the plant journal

SEB Society for Experimental Biology

The Plant Journal (2017) **90,** 994–1006 doi: 10.1111/tpj.13525

Concomitant loss of NDH complex-related genes within chloroplast and nuclear genomes in some orchids

Choun-Sea Lin[1,*,†], Jeremy J. W. Chen[2,†], Chi-Chou Chiu[3,†], Han C. W. Hsiao[4,†], Chen-Jui Yang[5,†], Xiao-Hua Jin[6,†], James Leebens-Mack[7,†], Claude W. de Pamphilis[8], Yao-Ting Huang[9], Ling-Hung Yang[1], Wan-Jung Chang[1], Ling Kui[10], Gane Ka-Shu Wong[11], Jer-Ming Hu[5], Wen Wang[9] and Ming-Che Shih[1,*]

[1]*Agricultural Biotechnology Research Center, Academia Sinica, Taipei, Taiwan,*
[2]*Institute of Biomedical Sciences, National Chung-Hsing University, Taichung, Taiwan,*
[3]*Institute of Tropical Plant Sciences, National Cheng Kung University, Tainan, Taiwan,*
[4]*Department of Bioinformatics and Medical Engineering, Asia University, Taichung City, Taiwan,,*
[5]*Institute of Ecology and Evolutionary Biology, National Taiwan University, Taipei, Taiwan,*
[6]*Institute of Botany, Chinese Academy of Sciences, Beijing, China,*
[7]*Department of Plant Biology, University of Georgia, Athens, GA, USA,*
[8]*Department of Biology, Pennsylvania State University, State College, PA, USA,*
[9]*Department of Computer Science and Information Engineering, National Chung Cheng University, Chiayi, Taiwan,*
[10]*State Key Laboratory of Genetic Resources and Evolution, Kunming Institute of Zoology, The Chinese Academy of Sciences, Kunming, China, and*
[11]*Department of Biological Sciences, University of Alberta, Edmonton, AB, Canada*

Received 13 December 2016; revised 20 February 2017; accepted 23 February 2017; published online 4 March 2017.
*For correspondence (e-mails cslin99@gate.sinica.edu.tw; mcshih@gate.sinica.edu.tw).
†These authors contributed equally to this work.

SUMMARY

The chloroplast NAD(P)H dehydrogenase-like (NDH) complex consists of about 30 subunits from both the nuclear and chloroplast genomes and is ubiquitous across most land plants. In some orchids, such as *Phalaenopsis equestris*, *Dendrobium officinale* and *Dendrobium catenatum*, most of the 11 chloroplast genome-encoded *ndh* genes (cp-*ndh*) have been lost. Here we investigated whether functional cp-*ndh* genes have been completely lost in these orchids or whether they have been transferred and retained in the nuclear genome. Further, we assessed whether both cp-*ndh* genes and nucleus-encoded NDH-related genes can be lost, resulting in the absence of the NDH complex. Comparative analyses of the genome of *Apostasia odorata*, an orchid species with a complete complement of cp-*ndh* genes which represents the sister lineage to all other orchids, and three published orchid genome sequences for *P. equestris*, *D. officinale* and *D. catenatum*, which are all missing cp-*ndh* genes, indicated that copies of cp-*ndh* genes are not present in any of these four nuclear genomes. This observation suggests that the NDH complex is not necessary for some plants. Comparative genomic/transcriptomic analyses of currently available plastid genome sequences and nuclear transcriptome data showed that 47 out of 660 photoautotrophic plants and all the heterotrophic plants are missing plastid-encoded cp-*ndh* genes and exhibit no evidence for maintenance of a functional NDH complex. Our data indicate that the NDH complex can be lost in photoautotrophic plant species. Further, the loss of the NDH complex may increase the probability of transition from a photoautotrophic to a heterotrophic life history.

Keywords: NDH complex, orchid, *Apostasia odorata*, gene loss, photoautotrophic plant.

INTRODUCTION

Plant cells, in contrast to animal cells, contain chloroplasts (cp) that were derived from cyanobacteria through endosymbiosis approximately 1.5 billion years ago (Yoon *et al.*, 2004). As the endosymbiont evolved into an organelle, the plastid retained a greatly reduced genome with many of the cyanobacterial genes being transferred to the host nuclear genome. Although most functional cp-to-nuclear gene transfers occurred early in plastid evolution,

such transfers have been a continual evolutionary process (Timmis *et al.*, 2004), and the cp genome (plastome) of photosynthetic land plants typically encodes 68–82 conserved protein-coding genes (Martin and Herrmann, 1998; Simpson and Stern, 2002; Wicke *et al.*, 2011; Barrett *et al.*, 2014; Tiller and Bock, 2014). However, some of the protein-coding genes have been independently lost in diverse land plant lineages. Some genes lost from the cp genome have been shown to have been transferred and retained in the nuclear genome (Jansen *et al.*, 2007). In addition to cp-to-nucleus gene transfer, functional gene transfer from the cp genome to the mitochondrial genome has also been shown to occur (Wang *et al.*, 2007, 2012).

The cp NAD(P)H-dehydrogenase-like (NDH) complex is located in the thylakoid membrane and mediates photosystem I cyclic electron transport (PSI-CET) and facilitates chlororespiration (Figure 1a) (Peltier *et al.*, 2016; Yamori and Shikanai, 2016). Although it has been suggested that the cp NDH complex and mitochondrial respiratory complex I may share a common evolutionary origin, the genes encoding for them are completely distinct (Braun *et al.*, 2014; Shikanai, 2016). In etioplasts (non-green plastids), the NDH complex exists as a monomer (Peng *et al.*, 2008). In green tissues, the NDH complex forms a supercomplex with PSI through linker proteins Lhca5/6 (Figure 1b) (Peng and Shikanai, 2011). Lhca5/6 is important for stabilizing the NDH complex but not the activity monitored by chlorophyll fluorescence analysis (Peng *et al.*, 2009). The NDH complex comprises approximately 30 subunits and can be divided into five subcomplexes: A, B, lumenal (L), membrane (M) and electron donor-binding (EDB) (Figure 1b) (Shikanai, 2016); the subunits of the NDH complex are encoded by a combination of genes residing in the plastid and nuclear genomes. Among the some 30 subunits, 11 are located in the cp genome (referred as to cp-*ndh*A to cp-*ndh*K), and this number is highly conserved in most land plants. In addition to the genes encoding for the subunits, at least 10 genes in the nuclear genome encode regulatory components for the expression and assembly of the NDH complex. Mutants with impaired subunits or regulatory components often show disassembly and dysfunction of the NDH complex (Shikanai, 2016). In some photoautotrophic plants, such as some species of Pinales, Geraniaceae and Orchidaceae, almost the entire set of cp-*ndh* genes has been lost (Wakasugi *et al.*, 1994; Chang *et al.*, 2006; Wu *et al.*, 2010; Lin *et al.*, 2015; Bellot and Renner, 2016). It has been reported that all of the cp-encoded *ndh* genes are missing in the orchid *Erycina pusilla*. Although some fragments of cp-*ndh* genes have been transferred to the mitochondrial genome in *E. pusilla*, the cp-*ndh* genes included in these fragments are not intact or functional (Lin *et al.*, 2015), suggesting that functional cp-*ndh* genes may either be transferred to the nuclear genome or completely lost. In the orchids *Cypripedium formosanum*, *Habenaria longidenticulata*, *Goodyera fumata* and *Masdevallia picturata* with plastomes having a complete set of 11 cp-*ndh* genes, gene order is conserved in blocks (Lin *et al.*, 2015): there is a seven-*ndh* gene block (*ndhF-D-E-G-I-A-H*) with *rpl32* and *ccsA* located between *ndhF* and *ndhD*, and *psaC* located between *ndhD* and *ndhE*, and these 10 genes are flanked by inverted repeats composed of *ycf1*–*rps15*. In orchids in which all these seven *ndh* genes have become non-functional, *ndhF*, *ndhI* and *ndhA* are lost and *ndhD*, *ndhE*, *ndhG* and *ndhH* are pseudogenized. The order of functional genes in this region has become *ycf1*–*rps15*–*rpl32*–*ccsA*–*psaC*–*rps15*–*ycf1*.

Recently, the nuclear genomes of three orchids that are missing cp-*ndh* genes, *Phalaenopsis equestris* (Cai *et al.*, 2015), *Dendrobium officinale* (Yan *et al.*, 2015) and *Dendrobium catenatum* (Zhang *et al.*, 2016), were published. These whole-genome sequences were analyzed to test whether functional cp-*ndh* genes were transferred to the nuclear genomes when they were lost from the plastid genome. Our results showed that the lost cp-*ndh* genes were not transferred and retained in the nuclear genome. Additionally, we observed that in the orchids without cp-*ndh* genes, genes for nuclear-encoded NDH subunits and regulatory components of the NDH complex (hereafter called nuclear NDH-related genes) are also lost, suggesting that NDH activity is not required in some plants. Our results support the model recently published by Wicke *et al.* (2016), which posits that the loss of the NDH complex may be related to the transition from photoautotroph to heterotroph (Wicke *et al.*, 2016).

