Lineage-Specific Reductions of Plastid Genomes in an Orchid Tribe with Partially and Fully Mycoheterotrophic Species

Yan-Lei Feng[1,9,†], Susann Wicke[2,†], Jian-Wu Li[3], Yu Han[4], Choun-Sea Lin[5], De-Zhu Li[6], Ting-Ting Zhou[1,9], Wei-Chang Huang[7], Lu-Qi Huang[8], and Xiao-Hua Jin[1,*]

[1]State Key Laboratory of Systematic and Evolutionary Botany, Institute of Botany, Chinese Academy of Sciences, Beijing, China

[2]Institute for Evolution and Biodiversity, University of Muenster, Germany

[3]Xishuangbanna Tropical Botanical Garden, Chinese Academy of Sciences, Menglun Township, Mengla County, Yunnan, China

[4]Nanchang University, Jiangxi, China

[5]Agricultural Biotechnology Research Center, Academia Sinica, Taipei, Taiwan

[6]Key Laboratory of Plant Diversity and Biogeography of East Asia, Kunming Institute of Botany, Chinese Academy of Sciences, Kunming, Yunnan, China

[7]Chenshan Shanghai Botanical Garden, Shanghai, Songjiang, China

[8]National Resource Centre for Chinese Materia Medica, China Academy of Chinese Medical Science, Beijing, China

[9]University of Chinese Academy of Sciences, Beijing, China

[†]These authors contributed equally to this work.

*Corresponding author: E-mail: xiaohuajin@ibcas.ac.cn.

Accepted: June 8, 2016

Data deposition: This project has been deposited at NCBI under the accession number KU551262-Ku551272.

Abstract

The plastid genome (plastome) of heterotrophic plants like mycoheterotrophs and parasites shows massive gene losses in consequence to the relaxation of functional constraints on photosynthesis. To understand the patterns of this convergent plastome reduction syndrome in heterotrophic plants, we studied 12 closely related orchids of three different lifeforms from the tribe Neottieae (Orchidaceae). We employ a comparative genomics approach to examine structural and selectional changes in plastomes within Neottieae. Both leafy and leafless heterotrophic species have functionally reduced plastid genome. Our analyses show that genes for the NAD(P)H dehydrogenase complex, the photosystems, and the RNA polymerase have been lost functionally multiple times independently. The physical reduction proceeds in a highly lineage-specific manner, accompanied by structural reconfigurations such as inversions or modifications of the large inverted repeats. Despite significant but minor selectional changes, all retained genes continue to evolve under purifying selection. All leafless *Neottia* species, including both visibly green and nongreen members, are fully mycoheterotrophic, likely evolved from leafy and partially mycoheterotrophic species. The plastomes of Neottieae span many stages of plastome degradation, including the longest plastome of a mycoheterotroph, providing invaluable insights into the mechanisms of plastome evolution along the transition from autotrophy to full mycoheterotrophy.

Key words: heterotrophy, plastid genome reduction, Neottieae, Orchidaceae, relaxed selection.

Introduction

Plastids are essential cell organelles and the site of many metabolic processes, including photosynthesis and carbon fixation, the biosynthesis of starch, lipids, pigments, and amino acids, among others (Neuhaus and Emes 2000; McNeal et al. 2007; Green 2011; Rousseau-Gueutin et al. 2011; Wicke et al. 2011). The importance of these functions, especially of photosynthesis, is also reflected in the plastid genome (plastome). Because of the constantly high selection pressure to suppress of eliminate potentially disadvantageous mutations, most land plant plastomes are highly conserved concerning structure and gene contents. The two single copy regions (large and

© The Author 2016. Published by Oxford University Press on behalf of the Society for Molecular Biology and Evolution.
This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (http://creativecommons.org/licenses/by-nc/4.0/), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited. For commercial re-use, please contact journals.permissions@oup.com

small single-copy region, i.e. LSC and SSC) and the two identical large inverted repeats (IRs) separating LSC and SSC normally encode 30 unique transfer RNAs, two sets of four ribosomal RNAs, and less than 100 protein-coding genes (Douglas 1998; Green 2011). And several hypotheses have been proposed to understand why chloroplasts retain their own genomes and genetic systems (Race et al. 1999; Barbrook et al. 2006; Allen 2015).

In heterotrophic land plants (including parasitic and mycoheterotrophic [MH] plants), photosynthesis-associated genes are no longer required, resulting in an extensive and rapid functional and physical reduction of their plastomes (Wolfe et al. 1992; Delannoy et al. 2011; Logacheva et al. 2011; Braukmann et al. 2013; Wicke et al. 2013). There appears to be a clear trend that genes for the thylakoid NAD(P)H dehydrogenase (*ndh* genes), the plastid-encoded polymerase (PEP; *rpo* genes), RuBisCO (*rbcL* locus), photosystems I and II (*psa* and *psb* genes, respectively), and the cytochrome b_6f complex (*pet* genes) are lost or pseudogenized soon after the transition to a heterotrophic and nonphotosynthetic lifeform, whereas the thylakoid ATP synthase complex (*atp* genes) and some ribosomal protein genes (*rpl* and *rps* genes) as well as tRNAs are lost later (Barrett and Davis 2012; Barrett et al. 2014; Cusimano and Wicke 2016). A case in point, the *rbcL* locus is a critical photosynthesis gene and therefore highly conserved in autotrophic plants (Kellogg and Juliano 1997). However, it displays a rather unexpected evolutionary pattern in MH plants (Barrett and Freudenstein 2008; Cafasso and Chinali 2012). In leafless *Corallorhiza*, *rbcL* is putatively functional in some visible green species and pseudogenized in other species (Barrett and Freudenstein 2008; Barrett et al. 2014). In *Neottia nidus-avis*, genomic rearrangements in the *rbcL* region and three distinct *rbcL* sequences had been detected (Cafasso and Chinali 2012). In *Epipogium* and *Rhizanthella*, *rbcL* has been completely lost in plastomes (Delannoy et al. 2011; Schelkunov et al. 2015).

Although mycoheterotrophs are very rare in nature, they are relatively common in Orchidaceae (Merckx et al. 2013a,b). About 235 leafless species in 43 genera across Orchidaceae are putatively full or nearly mycoheterotrophs (Merckx et al. 2013a,b). The tribe Neottieae (Epidendroideae, Orchidaceae) comprises about 100 species in six recognized genera, that is *Aphyllorchis*, *Cephalanthera*, *Epipactis*, *Limodorum*, *Neottia*, *Palmorchis*, out of which 50 species in four genera are putatively full or nearly mycoheterotrophs (Pridgeon et al. 2005; Xiang et al. 2012; Merckx et al. 2013a,b; Chase et al. 2015). There are three growth forms and trophic specializations in Neottieae: (1) autotrophic or partially mycoheterotrophic species with green leaves (Abadie et al. 2006; Liebel et al. 2010; Merckx et al. 2013a,b; Bellino et al. 2014; Gonneau et al. 2014) carrying out photosynthesis (e.g. *Neottia pinetorum*), (2) partially or entirely mycoheterotrophic leafless (the definition of leaf, see Beentje 2012) species with photosynthetic pigmentation (e.g. *Limodorum abortivum* (Girlanda et al. 2006; Merckx et al. 2013a,b; Bellino et al. 2014)), and (3) fully mycoheterotrophic and visibly nongreen taxa that have lost the ability to photosynthesize (e.g. *N. nidus-avis* (Zimmer et al. 2008)). It remains to be tested if all leafless species in *Aphyllorchis*, *Cephalanthera*, and *Neottia* truly are full mycoheterotrophs (Merckx et al. 2013a,b). While putatively fully mycoheterotrophic species of *Cephalanthera*, such as *Cephalanthera exigua*, are mostly visibly nongreen, containing only trace photosynthetic pigments (Cummingsand Welschmeyer 1998), spontaneous visibly nongreen variants (albinos) occur time and again in otherwise photosynthetic species such as *Cephalanthera damasonium* (Preiss et al. 2010; Stoeckel et al. 2011; Roy et al. 2013). These observations in Neottieae strongly suggest that plastomes of these partial mycoheterotrophs may be at the very beginning of degradation.

The tribe Neottieae is an excellent model system to understand the mechanisms underlying plastome degradation in heterotrophic plants and the series of functional losses, as this requires comparative analyses of closely related autotrophic and heterotrophic species, which, ideally, differ in their trophic specialization within a narrow phylogenetic lineage. Here, we compare the complete plastome sequences of 12 species from four genera of the Neottieae. Our sampling covers all three growth forms and trophic specializations within this orchid tribe (fig. 1) and allows us to infer the general modes and mechanisms of early-stage plastome reduction. Using comparative plastid genomics, we aim to (1) investigate the extent, progression, and tempo of reductive plastome evolution in Neottieae, and especially the molecular evolution of *rbcL*; (2) test the hypothesis that the leafless species in *Aphyllorchis*, *Cephalanthera*, and *Neottia* are fully mycoheterotrophic, (3) to know whether plastomes of leafless *Cephalanthera*, such as *Cephalanthera humilis*, are at the very beginning of degradation.

