanderson MJ, Copetti D, Burquez A, Bustamante E, Charboneau JL, Eqviarte LE, et al. Exceptional reduction of the plastid genome of saguaro cactus (Carnegiea gigantea): Loss of the ndh gene suite and inverted repeat. Am J Bot. 2015; 102(7):1115–1127. https://doi.org/10.3732/ajb.1500184 PMID: 26199368

30. Kim HT, Kim JS, Moore MJ, Neubig KM, Williams NH, Whitten WM, et al. 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. 2015; 10(11):e0142215. https://doi.org/10.1371/journal.pone.0142215 PMID: 26558895

31. Weng ML, Blazier JC, Govindu M, Jansen RK. Reconstruction of the ancestral plastid genome in Geraniaceae reveals a correlation between genome rearrangements, repeats, and nucleotide substitution rates. Mol Biol Evol. 2014; 31(3):645–659. https://doi.org/10.1093/molbev/mst257 PMID: 24396877

32. Kim KA, Cheon KS, Jang SK, Yoo KO. Complete chloroplast genome sequence of Adenophora remotiflora (Campanulaceae). Mitochondrial DNA A DNA Mapp Seq Anal. 2016; 27(4):2963–2964. https://doi.org/10.3109/19401736.2015.1060461 PMID: 26119125

33. Kumar S, Hahn FM, McMahan CM, Cornish K, Whalen MC. Comparative analysis of the complete sequence of the plastid genome of Parthenium argentatum and identification of DNA barcodes to differentiate Parthenium species and lines. BMC Plant Biol. 2009; 9(1):131. https://doi.org/10.1186/1471-2229-9-131 PMID: 19917140

34. Li X, Zhang TC, Qiao Q, Ren Z, Zhao J, Yonezawa T, et al. Complete chloroplast genome sequence of the parasitic flowering plant species Cuscuta reflexa and Cuscuta gronovii (Chenopodiaceae). PLoS One. 2013; 8(3):e58747. https://doi.org/10.1371/journal.pone.0058747 PMID: 23554920

35. Wicke S, Muller KF, de Pamphilis CW, Quandt D, Wickett NJ, Zhang Y, et al. Mechanisms of functional and physical genome reduction in photosynthetic and nonphotosynthetic parasitic plants of the broomrape family. Plant Cell. 2013; 25(10):3711–3725. https://doi.org/10.1105/tpc.113.113373 PMID: 24143802

36. Wolfe KH, Morden CW, Ems SC, Palmer JD. Rapid evolution of the plastid translational apparatus in a nonphotosynthetic plant: loss or accelerated sequence evolution of tRNA and ribosomal protein genes. J Mol Evol. 1992; 35(4):304–317. PMID: 1404416

37. Funk HT, Berg S, Krupinska K, Maier UG, Krause K. Complete DNA sequences of the plastid genomes of two parasitic flowering plant species, Cuscuta reflexa and Cuscuta gronovii. BMC Plant Biol. 2007; 7:45. https://doi.org/10.1186/1471-2229-7-45 PMID: 17714582

38. Kim JS, Kim HT, Kim J-H. The largest plastid genome of monocots: a novel genome type containing AT residue repeats in the slipper orchid Cypripedium japonicum. Plant Mol Biol Rep. 2014; 33(5):1210–1220. https://doi.org/10.1007/s11105-014-0833-y
39. Pan IC, Liao DC, Wu FH, Daniell H, Singh ND, Chang C, et al. Complete chloroplast genome sequence of Neottia nidus-avis. Plant Syst. Evol. 2012; 190:62–73. https://doi.org/10.1007/j.10735 -3871/2012-04.01 PMID: 22608520

40. Wu FH, Chan MT, Liao DC, Hsu CT, Lee YW, Daniell H, et al. Complete chloroplast genome of Oncidium Gower Ramsey and evaluation of molecular markers for identification and breeding in Oncidiinae. BMC Plant Biol. 2010; 10:68. https://doi.org/10.1186/1471-2229-10-68 PMID: 20398375

41. Jheng CF, Chen TC, Lin JY, Chen TC, Wu WL, Chang CC. The comparative chloroplast genomic analysis of photosynthetic orchids and developing DNA markers to distinguish Phalaenopsis orchids. Plant Sci. 2012; 190:62–73. https://doi.org/10.1016/j.plantsci.2012.04.01 PMID: 22608520

42. Pan IC, Liao DC, Wu FH, Daniell H, Singh ND, Chang C, et al. Complete chloroplast genome sequence of an orchid model plant candidate: Erycina pusilla apply in tropical Oncidium breeding. PLoS One. 2012; 7(4):e34738. https://doi.org/10.1371/journal.pone.0034738 PMID: 22496851

43. Yang JB, Tang M, Li HT, Zhang ZR, Li DZ. Complete chloroplast genome of the genus Cymbidium: insights into the species identification, phylogenetic implications and population genetic analyses. BMC Evol Biol. 2013; 13:84. https://doi.org/10.1186/1471-2148-13-84 PMID: 23597078