RESULTS

Most of the *ndh* genes lost from the cp genome are not retained in the nuclear genome

All the functional cp-*ndh* genes are absent from the cp genomes in *P. equestris*, *D. officinale* and *D. catenatum*, except for *ndhB* and *ndhE* that are retained intact in *D. officinale* and *D. catenatum* (Figure 1c, Chloroplast). To investigate whether the *ndh* genes lost in the cp genome were transferred and retained in the nuclear genome in *P. equestris*, *D. officinale* and *D. catenatum*, tBLASTN searches against the respective nuclear genomes were performed using peptide sequences of the 11 cp-*ndh* genes from *A. odorata*. Although some sequences were identified in *P. equestris* that showed similarities to *Apostasia* cp-*ndh* genes, their coding sequences were not intact (Figure S1). Similar results were obtained from the searches against *D. catenatum*, except that instead of leaving the *ndhE* gene intact in the cp genome (Figure 1c), a copy of intact *ndhE* was also transferred and retained in the nuclear genome (Figure S1 in the Supporting Information). In *D. officinale*, surprisingly, our initial searches showed that all the cp-*ndh* genes except *ndhI* were found in its

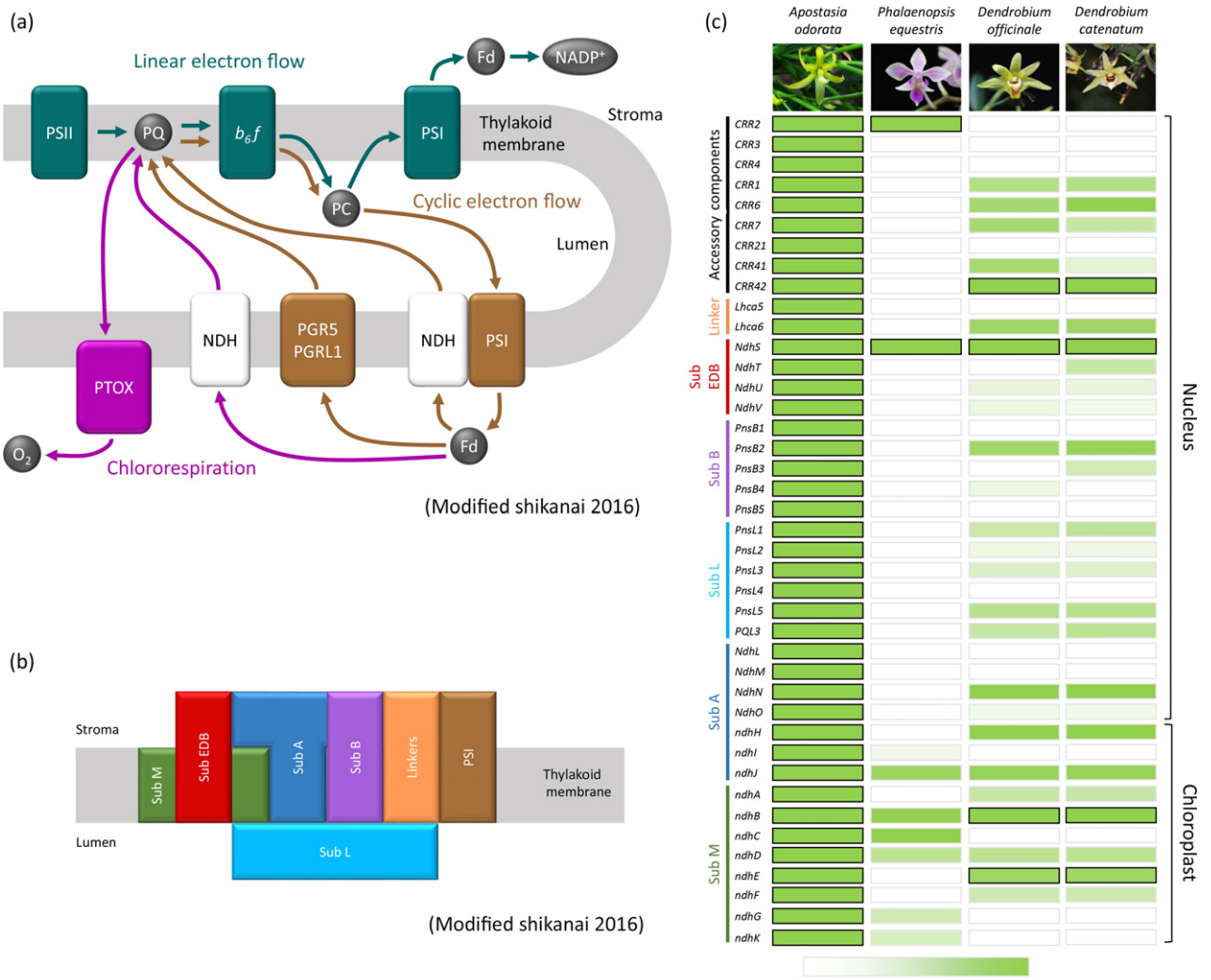


Figure 1. The NDH complex is not retained in *Phalaenopsis* and *Dendrobium*.
(a) Schematic models of the proposed functions of the NDH complex. Linear electron flow, cyclic electron flow and chlororespiration are shown. Arrowhead, electron flow direction; b_6f, cytochrome b_6/f; PSI/II, photosystem I/II; PTOX, plastid terminal oxidase; Fd, ferredoxin; PQ, plastoquinone; PC, plastocyanin. Protein complexes not present in *Phalaenopsis* and *Dendrobium* are shown in white (this study).
(b) Schematic model of the NDH-PSI supercomplex from Arabidopsis. The NDH complex can be divided into five subcomplexes (Sub): A, B, lumenal (L), membrane (M) and electron donor-binding (EDB).
(c) Genes encoding the NDH complex in the chloroplast and nuclear genomes are concomitantly lost in *Phalaenopsis* and *Dendrobium*. The amino acid sequences of the NDH-related genes in Arabidopsis were used as queries to identify the NDH-related genes in *Apostasia odorata*. Genes identified from *A. odorata* were then used as queries to identify NDH-related genes of the corresponding cp or nuclear genome in *Phalaenopsis equestris*, *Dendrobium officinale* and *Dendrobium catenatum*. The percentage of the query sequence that overlaps the subject sequence is represented with a white-to-green color gradient. Genes deduced to encode for an intact open reading frame are marked with black frames. Genes encoded in the nuclear or chloroplast genome, and the subcomplex they belong to are indicated. [Colour figure can be viewed at wileyonlinelibrary.com].

nuclear genome. It has been reported that *D. officinale* and *D. catenatum* are closely related, or may even turn out to be the same species (Jin and Huang, 2015). This prompted us to reinvestigate the identities of the cp-*ndh* genes retained in the nuclear genome of *D. officinale*. Advanced searches showed that in *D. officinale*, among the ten transferred cp-*ndh* genes, eight are almost identical to the cp-*ndh* genes of *Morus mongolica*, suggesting that they were the result of contamination by cp DNA of *M. mongolica* during sampling. Thus, in *D. officinale*, *ndhB* is present in both the cp and the nuclear genome, and *ndhH* is the only gene which was deleted from the cp genome and retained in the nuclear genome. Together, our searches showed that in *P. equestris*, *D. officinale* and *D. catenatum*, most of the deleted cp-*ndh* genes were not transferred to the nuclear genomes. Similar results were also obtained from the searches against the nuclear genomes of two cp-*ndh*-free gymnosperms, *Picea abies* (Nystedt *et al.*, 2013; Ranade *et al.*, 2016) and *Pinus taeda* (Neale *et al.*, 2014) (Figure S2).

Nuclear NDH-related genes are also lost in orchids that do not have cp-*ndh* genes

Mutants impaired in the regulatory components or subunits of the NDH complex often exhibit disassembly and dysfunction of the NDH complex (Peltier *et al.*, 2016). We were next interested to know whether the nuclear NDH-related genes were also lost in those orchids in which cp-*ndh* genes had been deleted. To this end, nuclear genomic sequences of an orchid that contains a complete complement of cp-*ndh* genes (hereafter called cp-*ndh*-complete orchid) were needed for comparison. We therefore sequenced the nuclear genome of *A. odorata*, a species with a complete set of plastome-encoded cp-*ndh* genes and a small nuclear genome size (Jersakova *et al.*, 2013). Thirty nuclear-encoded NDH-related genes were selected as a reference for the NDH complex in the *A. odorata* nuclear genome (Figure 1c). The tBLASTN searches using 30 nuclear NDH-related peptide sequences from *A. odorata* were performed against nuclear genomes of *P. equestris*, *D. officinale* and *D. catenatum*. Most of the nuclear-encoded NDH complex subunit genes were not found in *P. equestris*, *D. officinale* and *D. catenatum*. Although some sequence fragments showed similarities to *Apostasia* NDH-related genes, they were not intact (Figure 1c). Notably, in the cp-*ndh*-free orchids analyzed, only *NdhS* was still present. With regard to the regulatory components, *CRR2*, a gene essential for the expression of *ndhB*, was detected in *P. equestris*, and *CRR42*, a gene that functions in the assembly of the NDH complex, was detected in both *D. officinale* and *D. catenatum*.

To confirm that the observation of losses of nuclear NDH-related genes in *Pha. equestris*, *D. officinale* and *D. catenatum* was not due to the low-coverage genome sequencing of the genomes (*Phalaenopsis*, 1.16 Gb; *Dendrobium*, 1.35 Gb), genes related to the other protein complexes in the thylakoid membrane were searched. Our searches showed that nuclear genes encoding for PSI, PSII, cytochrome b_6/f complex (cyt b_6/f), F-type ATPase and light-harvesting complexes (LHCs) are all present in *P. equestris*, *D. officinale* and *D. catenatum* (Figure 2, Data S1, Genomes panel), suggesting that the absence of nuclear NDH-related genes is not due to the low-coverage genome sequencing. *Phalaenopsis equestris* and *D. officinale* both have transcriptome data available. No NDH-related gene transcripts were detected in RNA-Seq data, supporting the notion that the nuclear NDH-related genes are indeed lost from the *P. equestris*, *D. officinale* and *D. catenatum* lineages (Data S2). Analysis of genomes for the conifer species *Picea abies* (Nystedt *et al.*, 2013; Ranade *et al.*, 2016) and *Pinus taeda* (Neale *et al.*, 2014) also revealed correlated losses of plastid and nuclear NDH-related genes (Figure S3).