Materials and Methods

Taxon Sampling, DNA Isolation, Library Preparation, and Sequencing

We sampled 11 species from four Neottieae genera, including the leafy *Cephalanthera longifolia*, *Epipactis veratrifolia*, *Epipactis mairei*, *Neottia ovata*, *Neottia fugongensis*, *N. pinetorum*, the leafless but green *Neottia camtschatea* and *Neottia listeroides*, and the leafless visible nongreen *Aphyllorchis montana*, *Ce. humilis*, and *Neottia acuminata* (table 1). The previously published plastomes of leafless, visibly nongreen, and holo-mycoheterotrophic *N. nidus-avis* (Logacheva et al. 2011) and leafy, green, and autotrophic *Calanthe triplicata* (with 3–4 plicate leaves) (GenBank ID: NC_024544) complemented our analyses.

Total genomic DNA was extracted from silica-dried materials using a modified CTAB-method (Li et al. 2013). DNA with a concentration higher than 100 ng/μl were sheared to 500 bp, and sequencing libraries were prepared using the

Fig. 1.—Growth forms in Neottieae. (*A*) Leafy and photosynthetic *E. veratrifolia*, and (*B*) *N. pinetorum*. (*C*) Leafless and visibly green *N. camtschatea*. (*D*) Leafless and visible nongreen (nonphotosynthetic) *N. acuminata*, and (*E*) *Ce. humilis*.

Table 1
Length and GC Content of Plastid Genomes in Neottieae

	Length (in bp)				GC Content (in %)								
Species	**Length**	**LSC**	**SSC**	**IR**	**Total**	**LSC**	**SSC**	**IR**	**One IR**	**cds**	***rrn***	***trn***	**Noncoding**
A. montana	94,559				37.1				37.1	36.8	55.5	52.6	35.3
Ce. humilis	157,011	86,908	15,133	27,485	37.3	34.9	30.8	42.9	36.1	37.9	55.2	52.7	32.9
Ce. longifolia	161,877	88,806	19,187	26,942	37.2	35.0	30.6	43.2	36.0	37.9	55.2	52.7	31.7
E. mairei	159,019	86,377	18,816	26,913	37.3	35.1	30.5	43.2	36.1	38.0	55.2	52.8	31.6
E. veratrifolia	159,719	87,043	18,854	26,911	37.3	35.2	30.7	43.2	36.2	38.0	55.3	52.8	31.8
N. ovata	156,978	85,433	18,071	26,737	37.6	35.4	31.3	43.2	36.4	37.9	55.3	52.9	32.3
N. fugongensis	156,536	85,357	18,311	26,434	37.4	35.3	30.6	43.3	36.3	37.9	55.3	52.9	32.0
N. pinetorum	155,959	84,449	18,104	26,703	37.5	35.4	30.8	43.1	36.4	37.9	55.2	52.9	32.2
N. acuminata	83,190	51,145	5,371	13,337	36.6	33.1	26.7	45.2	34.9	34.3	54.9	52.1	32.0
N. camtschatea	106,385	52,960	9,273	22,076	37.2	33.5	28.8	43.3	35.6	35.3	54.9	52.5	33.3
N. listeroides	110,246	45,021	95,97	27,814	37.2	33.6	28.9	41.6	35.8	35.4	55.1	52.3	33.5
N. nidus-avis	92,060	36,423	78,21	23,908	34.4	29.2	25.3	39.9	32.5	33.2	54.4	51.0	28.4

total—total GC content, one IR—plastome excluding one IR, cds—protein-coding regions, *rrn*—ribosomal RNA genes, *trn*—tRNA genes; gray shading highlights leafless taxa.

NEBNext Ultra DNA Library Prep Kit (according to the manufacturer's protocol) for sequencing on an Illumina Hiseq 2500.

Read Mapping, Sequence Assembly, and Plastome Finishing

The raw reads (100 bp paired-end sequences) were filtered using NGSQCTOOLKIT v 2.3.3 (Patel and Jain 2012). We trimmed all bases with a quality lower than PHRED 20 and discarded all reads shorter than 70 bp after the trimming. The cleaned reads were mapped to *Ca. triplicata* in GENEIOUS 8.0 (Biomatters, Inc., Auckland, New Zealand; http://www.geneious.com, last accessed May 8, 2016) with medium-low sensitivity in five iterations. Used reads were exported and assembled using SOAPDENOVO2 (Luo et al. 2012), for which we chose a k-mer size from 37 to 67. Scaffolding was carried out in SSPACE (Boetzer et al. 2011). We additionally assembled all fully mycoheterotrophs *de novo* using VELVET (Zerbino and Birney 2008) with a k-mer size ranging from 37 to 45 and an auto-adjustment of the coverage cutoffs. The plastid sequences were extracted from the contigs pool using BLASTN 2.2.29+ (Camacho et al. 2009) and the *Ca. triplicata* plastome (Genbank ID: NC_024544) as the subject (reference) sequence. The extracted contigs were re-assembled and scaffolded using SSPACE, finished as described above, and reoriented to match the *Ca. triplicata* plastome reference; IR boundaries were determined by BLAST. All reads were mapped to the plastome scaffold to correct for assembly errors. Remaining gaps and stretches of

undetermined bases were closed by PCR and Sanger sequencing. We also corrected a one bp-deletion in the *matK* gene of the *N. nidus-avis* plastome by Sanger resequencing (voucher: Jin X.H. 14261, PE; primers and PCR conditions: Raskoti et al. 2016).

Annotation

We annotated the reconstructed plastomes of Neottieae using DOGMA with an *e* value of 5, a percent identity cutoff for protein coding genes and tRNAs of 60 or 80, respectively (Wyman et al. 2004), as well as the GENEIOUS annotation tool with the plastomes of *Ca. triplicata* and *Oncidium* Gower Ramsey (Genbank ID: NC_014056.1) as references. We annotated genes as pseudogenes if they showed disruptions of the reading frame, premature stop codons or frameshifts caused by nontriplet insertions and deletions; their assemblies were inspected by eye to confirm that the pseudogenes were not due to assembly errors. We used OGDRAW (Lohse et al. 2013) to draw linear plastome maps for all species (Lohse et al. 2013). All fully annotated plastome sequences were deposited in NCBI Genbank under accession numbers KU551262–KU551272.

Repeat Analyses

We removed one IR region for all repeat analyses, for which we employed REPUTER to quantify forward and palindromic repeats longer than 25 nt, setting the Hamming distance to 3; all duplicate hits were removed. We quantified tandem repeats with MREPS (Kolpakov et al. 2003), which was run with a minimal and maximal period of 5 and 30, respectively. Statistical analyses and plots were performed using the R statistical software (http://www.r-project.org, last accessed May 8, 2016).

GC Content and Codon Usage

We analyzed the GC content using GENEIOUS and/or a custom Perl script. Codon usage, base composition, and the effective number of codons (ENC) in protein-coding genes were computed using CODONW (http://codonw.sourceforge.net/, last accessed May 8, 2016). The subsequent statistical analyses and hypothesis testing were performed in the R statistical framework. Pairwise Wilcoxon tests were employed to evaluate differences in GC content at different codon positions (GCn) between pairs of species for 27 genes commonly retained in the investigated species. We specifically tested if the change of GC content influences the codon usage in the retained genes of Neottieae. Published plastomes of Epidendroideae (Orchidaceae) were analyzed and the well-annotated *Ca. triplicata* (Genbank ID: NC_024544) was selected as a reference to build a relative distance of GC content at the different codon positions. The results were illustrated using R in combination with the ggplot2 package (http://ggplot2.org/, last accessed May 8, 2016).