44. Luo J, Hou BW, Niu ZT, Liu W, Xue QY, Ding XY. Comparative chloroplast genomes of photosynthetic orchids: insights into evolution of the Orchidaceae and development of molecular markers for phylogenetic applications. PLoS One. 2014; 9(6):e99016. https://doi.org/10.1371/journal.pone.0099016 PMID: 24911363

45. Delannoy E, Fujii S, Colas des Francs-Small C, Brundrett M, Small I. Rampant gene loss in the underground orchid Rhizanthella gardneri highlights evolutionary constraints on plastid genomes. Mol Biol Evol. 2011; 28(7):2077–2086. https://doi.org/10.1093/molbev/msr028 PMID: 21289370

46. Logacheva MD, Schelkunov MI, Penin AA. Sequencing and analysis of plastid genome in mycoheterotrophic orchid Neottia nidus-avis. Genome Biol Evol. 2011; 3:1296–1303. https://doi.org/10.1093/gbe/ evr102 PMID: 21971517

47. Barrett CF, Davis JI. The plastid genome of the mycoheterotrophic Corallorhiza striata (Orchidaceae) is in the relatively early stages of degradation. Am J Bot. 2012; 99(9):1513–1523. https://doi.org/10.3732/ ajb.1200256 PMID: 22935964

48. Barrett CF, Freudenstein JV, Li J, Mayfield-Jones DR, Perez L, Pires JC, et al. Investigating the path of plastid genome degradation in an early-transitional clade of heterotrophic orchids, and implications for heterotrophic angiosperms. Mol Biol Evol. 2014; 31(12):3095–3112. https://doi.org/10.1093/molbev/msu252 PMID: 25172958

49. Schelkunov MI, Shratnikova VY, Nuraliev MS, Selosse MA, Penin AA, Logacheva MD. Exploring the limits for reduction of plastid genomes: a case study of the mycoheterotrophic orchids Epipogium aphyllum and Epipogium roseum. Genome Biol Evol. 2015; 7(4):1179–1191. https://doi.org/10.1093/gbe/ evr102 PMID: 25835040

50. Du Puy D, Cribb P. The genus Cymbidium: Christopher Helm: Portland, Oregon.: London & Timber Press; 1988.

51. Chase MW, Cameron KM, Freudenstein JV, Pridgeon AM, Salazar G, van den Berg C, et al. An updated classification of Orchidaceae. Bot J Linn Soc. 2015; 177(2):151–174. https://doi.org/10.1111/ bjl.12234

52. Palmer JD, Herbon LA. Plant mitochondrial DNA evolved rapidly in structure, but slowly in sequence. J Mol Evol. 1988; 28(1–2):87–97. PMID: 3148748

53. Yang P, Zhou H, Qian J, Xu H, Shao Q, Li Y, et al. The complete chloroplast genome sequence of Dendrobi um officinarum. Mitochondrial DNA A DNA Mapp Seq Anal. 2016; 27(2):1262–1264. https://doi.org/10.1080/19401736.2014.945547 PMID: 25103425

54. Ruhlman TA, Jansen RK. The plastid genomes of flowering plants. Methods Mol Biol. 2014; 1132:3–38. https://doi.org/10.1007/978-1-62703-995-6_1 PMID: 24599844

55. Wang S, Shi C, Gao LZ. Plastid genome sequence of a wild woody oil species, Prinsepia utilis, provides insights into evolutionary and mutational patterns of Rosaceae plastid genomes. PLoS One. 2013; 8(9):e73946. https://doi.org/10.1371/journal.pone.0073946 PMID: 24023915

56. Yi DK, Kim KJ. Complete chloroplast genome sequences of important oilseed crop Sesamum indicum L. PLoS One. 2012; 7(5):e35872. https://doi.org/10.1371/journal.pone.0035872 PMID: 22606240

57. Xu Q, Xiong G, Li P, He F, Huang Y, Wang K, et al. Analysis of complete nucleotide sequences of 12 Gossypium chloroplast genomes: origin and evolution of allotetraploids. PLoS One. 2012; 7(8):e37128. https://doi.org/10.1371/journal.pone.0037128 PMID: 22876273

58. Palmer JD. Plastid chromosomes: structure and evolution. The molecular biology of plastids. 1991; 7:5–53.
59. Matsuoka Y, Yamazaki Y, Ogihara Y, Tsunewaki K. Whole chloroplast genome comparison of rice, maize, and wheat: implications for chloroplast gene diversification and phylogeny of cereals. Mol Biol Evol. 2002; 19(12):2084–2091. PMID: 12446800

60. Ogihara Y, Terachi T, Sasakuma T. Intramolecular recombination of chloroplast genome mediated by short direct-repeat sequences in wheat species. Proc Natl Acad Sci U S A. 1988; 85(22):8573–8577. PMID: 3186748

61. Aldrich J, Cherney BW, Merlin E. The role of insertions/deletions in the evolution of the intergenic region between psbA and trnH in the chloroplast genome. Curr Genet. 1988; 14(2):137–146. PMID: 3180272