Losses of nuclear NDH-related genes are locus-specific

The losses of functional *ndh* genes are limited to those taxa with cp genomes devoid of cp-*ndh* genes (Wu *et al.*, 2010). We first searched whether the gene orders in the vicinity of nuclear NDH-related genes are conserved among five selected angiosperms, *Amborella trichopoda*, *Arabidopsis thaliana*, banana (*Musa acuminata*), rice (*Oryza sativa* Japonica) and Moso bamboo (*Phyllostachys heterocycla*). In the five angiosperm genomes analyzed, although the gene order around the nuclear NDH-related genes is not strictly conserved, phylogenetically closely related species show relatively similar structural organization for some nuclear NDH-related genes (Data S3). Among the 30 nuclear NDH-related genes, 21 showed that their neighboring genes in Moso bamboo and rice are orthologous. For example, orthologs of Arabidopsis *At2g45720* in Moso bamboo and rice are located adjacent to *CRR1*, and this is not observed in *Amborella*, Arabidopsis and banana. Orthologs of *At5g08060* and *At3g29400* in Moso bamboo and rice are located near *CRR7*. Moreover, the two genes upstream and two genes downstream of *CRR6*, *NdhN* and *Lhca5* in both Moso bamboo and rice are all orthologous and arranged in the same or the opposite order. Hence, in order to better investigate whether the losses of functional nuclear NDH-related genes in *P. equestris*, *D. officinale* and *D. catenatum* are locus-specific, the gene orders in the vicinity of nuclear NDH-related genes in the cp-*ndh*-complete orchid *A. odorata* were retrieved for comparison. In the nuclear genome of *A. odorata*, all 30 nuclear NDH-related genes were located on separate contigs. Five contigs in *A. odorata* were removed from our analyses because these contigs contain only the NDH-related gene without any neighboring genes for further comparison. Nine contigs in *P. equestris*, two in *D. officinale* and one in *D. catenatum* containing only the orthologous gene of one of the neighboring genes of the nuclear NDH-related genes in *A. odorata* were also removed for the same reason. We observed that functional genes flanked by seven, nine and eight nuclear NDH-related genes in *A. odorata* are neighbors in *P. equestris*, *D. officinale* and *D. catenatum*, respectively (Data S2). For instance, as shown in Figure 3, *NdhM* in *A. odorata* is located between genes encoding for translin family protein and brassinosteroid-responsive RING-H2 (BRH1). In the genomes of *P. equestris*, *D. officinale* and *D. catenatum*, these two genes become adjacent. No annotated genes or traces of the *NdhM* sequence were detected between genes for the translin family protein and *BRH1*. These results further confirm concomitant loss of cp-*ndh* genes and NDH-related genes in the *P. equestris*, *D. officinale* and *D. catenatum* lineages.

The NDH complex plays a role in cyclic electron transport in photosynthesis (Peltier *et al.*, 2016), and may share

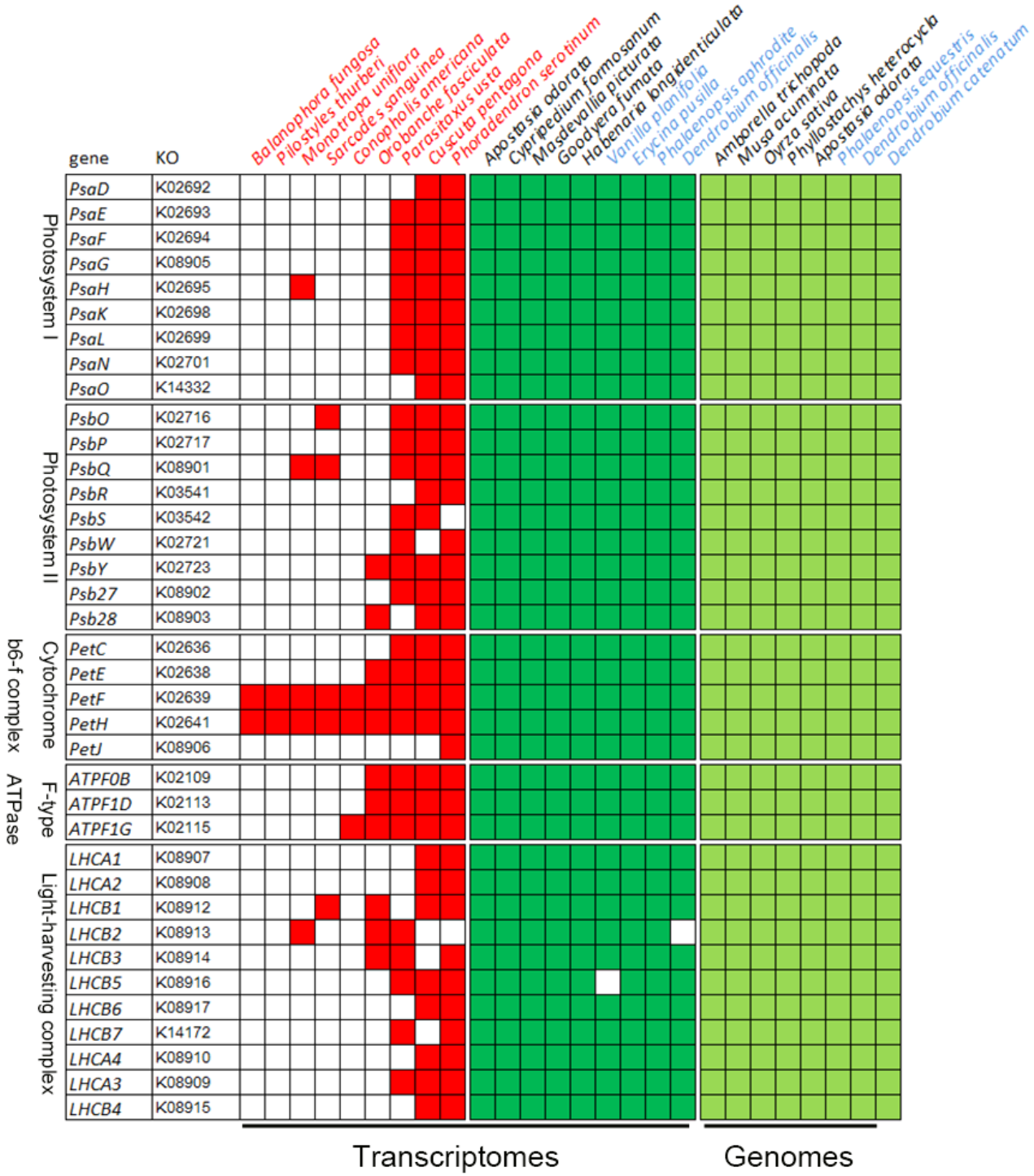


Figure 2. Transcriptomic/genomic analyses of photosynthesis-related genes in selected photoautotrophic and heterotrophic plants (with NDH-related genes excluded).

The amino acid sequences of Arabidopsis in KEGG photosynthesis pathways were used as queries for tBLASTN searches against public plant transcriptome/genome databases as described previously (E-value $< 1 \times 10^{-4}$) (Matasci *et al.*, 2014; Ruhlman *et al.*, 2015). The color-filled boxes indicate the detection of genes in the genome or the detection of gene expression in the transcriptome: olive, genomes of photoautotrophic plants; dark-green, orchid transcriptomes; red, transcriptomes of heterotrophic plants. Species names: blue, photoautotrophic cp-*ndh*-free orchids; black, cp-*ndh*-complete photoautotrophs; red, heterotrophic plants. Detailed information is shown in Data S1. [Colour figure can be viewed at wileyonlinelibrary.com].

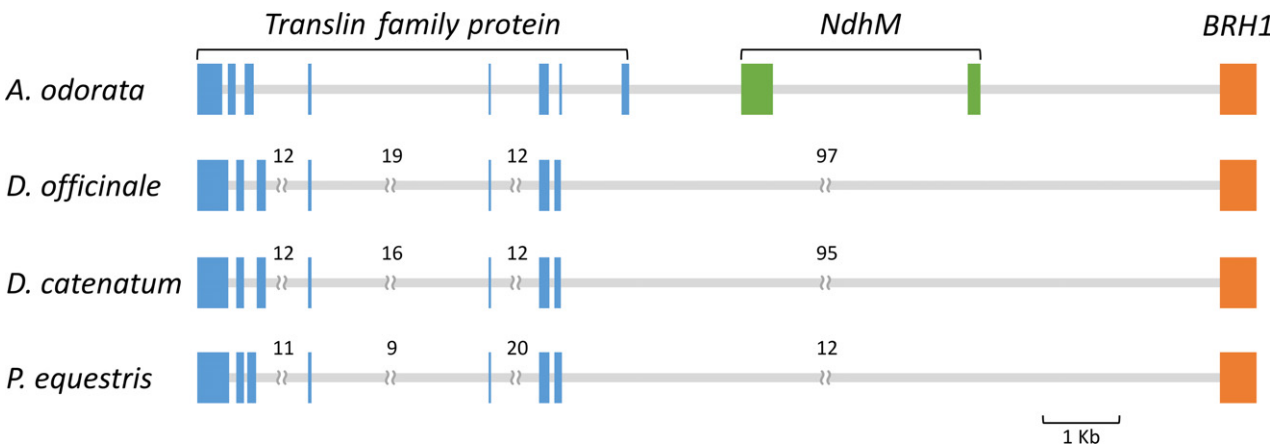


Figure 3. An example of gene-specific loss of nuclear *Ndh* genes in *Phalaenopsis* and *Dendrobium*.
In the *NdhM* genomic region, synteny was noted in *Apostasia odorata*, *Dendrobium officinale*, *Dendrobium catenatum* and *Phalaenopsis equestris*. Gene order remains unchanged as in *A. odorata* except that in *P. equestris*, *D. officinale* and *D. catenatum*, *NdhM* is lost. Box, exon. The number above the grey line indicates the size (kb) of the intron or intergenic region. [Colour figure can be viewed at wileyonlinelibrary.com].

a common evolutionary origin with mitochondrial respiratory complex I (Shikanai, 2016). We next tested whether these lineages missing their cp-*ndh* genes exhibited changes in the expression of (i) nuclear genes functionally related to photosynthesis, such as genes encoding for PSI, PSII, cyt b_6/f, F-type ATPase and LHCs, and (ii) nuclear genes encoding for mitochondrial respiratory complex I, although these genes are present in the genome (Figure 2, Genome panel, and Data S4, S5). Transcriptome data from five cp-*ndh*-complete orchids and four cp-*ndh*-free orchids were retrieved for comparison. As shown in Figure 2 (Transcriptomes panel) and Data S1, in the photoautotrophic orchids, no matter whether the cp-*ndh* gene set was intact or deleted, the majority of photosynthesis-related genes were expressed, except *Lhcb2* and *Lhcb5*, which were not detected in *D. officinale* and *Vanilla planifolia*, respectively. Most of the genes encoding for mitochondrial respiratory complex I are expressed in all the photoautotrophic orchids. Taken together, these results indicated that in orchids devoid of cp-*ndh* genes, the expression of nuclear genes functionally related to photosynthesis or to mitochondrial complex I is not affected. Similar results were also observed in the cp genomes of other cp-*ndh*-free orchids (Chang *et al.*, 2006; Wu *et al.*, 2010).