Phylogenetic Analyses and Molecular Dating

For our phylogenetic analyses using complete plastid genome sequences, we used *Ca. triplicata* (Genbank ID: NC_024544), *Dendrobium officinale* (Genbank ID: KJ862886), and *Erycina pusilla* (Genbank ID: NC_018114) as outgroups. We subdivided the plastomes into 26 conserved and unique segments based on the position of housekeeping genes and the deletion of photosynthetic genes of holo-mycoheterotrophic species; the orientation of these segments were adjusted accurately. All segments were aligned using MAFFT under the automatic model selection option (Katoh et al. 2002) with some manual adjustments, and subsequently combined into a single plastome matrix in SEQUENCEMATRIX (Vaidya et al. 2011). We analyzed the aligned data with maximum likelihood (ML) using RAXML (Stamatakis 2014) under the GTR + Γ model on the CIPRES SCIENCE GATEWAY webserver (www.phylo.org, last accessed May 8, 2016), and with Bayesian inference (BI) using MrBayes 3.2.3 (Ronquist et al. 2012). The best model (Lset nst = 6 rates = invgamma; Prset statefreqpr = dirichlet(1,1,1,1)) for BI analyses was selected using MrModeltest v2.3 (Nylander et al. 2004). We used two separate MrBayes runs with four parallel Monte Carlo Markov chains (MCMC), sampling every 1000th generation over ten million replicates, and computed a majority rule (>50%) consensus tree after removing 25% of the samples as 'burn-in'.

Molecular dating among Neottieae lineages was assessed using r8s 1.7 under a penalized likelihood (PL) method (Sanderson 2003) and the truncated Newton algorithm on the ML tree. A previous molecular phylogenetic dating of the Orchidaceae estimated the split between *Palmorchis* and remaining Neottieae at 30.9 Ma (Xiang et al. 2016).

Analyses of Selective Regimes

The synonymous and nonsynonymous substitution rates (d*S*, d*N*) of the concatenated 27 protein-coding housekeeping genes present in the plastome of each species were analyzed in PAML v.4.8 using the CODEML module (Yang et al. 2007). Selective regimes among branches were analyzed in PAML v.4.8, using the "branch" models in the CODEML module to test if the ratio of nonsynonymous to synonymous changes ($dN/dS = \omega$) differs between branches in the 27 commonly retained genes (Yang 1997, 2007). We specifically tested for differences in ω between (1) Neottieae and autotrophic outgroups, (2) leafless Neottieae and leafy Neottieae plus outgroups, (3) visibly nongreen Neottieae and visibly green Neottieae plus leafy outgroups, and (4) *Neottia* s.l. and autotrophic outgroup. We performed a likelihood ratio tests (LRTs) against a null model that assumes a global ω as described earlier (Barrett et al. 2014); LRT *P* values were Bonferroni-corrected to account for multiple comparisons (Yang 2007). To assess the extent of relaxed selection, that is either relaxed purifying or positive selection, in retained plastid genes during and after the transition to mycoheterotrophy in orchids

in general, we used the RELAX branch-site random effects likelihood method (Wertheim et al. 2015), implemented in the Datamonkey web server (Delport et al. 2010). To this end, plastid gene data of all orchid plastomes published in Genbank as of December 05, 2015 were downloaded (47 species) and complemented with our Neottieae data, and a dataset comprising all housekeeping genes was prepared as described earlier (Cusimano and Wicke 2016). We ran two tests with different branch-partitioning patterns to evaluate, firstly, if the relaxation of selection occurs only in nonphotosynthetic species compared to photosynthetic ones, and, secondly, if the relaxation of selection already occurs with the transition to a mycoheterotrophic lifestyle and irrespective of the photosynthetic capability.

Results

Size, Gene Content, and Structure of Plastomes of Neottieae

Plastid genomes of Neottieae range between 83,190 bp in *N. acuminata* and 161,877 bp in *C. longifolia* (fig. 2, table 1). The fully mycoheterotrophic *Ce. humilis* retains a plastome of 157,011 bp in length, and, hence, the longest plastome of any fully mycoheterotrophic species sequenced so far. The proportion of plastid reads in whole-genome shotgun data amounted to usually >1% in leafy species (except for *N. ovata*, 0.43%), whereas reads of plastid origin were less than 0.4% in the leafless taxa (supplementary table S1, Supplementary Material online).

Plastomes of Neottieae are mostly collinear with those of autotrophic orchids, with only a few exceptions. For example, the plastome of *N. acuminata* has an 11 kb inversion in the LSC between its *psbA* pseudogene (Ψ*psbA*) and *rps2*. Considerably shorter inversions occur in the SSC of *C. humilis* between *rpl32* and Ψ*ccsA*, and in the intergenic spacer between *atpA* and *atpF* of *A. montana*. Variations of the IRs is especially common in leafless Neottieae (supplementary fig. S1, Supplementary Material online). Whereas *A. montana* has lost one copy of the IR (fig. 2), *N. listeroides* has a slightly expanded IR that extends to the *rpl16* region. In *N. acuminata*, IRs range beyond the *trnL-CAA* gene but no longer enclose *the ycf2* gene. Similar to other plants, plastomes of Neottieae are rare in dispersed repeats in general, with only a handful of repeats longer than 60 bp (supplementary fig. S2, Supplementary Material online). Only the plastome of *N. listeroides* harbors a considerably higher number of tandem repeats.

Pseudogene and Gene Loss

Similar to the structural changes to the plastome, physical deletions of genes are lineage-specific despite a common pattern in the series of losses (fig. 3 and supplementary fig. S3, Supplementary Material online). Except *Ce. humilis* (fig. 4 and supplementary table S2, Supplementary Material online), leafless Neottieae have functionally lost genes related to photosynthesis (i.e. *ndh*, *pet*, *psa*, *psb*, *rpo*, *atp* genes, plus *ccsA*, *cemA*, *rbcL*, *ycf3*, and *ycf4*) and PEP (*rpo* genes) (fig. 4 and supplementary table S2, Supplementary Material online). In *Ce. humilis*, most *psa* and *psb* genes have intact open reading frames (ORFs), except for *psaA*, *psaB*, *psbK*, whereas *A. montana* and *N. camtschatea* both retain only five intact *psa* and *psb* genes, respectively (fig. 4 and supplementary table S2, Supplementary Material online). The *rbcL* locus is intact in *Ce. humilis* but exhibits many indels, frameshift mutations or premature stop codons in the plastomes of the other five leafless *Neottieae* species. There are six premature stop codons in *the rbcL* region of *N. acuminata* and 12 indels in *N. listeroides*. *Aphyllorchis montana* retains only a truncated copy of *rbcL* (1–885 bp) (fig. 5).

Housekeeping genes such as *accD*, *matK*, *clpP*, and *ycf1* have intact ORFs in most Neottieae, except for *N. nidus-avis* in which also *rps16* and *rps18* both have premature stop codons. Some housekeeping genes such as *rps4*, *rps18*, and *rpl16* have accumulated unique indels or frameshift mutations towards their 3′ ends in various leafless or nonphotosynthetic species (supplementary table S2, Supplementary Material online). The *clpP* genes of *N. camtschatea*, *N. listeroids*, and *N. nidus-avis* has a premature stop codon 30 bp upstream of its known gene stop, thus possibly producing a 10 amino acid shorter protein in these species. Although all seven group-IIA introns that require *matK* for splicing are present in Neottieae, this plastid-encoded maturase gene shows a premature stop codon 33 bp downstream of the initiator in some *Neottia* species (supplementary fig. S3, Supplementary Material online). This premature stop is, however, followed immediately by another potential start codon (supplementary fig. S3, Supplementary Material online).

Leafy and green Neottieae (*E. mairei*, *N. pinetorum*, and *E. veratrifolia*) have functionally lost one or two *ndh* genes. Some of these species (e.g. *N. fugongensis*) also show evidence for the nonfunctionalization of PEP, although no genes have been physically deleted from their plastomes.

GC Content and Codon Usage

The GC content differs notably in Neottieae, ranging from 34.4% in *N. nidus-avis* up to 37.6% in *N. ovata* (table 1). In protein-coding regions, GC reduction amounts to between 0.06% and 4.3% in the leafless species, while in structural RNAs it remains nearly unaltered (table 1). The IR-lacking fully mycoheterotrophic *A. montana* possesses the highest GC content in its noncoding regions (35.3%), compared to 32.5% in the leafy species and <34% in other leafless species (table 1). The three leafless species *N. acuminata, N. camtschatea*, and *N. nidus-avis* show a marked AT richness in both the first and second position, whereas the GC content in *A. montana* is