62. Muller AE, Kamisugi Y, Gruneberg R, Niedenhof I, Horold RJ, Meyer P. Palindromic sequences and A +T-rich DNA elements promote illegitimate recombination in Nicotiana tabacum. J Mol Biol. 1999; 291(1):29–46. https://doi.org/10.1006/jmbi.1999.2957 PMID: 10438604

63. Goulding SE, Olmstead RG, Morden CW, Wolfe KH. Ebb and flow of the chloroplast inverted repeat. Mol Gen Genet. 1996; 252(1–2):195–206. PMID: 8804393

64. Cai J, Liu X, Vanneste K, Proost S, Tsai WC, Liu KW, et al. The genome sequence of the orchid Phalaenopsis equestris. Nat Genet. 2015; 47(1):65–72. https://doi.org/10.1038/ng.3149 PMID: 25420146

65. Fang Y, Wu H, Zhang T, Yang M, Yin Y, Pan L, et al. A complete sequence and transcriptomic analyses of date palm (Phoenix dactylifera L.) mitochondrial genome. PLoS One. 2012; 7(5):e37164. https://doi.org/10.1371/journal.pone.0037164 PMID: 22655034

66. Rodriguez-Moreno L, Gonzalez VM, Benjak A, Martí MC, Puigdomènech P, Aranda MA, et al. Determination of the melon chloroplast and mitochondrial genome sequences reveals that the largest reported mitochondrial genome in plants contains a significant amount of DNA having a nuclear origin. BMC Genomics. 2011; 12(1):424. https://doi.org/10.1186/1471-2164-12-424 PMID: 21854637

67. Alverson AJ, Wei X, Rice DW, Stern DB, Barry K, Palmer JD. Insights into the evolution of mitochondrial genome size from complete sequences of Citrullus lanatus and Cucurbita pepo (Cucurbitaceae). Mol Biol Evol. 2010; 27(6):1436–1448. https://doi.org/10.1093/molbev/msq029 PMID: 2018192

68. Wang D, Wu YW, Shih AC, Wu CS, Wang YN, Chaw SM. Transfer of chloroplast genomic DNA to mitochondrial genome occurred at least 300 MYA. Mol Biol Evol. 2007; 24(9):2040–2048. https://doi.org/10.1093/molbev/msq077 PMID: 17609537

69. Yokoyama J. Molecular phylogeny and character evolution of Orchidaceae. Bull Nation Sci Mus, B (Tokyo). 2002; 28(4):129–139.

70. Gustafsson AL, Verola CF, Antonelli A. Reassessment of the temporal evolution of orchids with new fossils and a Bayesian relaxed clock, with implications for the diversification of the rare South American genus Hoffmannseggella (Orchidaceae: Epidendroideae). BMC Evol Biol. 2010; 10:177. https://doi.org/10.1186/1471-2148-10-177 PMID: 20546585

71. Park S, Grewe F, Zhu A, Ruhilman TA, Sabir J, Mower JP, et al. Dynamic evolution of Geranium mitochondrial genomes through multiple horizontal and intracellular gene transfers. New Phytol. 2015; 208(2):570–583. https://doi.org/10.1111/nph.13467 PMID: 25989702

72. van den Berg C, Ryan A, Cribb PJ, Chase MW. Molecular phylogenetics of Cymbidium (Orchidaceae: Maxillariinae): Sequence data from internal transcribed spacers (ITS) of nuclear ribosomal DNA and plastid matK. Lindleyana. 2002; 17(2):102–111.

73. Doyle JJ. A rapid DNA isolation procedure for small quantities of fresh leaf tissue. Phytochem bull. 1987; 19:11–15.

74. Kearse M, Moir R, Wilson A, Stones-Havas S, Cheung M, Sturrock S, et al. Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. Bioinformatics. 2012; 28(12):1647–1649. https://doi.org/10.1093/bioinformatics/bts199 PMID: 22543367

75. Leinonen R, Sugawara H, Shumway M. International Nucleotide Sequence Database C. The sequence read archive. Nucleic Acids Res. 2011; 39(Database issue):D19–D21.

76. Aittschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. J Mol Biol. 1990; 215(3):403–410. https://doi.org/10.1016/S0022-2836(05)80360-2 PMID: 2231712

77. Katoh K, Misawa K, Kuma K, Miyata T. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. Nucleic Acids Res. 2002; 30(14):3059–3066. PMID: 12136088

78. Darrida D, Taboada GL, Doallo R, Posada D. JModelTest 2: more models, new heuristics and parallel computing. Nat Methods. 2012; 9(8):772. https://doi.org/10.1038/nmeth.2109 PMID: 22847109

79. Ronquist F, Teslenko M, van der Mark P, Ayres DL, Darling A, Hohna S, et al. MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. Syst Biol. 2012; 61(3):539–542. https://doi.org/10.1093/sysbio/sys029 PMID: 22357727

80. Miller M, Pfeiffer W, Schwartz T, editors. Creating the CIPRES Science Gateway for inference of large phylogenetic trees. Gateway Computing Environments Workshop (GCE), 2010; 2010: IEEE.