NDH-related genes in the cp and nuclear genomes may also have been concomitantly lost in heterotrophic plant lineages

In many heterotrophic plants for which the cp genome has been sequenced, the cp-*ndh* genes are absent (Wicke *et al.*, 2016). We were therefore interested to know whether in heterotrophic plants the nuclear NDH-related genes are concomitantly lost in the same way as in orchids that are missing cp-*ndh* genes. Since some photosynthesis-related protein complexes such as PSI, PSII, cyt b_6/f and F-type ATPase are also encoded by the genes located in both the nuclear and cp genomes (Green, 2011), it was of interest to test whether the nuclear and cp genes encoding for these four photosynthesis-related protein complexes are also lost in a correlated fashion. As no sequenced nuclear genomes are available for heterotrophic plants, transcriptome data of the nucleus and the corresponding cp transcriptomes for the same heterotrophic plant species were retrieved from the 1000 Plants Project (1KP) (Matasci *et al.*, 2014) and analyzed instead. Among the nine heterotrophic plants analyzed, the whole set of cp-*ndh* genes were not expressed in six of the selected heterotrophs (Figure 4). Further, transcripts of the majority of nuclear NDH-related genes were not detected in these nine heterotrophic plants either (Figure 4). Notably, nuclear *NdhS*, which was shown to be retained in photoautotrophic cp-*ndh*-free orchids (Figure 1c) and pines (Figure S3), was not expressed in any of the heterotrophic plants analyzed except *Phoradendron serotinum* (Figure 4). The patterns of the other photosynthetic genes actively expressed in the cp (Data S6) and the nucleus (Figure 2) are rather diverse across various species. Most of the photosynthetic genes in some hemiparasitic heterotrophic plants, such as *Cuscuta pentagona* and *P. serotinum*, were expressed in both the cp and the nucleus. Both *Cuscuta* and *Phoradendron* are green and may be photosynthetically active for at least a portion of their life cycle (Marshall *et al.*, 1993; Revill *et al.*, 2005). While in some others, such as the holoparasites *Balanophora fungosa* and *Pilostyles thurben*, almost none were expressed. These data suggest that in all of the heterotrophic plants analyzed in this study, NDH-related genes in the cp and nuclear genomes may have been lost concomitantly in a similar manner to that observed in photoautotrophic cp-*ndh*-free orchids. In hemiparasites, the expression of both cp and nuclear NDH-related genes was

co-downregulated to undetectable levels. In addition, transcriptome data also suggested that, in some heterotrophic plants, photosynthetic genes in the cp and nuclear genome may have also been lost concomitantly (Figure 2; Data S6). These results are congruent with the model of plastid genome evolution previously proposed (Naumann *et al.*, 2016; Wicke *et al.*, 2016). The model predicts that cp-*ndh* genes should be lost from the plastome in the first phase of transition from autotroph to full heterotroph, accompanied by losses of the photosynthetic genes and some other genetic features. Our work shows that nuclear gene loss may also influence the evolution of heterotrophy.

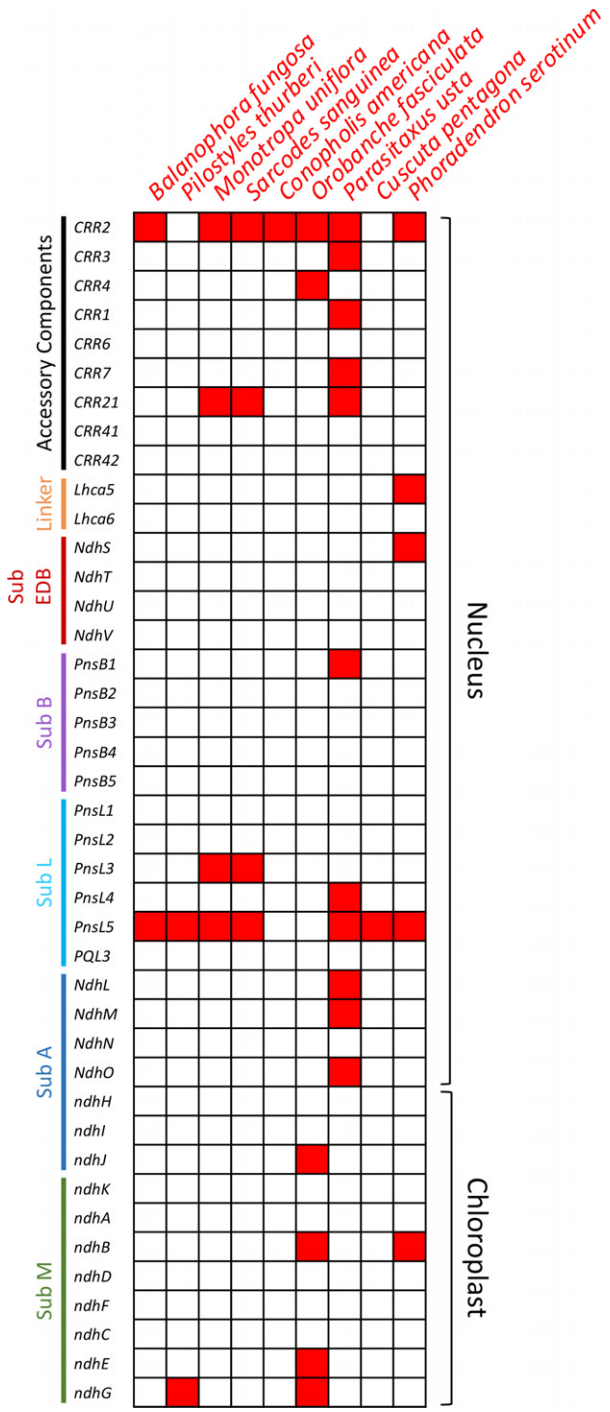


Figure 4. Transcriptome of chloroplast and nucleus encoded NDH-related genes in selected heterotrophic plants.
The amino acid sequences of Arabidopsis in the KEGG photosynthesis pathways were used as queries for tBLASTN searches against public plant transcriptome databases (E-value $< 1 \times 10^{-4}$) (Matasci *et al.*, 2014). The red color-filled box indicates that the gene is actively transcribed. Detailed information is shown in Data S6. [Colour figure can be viewed at wileyonlinelibrary.com].

Some cp-*ndh* genes in Orchidaceae are likely to be under relaxed purifying selection

In order to test for changes in selective constraint on NDH complex genes within the Orchidaceae, we first constructed the phylogenetic relationships of the 11 cp-*ndh* genes in our selected species of monocots. The relationships based on Bayesian inferences were similar to APG IV (The Angiosperm Phylogeny Group 2016) (Figure 5). The posterior probability for all branches is greater than 0.97, indicating a robust phylogenetic tree. The only topological difference from APG IV was the relationship within the commelinids. Our result strongly supports the relationships [[Arecales [Commelinales + Zingiberales]] [Dasypogonaceae + Poales]], which are incongruent with the analyses using full cp genomes (Barrett *et al.*, 2013, 2016). The phylogenetic incongruence may be caused by the strong heterogeneity of cp substitution rates (Barrett *et al.*, 2016) and/or sparse taxon sampling (Fitch and Bruschi, 1987; Venditti *et al.*, 2006).

To test whether the NDH complex evolved in Orchidaceae has a different selective regime from in the non-orchid monocots, we used the maximum likelihood approach to obtain the non-synonymous (dN)/synonymous (dS) ratio (ω) of substitution rate relative to other monocots (Table 1). Synonymous substitution is assumed to be neutral, the ω value can be regarded as an indicator of the direction and magnitude of natural selection. A ratio of less than one implies purifying selection, which means that most amino acid changes are deleterious, and a ratio of one indicates neutral selection. Only genes for which all taxa had complete open reading frames were included in the analyses. Unfortunately, except for commelinids and some orchids, the sequences of nuclear NDH-related genes for most monocots were not available in our current resources, so only 11 cp-*ndh* genes were analyzed. Various different branch models of selective change were compared with the null one-ratio model (Table 1). Though the difference in ω value between orchids and non-orchid monocots is very small, the likelihood ratio test showed significantly relaxed purifying selection of *ndhA*, *ndhD*, *ndhF* and *ndhH* in Orchidaceae relative to non-orchid monocots. Of these four genes, all the multiple-ratio branch models fit our data significantly better than the one-ratio model (M0), with the exception that the five-ratio

branch model was not favored for *ndhA*. In addition, neither the three-ratio nor the five-ratio branch models were favored over the two-ratio branch model for any of the cp-*ndh* genes when we specified that the whole orchid lineage – including *Apostasia* – had a different ω from the non-orchid monocots (Table S1). However, ω in the orchid lineage of *ndhA*, *ndhD*, *ndhF* and *ndhH* genes have not yet approached the expected equilibrium of 1.0 under neutral evolution. None of the multiple-ratio models for the other seven cp-*ndh* genes and *psbA* had a significantly better fit than M0 in likelihood ratio tests. These results showed that the functional constraint of cp-*ndh* genes has been decreased in Orchidaceae, although not all of them are statistically significant. In other words, if we assume that the NDH-related genes are under the same conditions, the need for the NDH complex is decreased in orchids and it may be dispensable. However, it remains unclear why the functional constraint is decreased in orchids and why the NDH-related genes were lost so rapidly that only very few species have vestigial pseudogenes.

Unexpectedly, the three-ratio model showed a significantly higher ω in core Orchidaceae (i.e. excluding the subfamily Apostasioideae) in *rbcL*, though two- and five-ratio models were not favored over M0 for *rbcL*. This unexpected result may be a consequence of the various photosynthetic types and diverse habits of core Orchidaceae species.