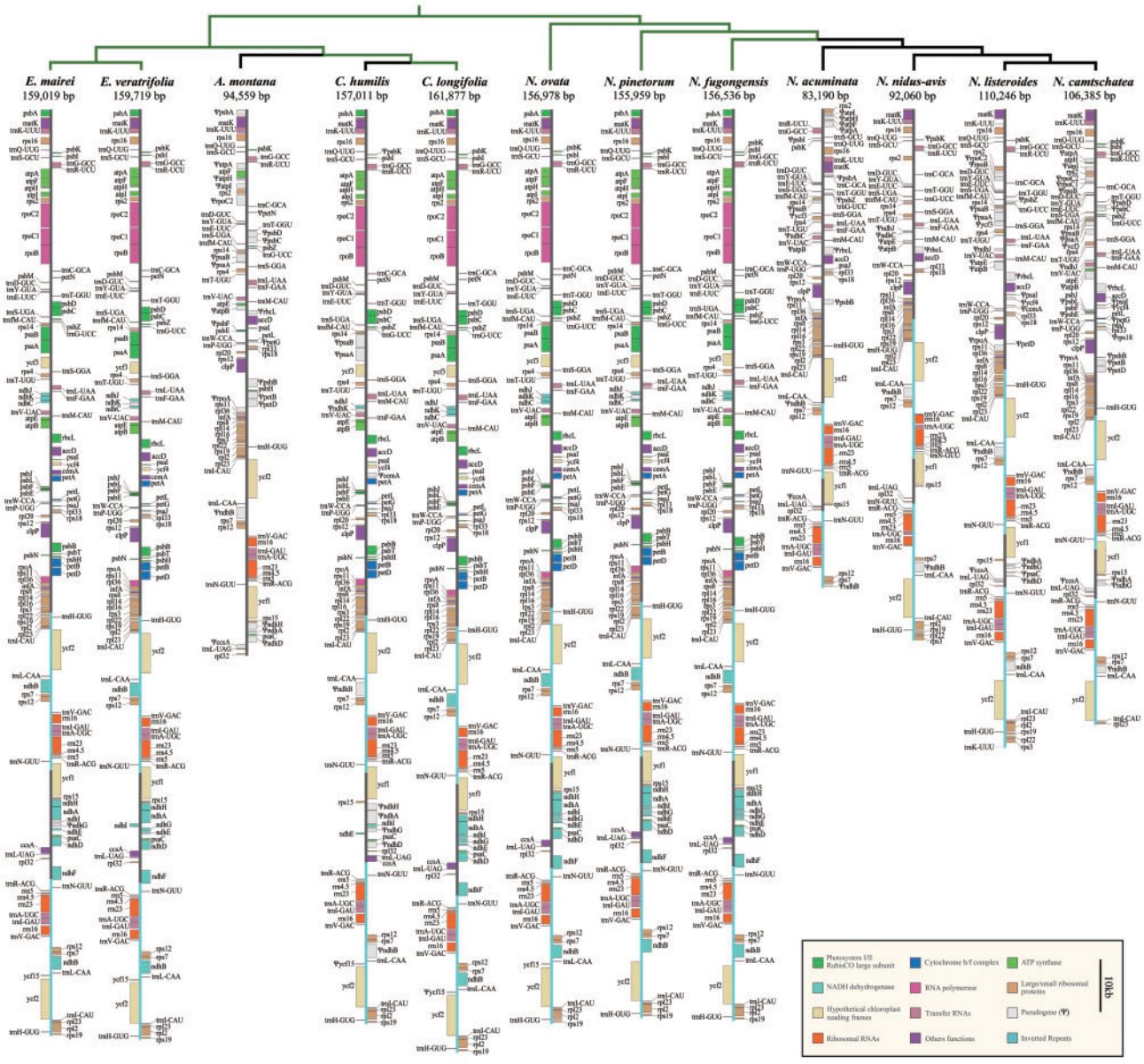

Fig. 2.—Plastid genome structures of 12 Neottieae species. Genes are colored by function. Pseudogenes are shown in gray and marked with a 'Ψ'.

significantly elevated only in the second and third position (supplementary fig. S4, Supplementary Material online).

Most leafless Neottieae utilize codons just like the leafy species, although we observe some minor under- or overuses (e.g. Val^{GUC}, Thr^{ACG} in *A. montana*; Arg^{AGA}, Asn^{AAU}, Gly^{GGA}, Pro^{CCG} in *N. acuminata*; Cys^{UGU} in *N. camtschatea*; Ala^{GCU}, Asp^{GAU}, Ile^{AUA}, Ser^{UCA}, Ter^{UAA} in *N. nidus-avis*; Ter^{UAG} in *N. listeroides*). The most pronounced differences are found for the Ile^{AUC}, whose usage rises from <14% to >21% in *A. montana*, *N. nidus-avis*, and *N. listeroides*, as well as for the terminator tRNA Ter^{UAA}, which is used <42% in *A. montana* compared to >59% in *N. nidus-avis* (supplementary table S3, Supplementary Material online).

Phylogenetic Analyses and Molecular Dating

The topology of the trees from our ML inference and BI results are mostly congruent with the MP results, with only one exception in the *Neottia* crown group. The relationships among Neottieae are mostly well resolved with high supported, except for the *Cephalanthera–Aphyllorchis* clade (ML = 70) (fig. 3A). *Neottia* species are sister to a clade formed by *Cephalanthera* and *Aphyllorchis*, which themselves are sister to *Epipactis*. Within *Neottia*, both ML and BI analyses show that the three leafy species *N. ovata*, *N. pinetorum*, and *N. fugongensis* are successive sisters to a grade of the four leafless *Neottia* (pp = 1, ML = 100). *N. acuminata* and *N. nidus-avis*

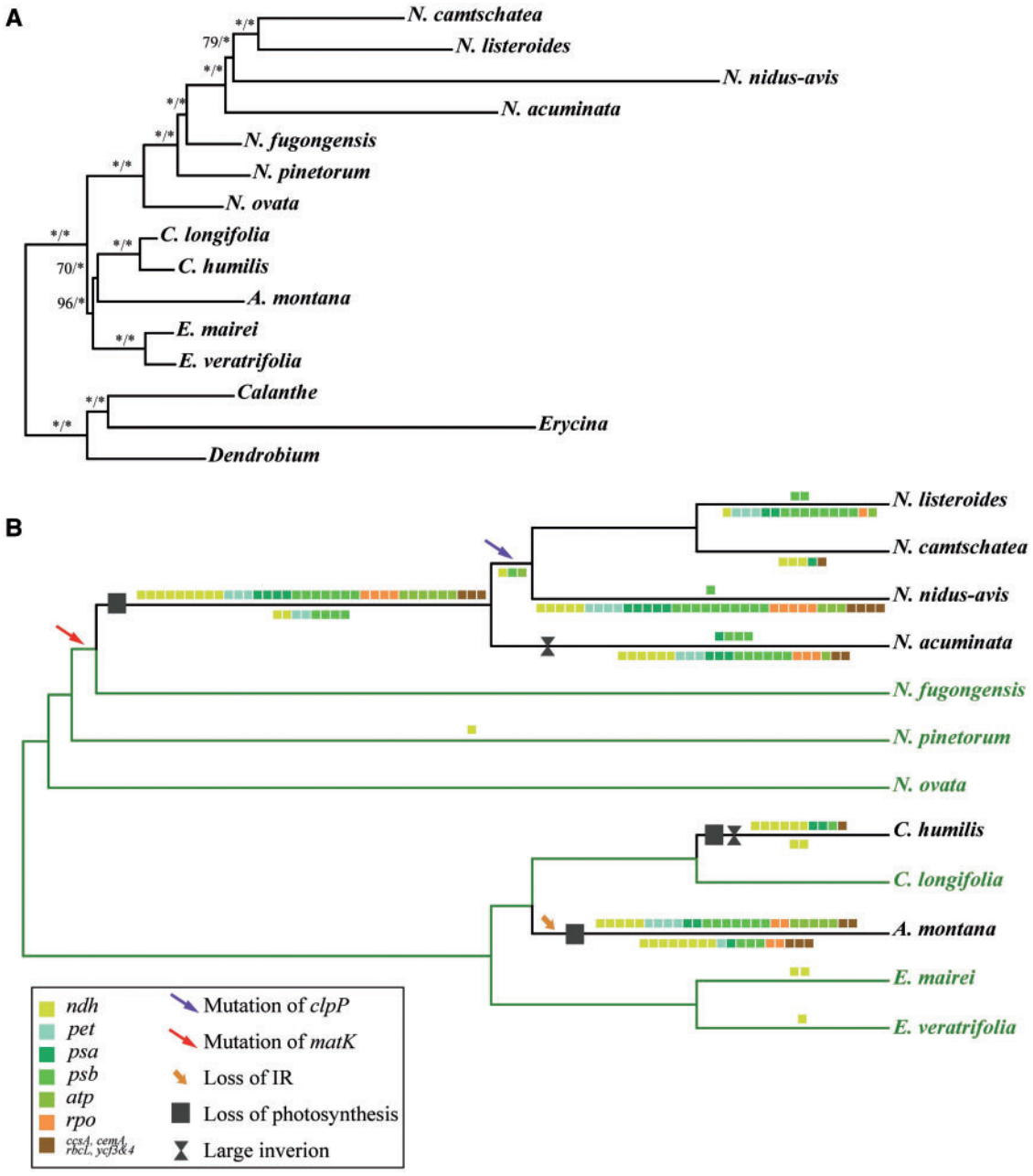

Fig. 3.—Phylogenetic results and the evolution of plastid genomes in Neottieae. (*A*) The maximum likelihood (ML, BI) trees inferred from whole plastomes. Number above branches are the bootstrap support (ML) and posterior probabilities (BI), respectively. (*B*) Functionally and physically lost genes are mapped on the plastome tree, showing pseudogenes above branches and genes that were deleted from the plastome below.

are successive sisters to the two visibly green *N. listeroides* and *N. camtschatea* in ML and BI results (fig. 3A). In our MP analysis, *N. acuminata* and *N. nidus-avis* together form a clade that is sister to *N. listeroides* and *N. camtschatea* (supplementary fig. S5, Supplementary Material online).

The split between *Neottia* and the remaining three genera was immediately followed by the divergence of *Epipactis* and the *Aphyllorchis–Cephalanthera* clade at 28.98 and 27.74 million years ago (Ma). We inferred the age of the four leafless species of *Neottia* at 20.63 Ma, *A. montana* at 27.74 Ma, and *Ce. humilis* at 10.07 Ma (fig. 5).

Analyses of Selective Regimes

Species of *Neottia*, including both leafy and leafless ones, have a much higher synonymous and nonsynonymous substitution

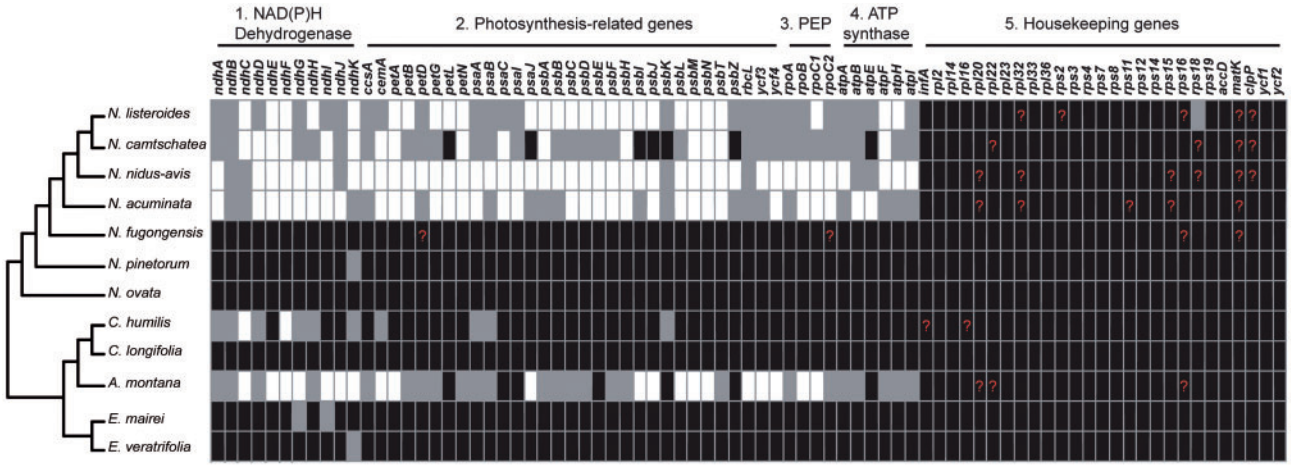

Fig. 4.—Protein-coding genes in Neottieae. Intact genes per species are indicated by black boxes, whereas gray and white boxes mark functional and or physical losses, respectively. Red question marks highlight uncertain calls. PEP—plastid-encoded RNA polymerase.

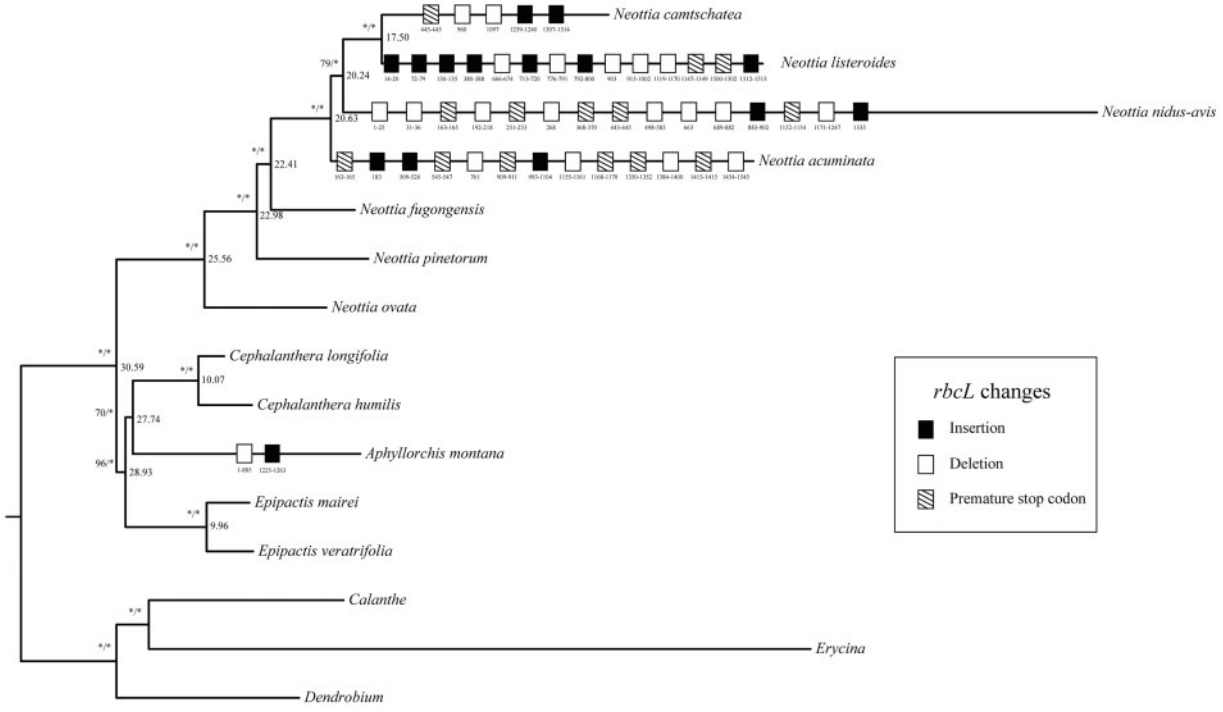

Fig. 5.—Changes in the *rbcL* locus over time in leafless Neottieae. Changes such as insertions (black boxes), deletions (white), and premature stop codons (striped) in *rbcL* are mapped on the branches of all leafless taxa. Divergence ages (in million years) inferred from a PL analysis are given in bold on the right-hand side of all Neottieae internodes.

rates (d*S*, d*N*) in the concatenated housekeeping genes than the remaining Neottieae (supplementary table S4, Supplementary Material online). Leafless and visibly nongreen *Neottia* have the highest d*S* and d*N* (supplementary table S4 and fig. S6, Supplementary Material online), and leafless *Ce. humilis* exhibits similar d*S* and d*N* as leafy *Ce. longifolia* (supplementary table S4 and fig. S7, Supplementary Material online). The ratio of d*N*/d*S* (ω) of Neottieae is <1, *A. montana* has the highest ω (supplementary table S4 and fig. S6, Supplementary Material online).

A "two-ratio" likelihood model (M_1), which allows all Neottieae to adopt a different ω than the outgroups, has a

significantly better fit for *rpl*16, *rpl*33, *rps*18, and *rps*2 than the basic "one-ratio" model, M_0 (Bonferroni-corrected LRTs between M_0 and M_1, P values all <0.05; supplementary table S4 Supplementary Material online). The "three-ratio" model, M_2, separating the leafless Neottieae from other Neottieae and from the outgroup applies only for *ycf1* (LRT M_2 vs. M_1: $P < 0.01$), but is outperformed by both a three- and a six-ratio model (M_3 and M_4, respectively) allowing for different ω between *Ce. humilis*, *A. montana*, and the leafless *Neottia*, or between *Ce. humilis*, *A. montana*, green and leafless *Neottia*, nongreen *Neottia*, and leafless *Neottia* (M_3 vs. $M_2 < 0.001$; M_2 vs. M_4: $P < 0.001$) (supplementary table S5, Supplementary Material online). When all housekeeping genes are concatenated for analyses, no model is significantly better than one-ratio model (supplementary table S5, Supplementary Material online).