Losses of cp-*ndh* genes occurred independently in various lineages in land plants

The results above show that plants with cp-*ndh* deletions also exhibit losses of their nuclear NDH-related genes. These findings motivated us to investigate how many plants, either photoautotrophic or heterotrophic, are cp-*ndh*-free, and whether loss of cp-*ndh* genes has occurred more often in certain evolutionary lineages. There are 11 cp-*ndh* genes in cp genomes. Although these cp-*ndh* genes have been shown to be lost or pseudogenized in parallel and gradually (Barrett *et al.*, 2014; Kim *et al.*, 2015), in the pool of our analyzed cp-*ndh*-free plants, no more than six cp-*ndh* genes were intact, functional and retained in the cp genome of any single species (Figure 6). If loss of the functional NDH complex is considered to have occurred by impairing most cp-*ndh* genes in a single evolutionary event, named cp-*ndh* deletion, 47 out of 660 photoautotrophic plants with available cp genome data have

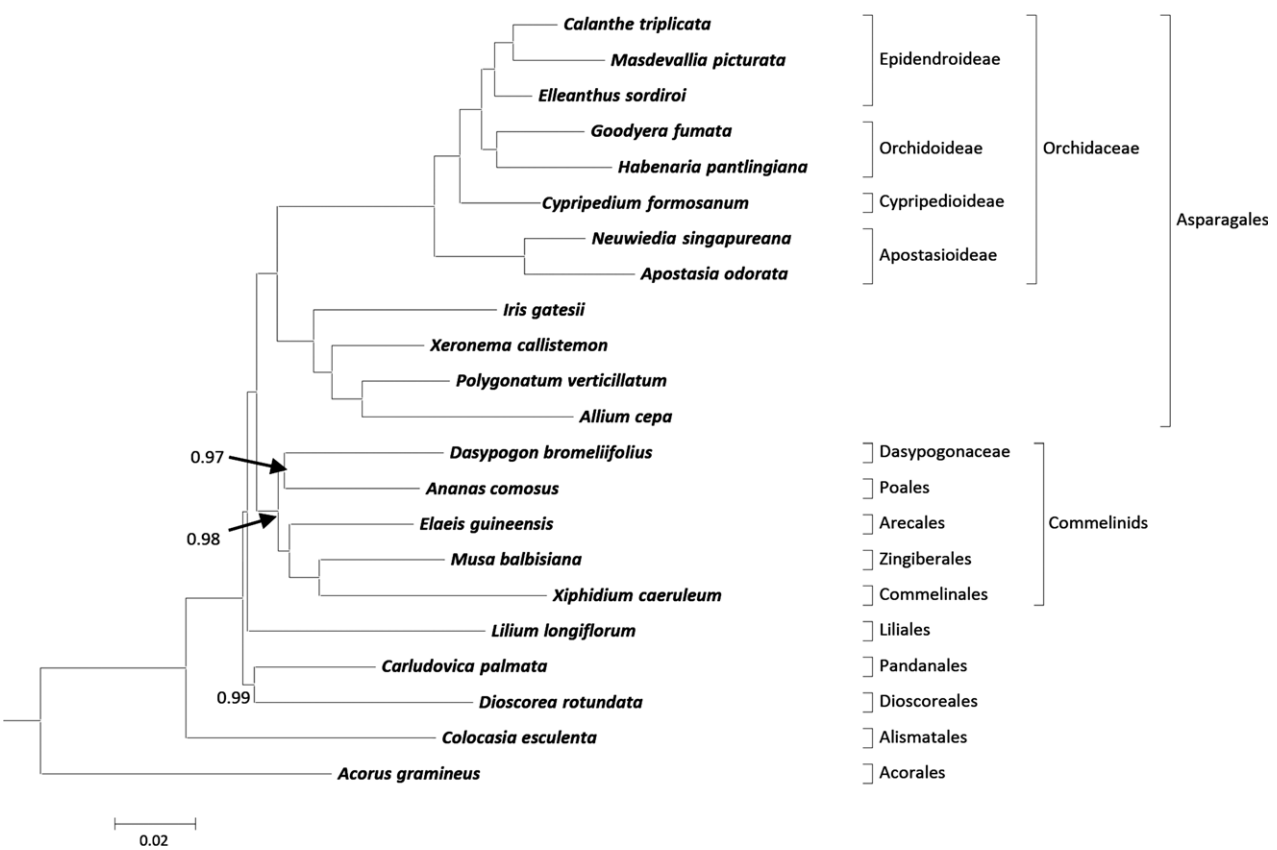


Figure 5. Phylogenetic tree of selected monocots constructed by the Bayesian inference method, using 11 chloroplast *ndh* genes.
Numbers at the nodes indicate posterior probability values. The posterior probability of nodes without a number is 1.00. Accession numbers of the selected species are listed in Table S3.

cp-*ndh* deletion (Figure 6a, Data S7; Daniell *et al.*, 2016), and loss of cp-*ndh* genes occurred independently at least eleven times among photoautotrophic land plants (Figure 6a, Table S2), twice in gymnosperms, five times in monocot angiosperms and four times in eudicot angiosperms. Some of the loss events may be ancient, such as within the ancestral lineage of the Pinaceae and Gnetales, which may share a single early cp-*ndh* deletion event (Braukmann *et al.*, 2009). In contrast, some loss events probably occurred rather recently, for example within the genus *Erodium* (Blazier *et al.*, 2011).

Partially and fully heterotrophic plants can be classified into several groups according to photosynthetic ability and/or the kingdom of the host, i.e. plants or fungi. Our analyses showed that all the 38 heterotrophic plants analyzed, including the liverwort *Aneura mirabilis* and 37 angiosperms including basal angiosperms, monocots and eudicots, exhibited cp-*ndh* deletion (Figure 6b). These results indicate that cp-*ndh*-free plants are not evolutionarily related and consist of a wide variety of taxa occurring in almost every major group of land plants.

DISCUSSION

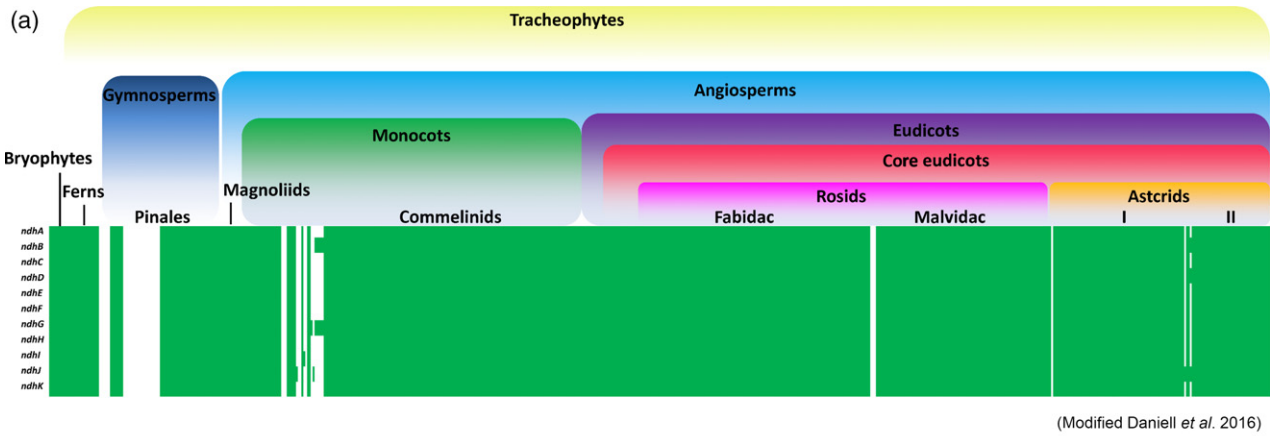
Loss of cp-*ndh* genes has been reported in a large number of plant species, including all heterotrophic and some photoautotrophic plants (Wakasugi *et al.*, 1994; Chang *et al.*, 2006; Wu *et al.*, 2010; Lin *et al.*, 2015; Bellot and Renner, 2016). However, whether the lost cp-*ndh* genes were transferred to the nuclear genome and whether the NDH-related genes in the nuclear genome were also lost in the cp-*ndh*-free plants is still unknown. In this study, through analyses of transcriptomes/genomes of related orchid species, we showed that the lost cp-*ndh* genes are not transferred to the nucleus and that losses of the nuclear NDH-related genes and cp-*ndh* genes are in tandem. Our data also showed that genes related to the NDH complex are the only genes for which concomitant losses have been observed for multiple genes in two distinct genomes that encode genes for the same protein complex. Analyses of the gene orders adjacent to NDH-related genes revealed that the losses of NDH-related genes in both the cp and nuclear genomes are locus-specific, that is, flanking genes unrelated to the NDH complex are still retained and are functional. The visual systems in animals such as the naked mole rat (Kim *et al.*, 2011), blind mole rat (Fang *et al.*, 2014) and cavefish (Yang *et al.*, 2016) are degenerated by mutations of multiple genes. However, in contrast to NDH complex-free orchids, in these animals, premature stop codons are created to the genes encoding for the visual systems by point mutation, insertion or deletion of some bases. Although these mutated genes relate to similar physiological function and development of the same organ, unlike the case of losses of NDH-related genes in orchids, the proteins encoded by these genes are not subunits of the same protein complex, and the genes encoding for other subunits of the complex are still retained and may have additional functions. NDH-related genes (including cp-*ndh* genes) interact functionally, so we expect co-evolution of these genes whether they are encoded by the nuclear or cp genomes. Correlated loss of these genes seems to occur over short evolutionary time periods, but it remains unknown whether concerted loss is adaptive or

Table 1 Estimated ω values and results of maximum likelihood tests evaluating different branch models for selectional changes of chloroplast (cp)-*ndh* genes and photosynthesis genes between Orchidaceae and non-orchid monocots