The multiple transitions to a mycoheterotrophic lifestyle within orchids (supplementary fig. S7, Supplementary Material online) coincide with significant changes in the selective pressures of the retained plastid genes. Housekeeping genes in the ten visibly nongreen lineages show a slightly relaxed purifying selection (selection intensity parameter $k = 0.75$, likelihood ratio (LR) = 9.06, $P = 0.003$). However, mean ω is only slightly different between visibly green ($\omega = 0.417$) and visibly nongreen orchids ($\omega = 0.553$), suggesting that housekeeping genes continue to evolve under purifying selection in the latter. A comparison of both visibly green and visibly nongreen mycoheterotrophs versus autotrophic orchids also indicates a weak trend for relaxed purifying selection ($k = 0.99$), which is, however, not significant ($P = 0.118$).

Discussion

Phylogenetic Relationships in Neottieae

The phylogeny of Neottieae remained unresolved in previous molecular systematics based on several loci (Bateman et al. 2004; Roy et al. 2009; Selosse and Roy 2009; Xiang et al. 2012). *Cephalanthera* sometimes was placed as sister to the clade formed by *Neottia* + *Aphyllorchis* + *Epipactis*, albeit with no statistic support (Bateman et al. 2004; Xiang et al. 2012). Our data resolves the phylogeny of Neottieae well, showing that *Neottia* is sister to a clade of *Epipactis*, *Aphyllorchis*, and *Cephalanthera*. These results are corroborated by several vegetative and floral traits (Burnsbalogh et al. 1987; Pridgeon et al. 2005; Chen et al. 2009). A substitution in the plastid *clpP* gene shared by three of the investigated Neottia species further supports our inferred close relationships of *N. camtschatea*, *N. listeroides*, and *N. nidus-avis*. The minor incongruency of the four mycoheterotrophic *Neottia* species seen in the ML and BI trees compared to the MP analyses may be the result of a long-branch attraction phenomenon in MP.

Plastome Reduction in Neottieae and Orchidaceae

Leafless members of Neottieae, including the two leafless green species, all show evidence of the degradation of genes that are directly involved in photosynthetic processes (fig. 3 and supplementary fig. S3, Supplementary Material online), albeit with an extraordinary variation in the number and nature of the genes lost or pseudogenized. Despite several intact photosynthesis genes (e.g. *petL*, *psaI*, *psbI*, *psbJ*, *psbK*, and *psbZ*), most species have lost subunits required for light harvesting and CO_2 fixation. Unlike in *Corallorhiza striata* var. *vreelandii* (Barrett et al. 2014) and *N. nidus-avis* (Logacheva et al. 2011), *ccsA*, *cemA*, *psaA*, and *psaB* are among the first to be lost in mycoheterotrophic Neottieae. These findings suggest that, despite the presence of chlorophyll, these plants likely are unable to carry out photosynthesis anymore (Girlanda et al. 2006; Cameron et al. 2009). Moreover, recent results indicate that the degradation of plastid genomes is independent of chlorophyll synthesis in some nonphotosynthetic parasites (Cameron et al. 2009; Wickett et al. 2011). The photosynthesis genes that have escaped deleterious mutations normally are short in length or benefit from some protection by nearby essential genes (Wicke et al. 2013). Genes for the thylakoid ATP synthase, which appears to play a significant role in some nonphotosynthetic plants (Kamikawa et al. 2015), are either pseudogenes or deleted from the plastomes of leafless Neottieae, except in *Ce. humilis*.

Although required under stress conditions in many plants (Peltier and Cournac 2002), *ndh* genes are often functionally or even physically lost in Orchidaceae (Wu et al. 2010; Lin et al. 2015), including in all mycoheterotrophic orchids studied so far. Although leafless Neottieae retain only a few *ndh* genes, the sets of retained *ndh* genes genes imply that the loss of the NAD(P)H dehydrogenase complex seems to have occurred at least four times independently within the tribe (fig. 4 and supplementary table S2, Supplementary Material online). For example, in *E. mairei* the genes *ndhG* and *ndhI* are pseudogenes, whereas intact ORFs of these exist in *E. veratrifolia*. In this species, we find a pseudogene of *ndhK*, which is, however, intact in *E. mairei*.

The transition to mycoheterotrophy, the progressive trophic specialization and the eventual loss of photosynthesis also alters the selective regimes in several housekeeping genes of Neottieae (supplementary tables S4, S5, and fig. S6, Supplementary Material online), although the pattern of selectional shifts itself is rather erratic. It seems that purifying selection has been relaxed in leafless non-Neottieae species than in leafless Neottieae. The loss of housekeeping genes often occurs rather late during the reductive plastome evolution in nonphotosynthetic plants. However, some Neottieae show evidence for the pseudogenization of genes for transcription (e.g. *rpoC2*), translation (e.g. *rps18*), and RNA maturation (e.g. *matK*). The latter has been shown recently to be expressed in Neottieae (Barthet et al. 2015), but our data

indicate that some of the mycoheterotrophs may utilize an alternative start codon. The gene *clpP* encoding a subunit of the *Clp* protease), which is retained also in most other nonphotosynthetic lineages except for Apodanthaceae (Bellot and Renner 2015), may produce a protein that is ten amino acids shorter in the three leafless *Neottia* species than in other Neottieae (fig. 3A), but otherwise apparently intact. Most of the remaining housekeeping genes, such as *rpl*, *rps*, *trn*, *rrn*, *rpo*, *rpl*, *ycf1*, and *ycf2*, are intact and putatively functional in most Neottieae (but see Logacheva et al. 2011).

Our results of the molecular dating and the analysis of plastid genome structures indicate that *C. humilis* might be one of the most recently evolved completely mycoheterotrophic species in Neottieae. *C. humilis* clearly is at the earliest stage of plastome reduction with only a few functional losses in the plastid photosystem genes. Its GC content is still very similar to that of close photosynthetic relatives (table 1). Although *Aphyllorchis* and *Cephalanthera* split ca. 27.74 Ma ago (fig. 5), *A. montana* may also be a recently evolved species. It, too, shows plastome reduction at an early stage, which we characterize as having a rather high GC content and a large number of putatively functional photosynthesis genes. It needs more work to elucidate if the loss of one copy of the IR in *A. montana* has occurred already in the common ancestor of *Aphyllorchis*. In contrast to *Cephalanthera* and *Aphyllorchis*, *N. acuminata*, *N. camtschatea*, *N. listeroides*, and *N. nidus-avis* are already at rather late stages. *Neottia* species already have begun to reduce the proportion of nonessential DNA after the massive pseudogenization of most photosynthesis-related genes. Their plastomes' GC contents decrease gradually alongside the physical reduction. Plastomes size varies considerably among different species, with reductions of up to 47% compared to the leafy congenerics.

Nuclear copies may have replaced a few plastid housekeeping genes (*rps16* and *rps18*), which seem to have been lost functionally from the plastome. The molecular evolution of the *rbcL* locus is highly lineage-specific in Neottieae. While *rbcL* is putatively functional in *C. humilis*, it is a pseudogene in the four leafless Neottia species and *A. montana*. There are about 44 indels in the *rbcL* locus of the four leafless *Neottia*, however, none of them is shared (fig. 5).

The Evolution of Mycoheterotrophy in Neottieae and Orchidaceae

Our findings indicate that the transition to a fully mycoheterotrophic lifestyle has evolved at least three times independently during the evolution of Neottieae (fig. 2A), that is, once in *Aphyllorchis*, in *Cephalanthera*, and in *Neottia*. In Neottieae, most investigated leafy species are partial mycoheterotrophic species (Bidartondo et al. 2004; Julou et al. 2005; Abadie et al. 2006; Girlanda et al. 2006; Selosse and Roy 2009; Liebel et al. 2010), and leafless species are full mycoheterotrophs (Cameron et al. 2009; Logacheva et al. 2011). In *Neottia*, the leafy partially mycoheterotrophic *N. ovata* is sister to the four fully mycoheterotrophs, which together nest deeply in the leafy clade (fig. 2A). It, thus, seems that the ancestor of *Neottia* was a leafy and partial mycoheterotroph from which the nonphotosynthetic mycoheterotrophic lifestyle evolved only once. In *Cephalanthera*, the visibly nongreen *C. austinae* also nests within leafy and partially mycoheterotrophic species (Abadie et al. 2006). In support of an earlier hypothesis (Abadie et al. 2006), we may thus assume that the last common ancestor of *Aphyllorchis* and *Cephalanthera* was partially mycoheterotrophic, too (Merkx et al. 2013).

All investigated plastomes of leafless partially or fully mycoheterotrophic orchids such as *N. nidus-avis.* (Logacheva et al. 2011), *Epipogium roseum* (Schelkunov et al. 2015), and *Epipogium aphyllum* [19] have lost genes for the photosynthesis machinery, with the notable exception of *Corallorhiza trifida* (Zimmer et al. 2008). The plastome of *C trifida* retains many intact photosynthesis-related genes (e.g. *psa*, *psb*, *pet* genes, and *rbcL*) (Barrett et al. 2014), although it physiologically resembles the fully mycoheterotrophic condition (Cameron et al. 2009). This observation tempts us to speculate that the loss of leaves precedes the reduction of plastomes in some partially mycoheterotrophic orchids.

In sum, our study demonstrates that the degradation of plastomes in Neottieae not only agrees with the previously proposed general trends (Barrett and Davis 2012; Wicke et al. 2013; Barrett et al. 2014), but also underlies highly lineage-specific patterns of physical and functional reductions with many potentially independent gene losses. The shift to a leafless growth form in Neottieae *per se* appears to be closely associated with a significant decrease in plastome size and changes of the selectional regimes in retained housekeeping genes. The loss of leaves in combination with the onset of plastome degeneration including several functional gene losses and relaxed purifying selection in some genes imply that the transition to an obligate mycoheterotrophic lifestyle relaxes selection on plastid gene function rather than only the eventual loss of photosynthesis. By shedding new light on the earliest functional changes during the transition from partial to fully mycoheterotrophy, the complete plastid genome sequences of Neottieae contribute significantly to our understanding of the various routes of plastome degeneration in heterotrophic plants in general.

Supplementary Material

Supplementary figures S1–S7 and tables S1–S5 are available at *Genome Biology and Evolution* online (http://www.gbe.oxfordjournals.org/).

Acknowledgments

This study was supported by grants from the National Natural Science Foundation of China (Grant No. 31470299), and the Chinese Special Fund for Medicine Research in the Public Interest (201407003).

Literature Cited

Abadie J-C, et al. 2006. *Cephalanthera longifolia* (Neottieae, Orchidaceae) is mixotrophic: a comparative study between green and nonphotosynthetic individuals. Can J Bot.-Revue Canadienne De Botanique 84:1462–1477. doi: 10.1139/b06-101

Allen JF. 2015. Why chloroplasts and mitochondria retain their own genomes and genetic systems: Colocation for redox regulation of gene expression. Proc Natl Acad Sci U S A. 112:10231–10238. doi: 10.1073/pnas.1500012112

Barbrook AC, Howe CJ, Purton S. 2006. Why are plastid genomes retained in non-photosynthetic organisms? Trends Plant Sci. 11:101–108. doi: 10.1016/j.tplants.2005.12.004

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

Barrett CF, Freudenstein JV. 2008. Molecular evolution of rbcL in the mycoheterotrophic coralroot orchids (*Corallorhiza* Gagnebin, Orchidaceae). Mol Phylogenet Evol. 47:665–679. doi: 10.1016/j.ympev.2008.02.014

Barrett CF, et al. 2014. Investigating the path of plastid plastome degradation in an early-transitional clade of heterotrophic orchids, and implications for heterotrophic angiosperms. Mol Biol Evol. 31:3095–3112. doi: 10.1093/molbev/msu252

Barthet MM, Moukarzel K, Smith KN, Patel J, Hilu KW. 2015. Alternative translation initiation codons for the plastid maturase *MatK*: unraveling the pseudogene misconception in the Orchidaceae. BMC Evol Biol. 15:210. doi: 10.1186/s12862-015-0491-1

Bateman RM, Hollingsworth PM, Squirrell J, Hollingsworth ML. 2004. Tribe Neottieae: phylogenetics. In: Pridgeon AM, Cribb PJ, Chase MW, Rasmussen FN, editors. Genera Orchidacearum (4): Epidendroideae (Part 1). Oxford, UK: Oxford University Press. p. 487–515.

Beentje H. 2012. The Kew Plant Glossary: an illustrated dictionary of plant terms. London: Kew Publishing.

Bellino A, et al. 2014. Nutritional regulation in mixotrophic plants: new insights from *Limodorum abortivum*. Oecologia 175:875–885. doi: 10.1007/s00442-014-2940-8

Bellot S, Renner SS. 2015. 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. doi:10.1093/gbe/evv1251.

Bidartondo MI, et al. 2004. Changing partners in the dark: isotopic and molecular evidence of ectomycorrhizal liaisons between forest orchids and trees. Proc R Soc B-Biol Sci. 271:1799–1806. doi: 10.1098/rspb.2004.2807

Boetzer M, Henkel CV, Jansen HJ, Butler D, Pirovano W. 2011. Scaffolding pre-assembled contigs using SSPACE. Bioinformatics 27:578–579. doi: 10.1093/bioinformatics/btq683

Braukmann T, Kuzmina M, Stefanovic S. 2013. Plastid genome evolution across the genus *Cuscuta* (Convolvulaceae): two clades within subgenus *Grammica* exhibit extensive gene loss. J Exp Bot. 64:977–989. doi: 10.1093/jxb/ers391

Burnsbalogh P, Szlachetko DL, Dafni A. 1987. Evolution, pollination, and systematics of the tribe Neottieae (Orchidaceae). Plant Syst Evol. 156:91–115. doi: 10.1007/bf00937204

Cafasso D, Chinali G. 2012. Multiple and different genomic rearrangements of the rbcL gene are present in the parasitic orchid *Neottia nidus-avis*. Genome 55:629–637. doi: 10.1139/g2012-057

Camacho C, et al. 2009. BLAST plus: architecture and applications. BMC Bioinformatics 10:421. doi: 10.1186/1471-2105-10-421

Cameron DD, Preiss K, Gebauer G, Read DJ. 2009. The chlorophyll-containing orchid *Corallorhiza trifida* derives little carbon through photosynthesis. New Phytol. 183:358–364. doi: 10.1111/j.1469-8137.2009.02853.x

Chase MW, et al. 2015. An updated classification of Orchidaceae. Bot J Linn Soc. 177:151–174. doi: 10.1111/boj.12234

Chen SC, et al. 2009. Flora of China. Beijing: Science Press.

Cummings MP, Welschmeyer NA. 1998. Pigment composition of putatively achlorophyllous angiosperms. Plant Syst Evol. 210:105–111. doi: 10.1007/bf00984730

Cusimano N, Wicke S. 2016. Massive intracellular gene transfer during plastid genome reduction in nongreen Orobanchaceae. New Phytol. 210:680–693. doi: 10.1111/nph.13784

Delannoy E, Fujii S, des Francs-Small CC, Brundrett M, Small I. 2011. Rampant gene loss in the underground orchid *Rhizanthella gardneri* highlights evolutionary constraints on plastid genomes. Molecular Biology and Evolution 28:2077–2086. doi: 10.1093/molbev/msr028

Delport W, Poon AFY, Frost SDW, Pond SLK. 2010. Datamonkey 2010: a suite of phylogenetic analysis tools for evolutionary biology. Bioinformatics 26:2455–2457. doi: 10.1093/bioinformatics/btq429

Douglas SE. 1998. Plastid evolution: origins, diversity, trends. Curr Opin Genet Dev. 8:655–661. doi: 10.1016/s0959-437x(98)80033-6

Girlanda M, et al. 2006. Inefficient photosynthesis in the Mediterranean orchid *Limodorum abortivum* is mirrored by specific association to ectomycorrhizal Russulaceae. Mol Ecol. 15:491–504. doi: 10.1111/j.1365-294X.2005.02770.x

Gonneau C, et al. 2014. Photosynthesis in perennial mixotrophic *Epipactis* spp. (Orchidaceae) contributes more to shoot and fruit biomass than to hypogeous survival. J Ecol. 102:1183–1194. doi: 10.1111/1365-2745.12274

Green BR. 2011. Chloroplast genomes of photosynthetic eukaryotes. Plant J. 66:34–44. doi: 10.1111/j.1365-313X.2011.04541.x

Julou T, et al. 2005. Mixotrophy in orchids: insights from a comparative study of green individuals and nonphotosynthetic individuals of *Cephalanthera damasonium*. New Phytol. 166:639–653. doi: 10.1111/j.1469-8137.2005.01364.x

Katoh K, Misawa K, Kuma K, Miyata T. 2002. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. Nucleic Acids Res. 30:3059–3066. doi: 10.1093/nar/gkf436

Kellogg EA, Juliano ND. 1997. The structure and function of RuBisCO and their implications for systematic studies. Am J Bot. 84:413–428. doi: 10.2307/2446015

Kolpakov R, Bana G, Kucherov G. 2003. mreps: efficient and flexible detection of tandem repeats in DNA. Nucleic Acids Res. 31:3672–3678. doi: 10.1093/nar/gkg617

Li J, Wang S, Yu J, Wang L, Zhou S. 2013. A modified CTAB protocol for plant DNA extraction. Chin Bull Bot. 48:72–78.