	One-ratio model (M0)	Two-ratio model			Three-ratio model				Five-ratio model					
Gene	ω_{all}	$\omega_{non\text{-}or}$	ω_{orchid}	P^a	$\omega_{non\text{-}or}$	ω_{Apo}	$\omega_{non\text{-}Apo}$	P^a	$\omega_{non\text{-}or}$	ω_{Apo}	ω_{Cyp}	ω_{Orc}	ω_{Epi}	P^a
ndhA	0.158	0.140	0.211	**0.005**	0.141	0.198	0.236	**0.013**	0.142	0.198	0.221	0.227	0.250	0.068
ndhB	0.301	0.300	0.304	0.964	0.295	0.243	0.364	0.801	0.295	0.244	0.274	0.407	0.368	0.959
ndhC	0.162	0.166	0.152	0.796	0.158	0.282	0.130	0.496	0.158	0.282	0.000	0.188	0.078	0.533
ndhD	0.172	0.146	0.255	**<0.001**	0.148	0.229	0.287	**<0.001**	0.148	0.228	0.404	0.224	0.296	**<0.001**
ndhE	0.130	0.122	0.151	0.472	0.120	0.235	0.144	0.441	0.120	0.236	0.052	0.363	0.104	0.105
ndhF	0.226	0.206	0.296	**<0.001**	0.211	0.302	0.386	**0.007**	0.211	0.304	0.228	0.351	0.273	**0.018**
ndhG	0.225	0.216	0.247	0.538	0.216	0.306	0.224	0.591	0.216	0.306	0.274	0.319	0.115	0.402
ndhH	0.085	0.073	0.125	**0.002**	0.074	0.145	0.122	**0.009**	0.074	0.145	0.125	0.142	0.098	**0.039**
ndhI	0.100	0.097	0.107	0.706	0.100	0.084	0.111	0.875	0.100	0.084	0.125	0.078	0.136	0.911
ndhJ	0.157	0.157	0.156	0.985	0.149	0.171	0.190	0.762	0.149	0.172	0.158	0.175	0.248	0.918
ndhK	0.132	0.132	0.135	0.926	0.131	0.104	0.165	0.650	0.131	0.103	0.267	0.245	0.078	0.388
psbA	0.024	0.023	0.027	0.672	0.021	0.026	0.035	0.551	0.021	0.026	0.015	0.063	0.029	0.529
rbcL	0.085	0.079	0.109	0.102	0.077	0.089	0.142	**0.025**	0.077	0.089	0.082	0.178	0.144	0.059

Abbreviations for ω-ratios: ω_{all}, all monocots; $\omega_{non\text{-}or}$, non-orchid monocots; ω_{orchid}, all orchids; ω_{Apo}, Apostasioideae; $\omega_{non\text{-}Apo}$, non-Apostasioideae orchids; ω_{Cyp}, Cypripedioideae; ω_{Orc}, Orchidoideae; ω_{Epi}, Epidendroideae.

[a]*P*-values are derived from likelihood-ratio tests (LRTs) versus the one-ratio model (see Experimental Procedures for details). *P*-values < 0.05 are given in bold type.



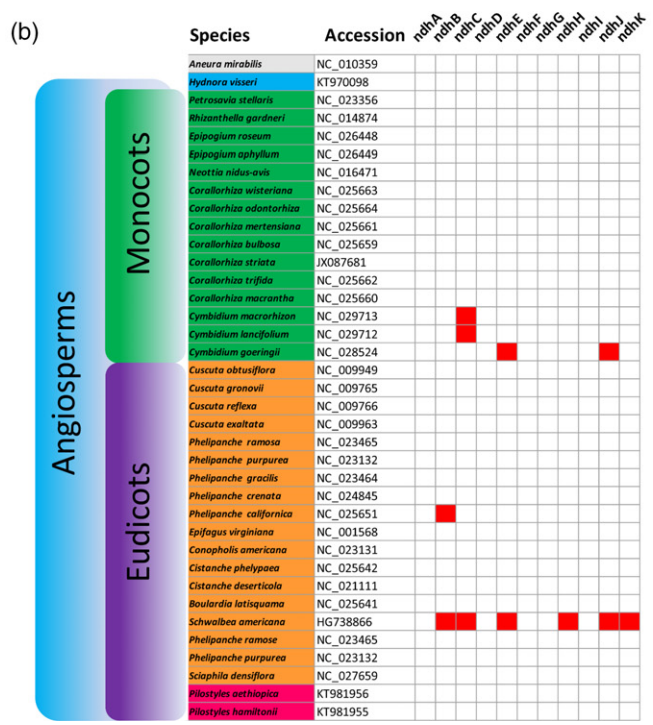


Figure 6. Loss of the functional chloroplast (cp)-*ndh* genes in the cp genome of photoautotrophic and heterotrophic plants.
(a) Loss of cp-*ndh* genes in the cp genome of photoautotrophic plants. Forty-seven out of 660 plants are cp-*ndh*-free. Only cp-*ndh* genes are shown. Presence and absence of cp-*ndh* genes are shown in green and white boxes, respectively. Detailed information is shown in Data S7.
(b) Deletion of cp-*ndh* genes in the cp genome of heterotrophic plants. No more than six cp-*ndh* potentially functional *ndh* genes are detected in the heterotrophic plants analyzed. Presence and absence of genes are shown in black and white boxes, respectively. [Colour figure can be viewed at wileyonlinelibrary.com].

neutral. Although our finding about cp-*ndh* genes lost in heterotroph plastomes supports the prediction of the model proposed by Wicke *et al.* (2016), it is unclear why cp-*ndh* genes and nuclear NDH-related genes have been lost in some autotrophic plants such as Pinaceae.

Our analyses of 660 photoautotrophic plants showed that about one-fifteenth have lost most of the cp-*ndh* genes from their chloroplast genomes (Figure 6a, Data S7). Multiple independent cp-*ndh* deletions have been revealed in various leafy photosynthetic lineages over the past few decades (Table S2), which suggests that the function of the NDH complex may not be indispensable for photosynthesis, or the significance of the NDH complex in photosynthesis may be restricted to certain plant species. PSI-CET generates a trans-thylakoid proton gradient and functions in ATP production without producing NADPH. In contrast to C3 photosynthesis, C4 photosynthesis requires more ATP to efficiently fix atmospheric CO_2, and NDH complex-mediated CET is shown to be crucial for this step (Takabayashi *et al.*, 2005; Ishikawa *et al.*, 2016a,b). Mutants deficient in the NDH complex have also been reported to be sensitive to various stress conditions, such as high light

(Endo *et al.*, 1999), water deficiency (Horvath *et al.*, 2000) and low temperature (Yamori *et al.,* 2011). The NDH complex is also shown to facilitate chlororespiration, the non-photochemical reduction and oxidation of plastoquinones (Peltier *et al.*, 2016; Yamori and Shikanai, 2016). Since the studies of mutants lacking the NDH complex are mostly performed under light-grown conditions, in which NDH-mediated PSI-CET may operate, it is uncertain whether chlororespiration mediated by the NDH complex plays an important or essential role.

Heterotrophic plants have different trophic strategies from photoautotrophic plants. They completely or partly acquire reduced carbon from hosts, such as plants and fungi, instead of by photosynthetic CO_2 fixation. Some hemiparasitic heterotrophic plants, such as *Cuscuta* and *Phoradendron*, are green and may be photosynthetically active for at least a portion of their life cycle (Marshall *et al.*, 1993; Revill *et al.*, 2005). However, these plants are considered to possess, at most, a low level of CO_2 fixation (Sherman *et al.*, 1999). Most of the cp-*ndh* genes are absent in all the heterotrophic plants analyzed (Wicke *et al.*, 2016). Our analyses of currently available transcriptome data showed that expression of the cp and the nuclear NDH-related genes is not detected in heterotrophic plants (Figure 4), suggesting that in heterotrophic plants NDH-related genes in the cp and the nuclear genomes may also have been lost concomitantly; furthermore, the loss of NDH-related genes may be related to the transition from photoautotroph to heterotroph or reduction of CO_2 fixation activity. It is interesting that orchid seeds generally contain a minute globular-shape embryo and a very limited food reserve, and, in a natural habitat, most germinating orchid seedlings are heterotrophic and rely on fungi to supply nutrients (Arditti, 1992; Pierik, 1997). Thus, it is also possible that loss of the NDH complex is related to, or necessary for, plants with heterotrophic phase in their life cycle. Maybe the existence of the NDH complex or subunit(s) of the NDH complex has a deleterious effect on the lifestyle of heterotrophs and therefore tends to be eliminated by selection. As suggested by Krause (2012), parasitism may prompt adaptive gene loss. Some orchids still contain the whole set of NDH-related genes (Data S7). The emergence of the heterotrophic stage of these orchids may have occurred very recently, and thus the NDH-related genes are still retained due to insufficient time for selection to eliminate them. Our analyses suggest that in contrast to non-orchid monocots, NDH complex genes in Orchidaceae may be evolving under relaxed selective constraints and may ultimately be lost (Table 1). Further investigations into the correlation between CO_2 fixation, growth habit, type of trophic strategy and NDH activity would be helpful to elucidate the mechanism and function of the NDH complex in plants. In addition to some photoautotrophic orchids, many non-orchid photoautotrophic plants also exhibit cp-*ndh* loss (Figure 6a). It is possible that an unidentified heterotrophic stage may occur in the life cycle of these plants; however, it is not excluded that some unidentified special needs evolved independently of the transition to the trophic strategy.

EXPERIMENTAL PROCEDURES

Plant material and genome sequencing

Plant tissues of *A. odorata* were collected in Yunnan, China. The voucher number is 'Jin X.H. 11267 (PE)'. Genomic DNA for genome sequencing was extracted using a Plant DNA Purification Kit (DP320, Tiangen, http://www.tiangen.com/en/). Paired-end genomic DNA was sequenced using an Illumina MiSeq platform (https://www.illumina.com/; 250-bp paired-end reads, 586-bp insert with adaptor). Total RNA was extracted using a Maxwell 16 LEV Plant RNA Kit (AS1430, Promega, http://www.promega.com/). The Illumina Hiseq 2000 sequencing platform was used for transcriptome sequencing. Paired-end libraries of RNA were constructed by the TruSeq library with 2 × 90 bp and the insert size was 270 bp.