Liebel HT, et al. 2010. C and N stable isotope signatures reveal constraints to nutritional modes in orchids from the Mediterranean and Macaronesia. Am J Bot. 97:903–912. doi: 10.3732/ajb.0900354.

Lin C-S, et al. 2015. The location and translocation of *ndh* genes of chloroplast origin in the Orchidaceae family. Sci Rep. 5:9040. doi: 10.1038/srep0

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

Lohse M, Drechsel O, Kahlau S, Bock R. 2013. Organellar GenomeDRAW—a suite of tools for generating physical maps of plastid and mitochondrial genomes and visualizing expression

data sets. Nucleic Acids Res. 41:W575–W581. doi: 10.1093/nar/gkt289
Luo R, et al. 2012. SOAPdenovo2: an empirically improved memory-efficient short-read de novo assembler. Gigascience 1. doi: 10.1186/2047-217x-1-18. p.1–6.
McNeal JR, Kuehl JV, Boore JL, de Pamphilis CW. 2007. Complete plastid genome sequences suggest strong selection for retention of photosynthetic genes in the parasitic plant genus *Cuscuta*. BMC Plant Biol. 7:57. doi: 10.1186/1471-2229-7-57
Merckx VSFT, et al. 2013a. Taxonomy and classification. In: Merckx VSFT, editor. Mycoheterotrophy: the biology of plants living on fungi. New York: Springer. p. 19–101.
Merkx VSFT, Mennes CB, Peay KG, Geml J. 2013b. Taxonomy and diversification. In: Merckx VSFT, editor. Mycoheterotrophy: the biology of plants living on fungi. New York: Springer. p. 215–244.
Neuhaus HE, Emes MJ. 2000. Nonphotosynthetic metabolism in plastids. Annu Rev Plant Physiol Plant Mol Biol. 51:111–140. doi: 10.1146/annurev.arplant.51.1.111
Nylander JAA, Ronquist F, Huelsenbeck JP, Nieves-Aldrey JL. 2004. Bayesian phylogenetic analysis of combined data. Syst Biol. 53:47–67. doi: 10.1080/10635150490264699
Patel RK, Jain M. 2012. NGS QC Toolkit: A toolkit for quality control of next generation sequencing data. PLoS One 7: doi: 10.1371/journal.pone.0030619
Peltier G, Cournac L. 2002. Chlororespiration. Annu Rev Plant Biol. 53:523–550. doi: 10.1146/annurev.arplant.53.100301.135242
Preiss K, Adam IKU, Gebauer G. 2010. Irradiance governs exploitation of fungi: fine-tuning of carbon gain by two partially myco-heterotrophic orchids. Proc R Soc B-Biol Sci. 277:1333–1336. doi: 10.1098/rspb.2009.1966
Pridgeon AM, Cribb PJ, Chase MW, Rasmussen FN. 2005. Genera Orchidacearum. Vol. 4. Epidendroideae (Part 1). Oxford, UK: Oxford University Press.
Race HL, Herrmann RG, Martin W. 1999. Why have organelles retained genomes? Trends Genet. 15:364–370. doi: 10.1016/s0168-9525(99)01766-7
Raskoti BB, et al. 2016. A phylogenetic analysis of molecular and morphological characters of *Herminium* (Orchidaceae, Orchideae): evolutionary relationships, taxonomy, and patterns of character evolution. Cladistics 32:198–210. doi: 10.1111/cla.12125
Ronquist F, et al. 2012. MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. Syst Biol. 61:539–542. doi: 10.1093/sysbio/sys029
Rousseau-Gueutin M, Ayliffe MA, Timmis JN. 2011. Conservation of plastid sequences in the plant nuclear genome for millions of years facilitates endosymbiotic evolution. Plant Physiol. 157:2181–2193. doi: 10.1104/pp.111.185074
Roy M, et al. 2013. Why do mixotrophic plants stay green? A comparison between green and achlorophyllous orchid individuals *in situ*. Ecol Monogr. 83:95–117. doi: 10.1890/11-2120.1
Roy M, et al. 2009. Two mycoheterotrophic orchids from Thailand tropical dipterocarpacean forests associate with a broad diversity of ectomycorrhizal fungi. BMC Biol. 7:51. doi: 10.1186/1741-7007-7-51
Kamikawa R, et al. 2015. Proposal of a twin arginine translocator system-mediated constraint against loss of ATP synthase genes from nonphotosynthetic plastid genomes. Mol Biol Evol. 32:2598–2604.
Sanderson MJ. 2003. r8s: inferring absolute rates of molecular evolution and divergence times in the absence of a molecular clock. Bioinformatics 19:301–302. doi: 10.1093/bioinformatics/19.2.301
Schelkunov MI, et al. 2015. Exploring the limits for reduction of plastid genomes: a case study of the mycoheterotrophic orchids *Epipogium aphyllum* and *Epipogium roseum*. Genome Biol Evol. 7:1179–1191. doi: 10.1093/gbe/evv019
Selosse M-A, Roy M. 2009. Green plants that feed on fungi: facts and questions about mixotrophy. Trends Plant Sci. 14:64–70. doi: 10.1016/j.tplants.2008.11.004
Stamatakis A. 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics 30:1312–1313. doi: 10.1093/bioinformatics/btu033
Stoeckel M, Meyer C, Gebauer G. 2011. The degree of mycoheterotrophic carbon gain in green, variegated and vegetative albino individuals of *Cephalanthera damasonium* is related to leaf chlorophyll concentrations. New Phytol. 189:790–796. doi: 10.1111/j.1469-8137.2010.03510.x
Vaidya G, Lohman DJ, Meier R. 2011. SequenceMatrix: concatenation software for the fast assembly of multi-gene datasets with character set and codon information. Cladistics 27:171–180. doi: 10.1111/j.1096-0031.2010.00329.x
Wertheim JO, Murrell B, Smith MD, Pond SLK, Scheffler K. 2015. RELAX: Detecting relaxed selection in a phylogenetic framework. Mol Biol Evol. 32:820–832. doi: 10.1093/molbev/msu400
Wicke S, et al. 2013. Mechanisms of functional and physical genome reduction in photosynthetic and nonphotosynthetic parasitic plants of the Broomrape family. Plant Cell 25:3711–3725. doi: 10.1105/tpc.113.113373
Wicke S, Schneeweiss GM, dePamphilis CW, Müller KF, Quandt D. 2011. The evolution of the plastid chromosome in land plants: gene content, gene order, gene function. Plant Mol Biol. 76:273–297. doi: 10.1007/s11103-011-9762-4
Wickett NJ, et al. 2011. Transcriptomes of the parasitic plant family Orobanchaceae reveal surprising conservation of chlorophyll synthesis. Curr Biol. 21:2098–2104. doi: 10.1016/j.cub.2011.11.011
Wolfe KH, Morden CW, Palmer JD. 1992. Function and evolution of a minimal plastid genome from a nonphotosynthetic parasitic plant. Proc Natl Acad Sci U S A. 89:10648–10652. doi: 10.1073/pnas.89.22.10648
Wu F-H, et al. 2010. Complete chloroplast genome of *Oncidium* Gower Ramsey and evaluation of molecular markers for identification and breeding in Oncidiinae. BMC Plant Biol. 10:68. doi: 10.1186/1471-2229-10-68
Wyman SK, Jansen RK, Boore JL. 2004. Automatic annotation of organellar genomes with DOGMA. Bioinformatics 20:3252–3255. doi: 10.1093/bioinformatics/bth352
Xiang X-G, et al. 2012. Phylogenetic placement of the enigmatic orchid genera *Thaia* and *Tangtsinia*: evidence from molecular and morphological characters. Taxon 61:45–54.
Xiang XG, et al. 2016. Biogeographical diversification of mainland Asian *Dendrobium* (Orchidaceae) and its implications for the historical dynamics of evergreen broad-leaved forests. J Biogeogr. doi:10.1111/jbi.12726.
Yang Z. 2007. PAML 4: phylogenetic analysis by maximum likelihood. Mol Biol Evol. 24:1586–1591. doi: 10.1093/molbev/msm088
Yang ZH. 1997. PAML: a program package for phylogenetic analysis by maximum likelihood. Comput Appl Biosci. 13:555–556.
Zerbino DR, Birney E. 2008. Velvet: algorithms for de novo short read assembly using de Bruijn graphs. Genome Res. 18:821–829. doi: 10.1101/gr.074492.107
Zimmer K, Meyer C, Gebauer G. 2008. The ectomycorrhizal specialist orchid *Corallorhiza trifida* is a partial myco-heterotroph. New Phytol. 178:395–400. doi: 10.1111/j.1469-8137.2007.02362.x

Associate editor: Shu-Miaw Chaw