Draft genome assembly

A draft assembly genome of *A. odorata* was generated using SOAPdenovo and ABySS. We used SOAPdenovo to assemble paired-end sequencing reads (2 × 250 bp, 83 million reads) to get the contigs from DNA sequence reads first, then combined mate-pair reads from a 2-kb genomic DNA library (Kmer = 41) to generate scaffolds (26 million reads). Then ABySS was used to assemble five different transcriptome sequencing reads and combine them with DNA contigs. We also used SSPACE to fill the gaps. The genome sequence is 330 Mbp and the raw read coverage 69×. The raw reads are deposited in the NCBI database (BioProject ID PRJNA335866).

Homolog identification

The Arabidopsis peptide sequences of NDH-, photosynthesis- and mitochondrial respiration-related genes were used as queries to identify the orthologs. Genome and transcriptome blast libraries were constructed using NCBI blast+ blast-2.2.29+. We aligned the reference gene to genomes and transcriptomes using tBLASTN (E-value $< 1 \times 10^{-4}$) and extracted the candidate sequences. These candidate sequences were blasted back to Arabidopsis using blastx to check that the top hit of the sequence was the corresponding reference gene (Ruhlman *et al.*, 2015).

Sequence alignment and phylogenetic construction

A phylogenetic tree was constructed based on 11 cp-*ndh* genes (*ndhA–K*) as the input tree for relaxed purifying selection tests. The open reading frames of 11 cp-*ndh* sequences were downloaded from GenBank. The nucleotide sequences were converted to amino acid sequences and aligned using ClustalW with default settings. The alignment was then converted back to nucleotide sequences for subsequent analyses. The stop codons were removed and the indels were deleted as a triplet codon to maintain the reading frames. The trimmed alignments were concatenated into a single alignment data matrix. The general time-reversible substitution model with a gamma and invariant site heterogeneity model was selected by MEGA v.6 (Tamura *et al.*, 2013) for Bayesian analyses. A phylogenetic tree was constructed by Bayesian inferences using MrBayes v.3.2.6 (Ronquist

et al., 2012). Four chains, three heated and one cold, were performed for a million generations, sampling every 1000 generations. The first 25% of sampled trees were discarded as 'burn-in'. The remaining trees were then used to calculate Bayesian inference posterior probability.

Detecting the genetic signatures of relaxed purifying selection

We used the maximum likelihood approach to estimate the ratio (ω) of the non-synonymous (dN) to synonymous (dS) substitution rate relative to other monocots with codon-based analysis (codeml) in the PAML v.4.8a package (Yang, 2007). The taxonomy and evolutionary placement of the selected monocot species are listed in Table S3. We selected one species from every order in the monocots with the exception that twelve species were selected from Asparagales and eight belonged to the family Orchidaceae, and we did not have access to data for the Petrosaviales.

Molecular evolutionary analyses were performed using contrasting branch models to test whether the selective regime was changed within specified focal lineages. First, the likelihood of an optimized two-ratio model (branch model), with a foreground ω for Orchidaceae and a background ω for non-orchid monocots, was compared with the likelihood of the null one-ratio model (M0). To test whether the *Apostasia* branch had a different ω value relative to other orchid lineages, a branch model with three-ratios was compared with the null (M0) model. We also compared the M0 model with a model in which each orchid subfamily was allowed to have a separate ω (branch model with five ratios). The likelihood ratio test statistic, $2\Delta \ln L$, was compared with the chi-square (χ^2) distribution, with a degree of freedom equal to the difference in the number of parameters for the models, in order to evaluate the fit of the data to alternative branch models.

ACKNOWLEDGEMENTS

We thank Miranda Loney for English editing. This work was supported by Agricultural Biotechnology Research Center, Academia Sinica of Taiwan (to CSL and MCS) and the Ministry of Science and Technology, Taiwan (grant MOST 104-2311-B-006-001 to CCC).

CONFLICT OF INTEREST

The authors declare no conflict of interest.

SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article.

Figure S1. Most cp-*ndh* genes lost from the chloroplast genome were not retained in the nuclear genomes in *Phalaenopsis* and *Dendrobium*.

Figure S2. Most cp-*ndh* genes lost from the chloroplast genome were not retained in the nuclear genomes in gymnosperms *Pinus taeda* and *Picea abies*.

Figure S3. Genes encoding for the NDH complex in the chloroplast and the nuclear genomes are concomitantly lost in the gymnosperms *Pinus taeda* and *Picea abies*.

Table S1. The *P*-values of likelihood ratio test for the two-ratio model versus the three-ratio or five-ratio models for four *ndhA*, *ndhD*, *ndhF* and *ndhH* genes.

Table S2. Overview of numbers of independent occurrences of cp-*ndh* deletion in photoautotrophic land plants.

Table S3. List of plastid genomes of monocot species used in the selection test in Table 1.

Data S1. Detailed data from the transcriptomic/genomic analyses of photosynthesis-related genes in some photoautotrophic and heterotrophic plants (excluding NDH-related genes) shown in Figure 2.

Data S2. The order of the orthologous genes of the adjacent genes of *Apostasia* NDH-related genes in *Dendrobium officinale*, *Dendrobium catenatum* and *Phalaenopsis equestris*.

Data S3. The order of annotated genes adjacent to the 30 nuclear NDH-related genes in the genomes of five selected angiosperms.

Data S4. Transcriptomic/genomic analyses of respiration-related genes in selected photoautotrophic and heterotrophic plants. Detailed information is shown in Data S5.

Data S5. Detailed data from the transcriptomic/genomic analyses of respiration-related genes in selected photoautotrophic and heterotrophic plants shown in Data S4.

Data S6. Transcriptome of nuclear NDH-related genes, cp-*ndh* genes and other genes in the chloroplast genome of selected heterotrophic plants.

Data S7. Detailed data about loss of the functional cp-*ndh* genes in the chloroplast genome of photoautotrophic plants.

REFERENCES

Arditti, J. (1992) *Fundamentals of Orchid Biology*. New York: John Wiley and Sons, Inc..

Barrett, C.F., Davis, J.I., Leebens-Mack, J., Conran, J.G. and Stevenson, D.W. (2013) Plastid genomes and deep relationships among the commelinid monocot angiosperms. *Cladistics*, **29**, 65–87.

Barrett, C.F., Freudenstein, J.V., Li, J., Mayfield-Jones, D.R., Perez, L., Pires, J.C. and Santos, C. (2014) Investigating the path of plastid genome degradation in an early-transitional clade of heterotrophic orchids, and implications for heterotrophic angiosperms. *Mol. Biol. Evol.* **31**, 3095–3112.

Barrett, C.F., Baker, W.J., Comer, J.R. *et al.* (2016) Plastid genomes reveal support for deep phylogenetic relationships and extensive rate variation among palms and other commelinid monocots. *New Phytol.* **209**, 855–870.

Bellot, S. and Renner, S.S. (2016) The plastomes of two species in the endoparasite genus Pilostyles (Apodanthaceae) each retain just five or six possibly functional genes. *Genome Biol. Evol.* **8**, 189–201.

Blazier, J., Guisinger, M.M. and Jansen, R.K. (2011) Recent loss of plastid-encoded *ndh* genes within *Erodium* (Geraniaceae). *Plant Mol. Biol.* **76**, 263–272.

Braukmann, T.W.A., Kuzmina, M. and Stefanovic, S. (2009) Loss of all plastid *ndh* genes in Gnetales and conifers: extent and evolutionary significance for the seed plant phylogeny. *Curr. Genet.* **55**, 323–337.

Braun, H.P., Binder, S., Brennicke, A *et al.* (2014) The life of plant mitochondrial complex I. *Mitochondrion*, **19** Pt B, 295–313.

Cai, J., Liu, X., Vanneste, K. *et al.* (2015) The genome sequence of the orchid *Phalaenopsis equestris*. *Nat. Genet.* **47**, 304–304.

Chang, C.C., Lin, H.C., Lin, I.P. *et al.* (2006) The chloroplast genome of *Phalaenopsis aphrodite* (Orchidaceae): Comparative analysis of evolutionary rate with that of grasses and its phylogenetic implications. *Mol. Biol. Evol.* **23**, 279–291.

Daniell, H., Lin, C.S., Yu, M. and Chang, W.J. (2016) Chloroplast genomes: diversity, evolution, and applications in genetic engineering. *Genome Biol.* **17**, 134.

Endo, T., Shikanai, T., Takabayashi, A., Asada, K. and Sato, F. (1999) The role of chloroplastic NAD(P)H dehydrogenase in photoprotection. *FEBS Lett.* **457**, 5–8.

Fang, X.D., Nevo, E., Han, L.J. *et al.* (2014) Genome-wide adaptive complexes to underground stresses in blind mole rats Spalax. *Nat. Commun.* **6**, 8051.

Fitch, W.M. and Bruschi, M. (1987) The Evolution of prokaryotic ferredoxins - with 10 a general-method correcting for unobserved substitutions in less branched 11 lineages. *Mol. Biol. Evol.* **4**, 381–394.

Green, B.R. (2011) Chloroplast genomes of photosynthetic eukaryotes. *Plant J.* **66**, 34–44.

Horvath, E.M., Peter, S.O., Joet, T., Rumeau, D., Cournac, L., Horvath, G.V., Kavanagh, T.A., Schafer, C., Peltier, G. and Medgyesy, P. (2000) Targeted inactivation of the plastid *ndhB* gene in tobacco results in an enhanced sensitivity of photosynthesis to moderate stomatal closure. *Plant Physiol.* **123**, 1337–1350.

Ishikawa, N., Takabayashi, A., Noguchi, K., Tazoe, Y., Yamamoto, H., von Caemmerer, S., Sato, F. and Endo, T. (2016a) NDH-mediated cyclic electron flow around photosystem I is crucial for C4 photosynthesis. *Plant Cell Physiol.* **57**, 2020–2028.

Ishikawa, N., Takabayashi, A., Sato, F. and Endo, T. (2016b) Accumulation of the components of cyclic electron flow around photosystem I in C4 plants, with respect to the requirements for ATP. *Photosynth. Res.* **129**, 261–277.

Jansen, R.K., Cai, Z., Raubeson, L.A. *et al.* (2007) Analysis of 81 genes from 64 plastid genomes resolves relationships in angiosperms and identifies genome-scale evolutionary patterns. *Proc. Natl. Acad. Sci. U. S. A.* **104**, 19369–19374.

Jersakova, J., Travnicek, P., Kubatova, B. *et al.* (2013) Genome size variation in Orchidaceae subfamily Apostasioideae: filling the phylogenetic gap. *Bot. J. Linnean Soc.* **172**, 95–105.

Jin, X.H. and Huang, L.Q. (2015) Proposal to conserve the name *Dendrobium officinale* against *D. stricklandianum*, *D. tosaense*, and *D. pere-fauriei* (Orchidaceae). *Taxon*, **64**, 385–386.

Kim, E.B., Fang, X.D., Fushan, A.A. *et al.* (2011) Genome sequencing reveals insights into physiology and longevity of the naked mole rat. *Nature*, **479**, 223–227.

Kim, H.T., Kim, J.S., Moore, M.J., Neubig, K.M., Williams, N.H., Whitten, W.M. and Kim, J.H. (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**, e0142215.

Krause, K. (2012) *Plastid Genomes of Parasitic Plants: A Trail of Reductions and Losses*. Berlin Heidelberg: Springer-Verlag.

Lin, C.S., Chen, J.J.W., Huang, Y.T. *et al.* (2015) The location and translocation of *ndh* genes of chloroplast origin in the Orchidaceae family. *Sci. Rep.* **5**, 9040.

Marshall, J.D., Dawson, T.E. and Ehleringer, J.R. (1993) Gender-Related Differences in gas-exchange are not related to host quality in the xylem-tapping mistletoe, *Phoradendron-Juniperinum* (Viscaceae). *Am. J. Bot.* **80**, 641–645.

Martin, W. and Herrmann, R.G. (1998) Gene transfer from organelles to the nucleus: How much, what happens, and why? *Plant Physiol.* **118**, 9–17.

Matasci, N., Hung, L.H., Yan, Z.X. *et al.* (2014) Data access for the 1,000 Plants (1KP) project. *Gigascience*, **3**, 17.

Naumann, J., Der, J.P., Wafula, E.K. *et al.* (2016) Detecting and characterizing the highly divergent plastid genome of the nonphotosynthetic parasitic plant *Hydnora visseri* (Hydnoraceae). *Genome Biol. Evol.* **8**, 345–363.

Neale, D.B., Wegrzyn, J.L., Stevens, K.A. *et al.* (2014) Decoding the massive genome of loblolly pine using haploid DNA and novel assembly strategies. *Genome Biol.* **15**, R59.

Nystedt, B., Street, N.R., Wetterbom, A. *et al.* (2013) The Norway spruce genome sequence and conifer genome evolution. *Nature*, **497**, 579–584.

Peltier, G., Aro, E.M. and Shikanai, T. (2016) NDH-1 and NDH-2 plastoquinone reductases in oxygenic photosynthesis. *Annual Rev. Plant Biol.* **67**, 55–80.

Peng, L.W. and Shikanai, T. (2011) Supercomplex formation with Photosystem I is required for the stabilization of the chloroplast NADH dehydrogenase-like complex in *Arabidopsis*. *Plant Physiol.* **155**, 1629–1639.

Peng, L., Shimizu, H. and Shikanai, T. (2008) The chloroplast NAD(P)H dehydrogenase complex interacts with photosystem I in *Arabidopsis*. *J. Biol. Chem.* **283**, 34873–34879.

Peng, L., Fukao, Y., Fujiwara, M., Takami, T. and Shikanai, T. (2009) Efficient operation of NAD(P)H dehydrogenase requires supercomplex formation with photosystem I via minor LHCI in *Arabidopsis*. *Plant Cell*, **21**, 3623–3640.

Pierik, R.L.M. (1997) *In Vitro Culture of Higher Plants Dordrecht*. Boston: Kluwer Academic Publishers.

Ranade, S.S., Garcia-Gil, M.R. and Rossello, J.A. (2016) Non-functional plastid *ndh* gene fragments are present in the nuclear genome of Norway spruce (*Picea abies* L. Karsch): insights from in silico analysis of nuclear and organellar genomes. *Mol. Genet. Genomics*, **291**, 935–941.

Revill, M.J.W., Stanley, S. and Hibberd, J.M. (2005) Plastid genome structure and loss of photosynthetic ability in the parasitic genus *Cuscuta*. *J. Exp. Bot.* **56**, 2477–2486.

Ronquist, F., Teslenko, M., van der Mark, P., Ayres, D.L., Darling, A., Hohna, S., Larget, B., Liu, L., Suchard, M.A. and Huelsenbeck, J.P. (2012) MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. *Syst. Biol.* **61**, 539–542.

Ruhlman, T.A., Chang, W.J., Chen, J.J.W. *et al.* (2015) *NDH* expression marks major transitions in plant evolution and reveals coordinate intracellular gene loss. *BMC Plant Biol.* **15**, 100.

Sherman, T.D., Pettigrew, W.T. and Vaughn, K.C. (1999) Structural and immunological characterization of the *Cuscuta pentagona* L-chloroplast. *Plant Cell Physiol.* **40**, 592–603.

Shikanai, T. (2016) Chloroplast NDH: A different enzyme with a structure similar to that of respiratory NADH dehydrogenase. *Biochim. Biophys. Acta*, **1857**, 1015–1022.

Simpson, C.L. and Stern, D.B. (2002) The treasure trove of algal chloroplast genomes. Surprises in architecture and gene content, and their functional implications. *Plant Physiol.* **129**, 957–966.

Takabayashi, A., Kishine, M., Asada, K., Endo, T. and Sato, F. (2005) Differential use of two cyclic electron flows around photosystem I for driving CO_2-concentration mechanism in C4 photosynthesis. *Proc. Natl. Acad. Sci. U. S. A.* **102**, 16898–16903.

Tamura, K., Stecher, G., Peterson, D., Filipski, A. and Kumar, S. (2013) MEGA6: Molecular evolutionary genetics analysis version 6.0. *Mol. Biol. Evol.* **30**, 2725–2729.

The Angiosperm Phylogeny Group (2016) An update of the Angiosperm Phylogeny Group classification for the orders and families of flowering plants: APG IV. *Bot. J. Linnean Soc.* **181**, 1–20.

Tiller, N. and Bock, R. (2014) The translational apparatus of plastids and its role in plant development. *Mol. Plant*, **7**, 1105–1120.

Timmis, J.N., Ayliffe, M.A., Huang, C.Y. and Martin, W. (2004) Endosymbiotic gene transfer: Organelle genomes forge eukaryotic chromosomes. *Nat. Rev. Genet.* **5**, 123–135.

Venditti, C., Meade, A. and Pagel, M. (2006) Detecting the node-density artifact in phylogeny reconstruction. *Syst. Biol.* **55**, 637–643.

Wakasugi, T., Tsudzuki, J., Ito, S., Nakashima, K., Tsudzuki, T. and Sugiura, M. (1994) Loss of all *ndh* genes as determined by sequencing the entire chloroplast genome of the black pine *Pinus thunbergii*. *Proc. Natl. Acad. Sci. U. S. A.* **91**, 9794–9798.

Wang, D., Wu, Y.W., Shih, A.C.C., Wu, C.S., Wang, Y.N. and Chaw, S.M. (2007) Transfer of chloroplast genomic DNA to mitochondrial genome occurred at least 300 MYA. *Mol. Biol. Evol.* **24**, 2040–2048.

Wang, D., Rousseau-Gueutin, M. and Timmis, J.N. (2012) Plastid sequences contribute to some plant mitochondrial genes. *Mol. Biol. Evol.* **29**, 1707–1711.

Wicke, S., Schneeweiss, G.M., dePamphilis, C.W., Muller, K.F. and Quandt, D. (2011) The evolution of the plastid chromosome in land plants: gene content, gene order, gene function. *Plant Mol. Biol.* **76**, 273–297.

Wicke, S., Muller, K.F., dePamphilis, C.W., Quandt, D., Bellot, S. and Schneeweiss, G.M. (2016) Mechanistic model of evolutionary rate variation en route to a nonphotosynthetic lifestyle in plants. *Proc. Natl. Acad. Sci. U. S. A.* **113**, 9045–9050.

Wu, F.H., Chan, M.T., Liao, D.C., Hsu, C.T., Lee, Y.W., Daniell, H., Duvall, M.R. and Lin, C.S. (2010) Complete chloroplast genome of *Oncidium* Gower Ramsey and evaluation of molecular markers for identification and breeding in Oncidiinae. *BMC Plant Biol.* **10**, 68.

Yamori, W., Sakata, N., Suzuki, Y., Shikanai, T. and Makino, A. (2011) Cyclic electron flow around photosystem I via chloroplast NAD(P)H dehydrogenase (NDH) complex performs a significant physiological role during photosynthesis and plant growth at low temperature in rice. *Plant J.* **68**, 966–976.

Yamori, W. and Shikanai, T. (2016) Physiological functions of cyclic electron transport around photosystem I in sustaining photosynthesis and plant growth. *Annual Rev. Plant Biol.* **67**, 81–106.

Yan, L., Wang, X., Liu, H. *et al.* (2015) The genome of *Dendrobium officinale* illuminates the biology of the important traditional Chinese orchid herb. *Mol. Plant*, **8**, 922–934.

Yang, Z.H. (2007) PAML 4: Phylogenetic analysis by maximum likelihood. *Mol. Biol. Evol.* **24**, 1586–1591.

Yang, J., Chen, X., Bai, J. *et al.* (2016) The Sinocyclocheilus cavefish genome provides insights into cave adaptation. *BMC Biol.* **14**, 1.

Yoon, H.S., Hackett, J.D., Ciniglia, C., Pinto, G. and Bhattacharya, D. (2004) A molecular timeline for the origin of photosynthetic eukaryotes. *Mol. Biol. Evol.* **21**, 809–818.

Zhang, G.Q., Xu, Q., Bian, C. *et al.* (2016) The *Dendrobium catenatum* Lindl. genome sequence provides insights into polysaccharide synthase, floral development and adaptive evolution. *Sci. Rep.* **6**, 19029.