

Sequencing and Analysis of Plastid Genome in Mycoheterotrophic Orchid *Neottia nidus-avis*

Maria D. Logacheva^{1,2,3,*}, Mikhail I. Schelkunov^{4,5}, and Aleksey A. Penin^{2,3,6}

¹Department of Evolutionary Biochemistry, A. N. Belozersky Institute of Physico-Chemical Biology, M. V. Lomonosov Moscow State University, Russia

²Evolutionary Genomics Laboratory, Faculty of Bioengineering and Bioinformatics, M. V. Lomonosov Moscow State University, Russia

³A. A. Kharkevich Institute for Information Transmission Problems, Russian Academy of Science, Moscow, Russia

⁴Center of Innovation and Technology of Biologically Active Compounds, Russian Academy of Sciences, Moscow, Russia

⁵V. A. Engelhardt Institute of Molecular Biology, Russian Academy of Sciences, Moscow, Russia

⁶Department of Genetics, Biological faculty, M. V. Lomonosov Moscow State University, Russia

*Corresponding author: E-mail: maria.log@gmail.com.

Accepted: 27 September 2011

Abstract

Plastids are the semiautonomous organelles that possess their own genome inherited from the cyanobacterial ancestor. The primary function of plastids is photosynthesis so the structure and evolution of plastid genomes are extensively studied in photosynthetic plants. In contrast, little is known about the plastomes of nonphotosynthetic species. In higher plants, plastid genome sequences are available for only three strictly nonphotosynthetic species, the liverwort *Aneura mirabilis* and two flowering plants, *Epifagus virginiana* and *Rhizanthella gardneri*. We report here the complete sequence of a plastid genome of nonphotosynthetic mycoheterotrophic orchid *Neottia nidus-avis*, determined using 454 pyrosequencing technology. It was found to be reduced in both genome size and gene content; this reduction is however not as drastic as in the other nonphotosynthetic orchid, *R. gardneri*. *Neottia* plastome lacks all genes encoding photosynthetic proteins, RNA polymerase subunits but retains most genes of translational apparatus. Those genes that are retained have an increased rate of both synonymous and nonsynonymous substitutions but do not exhibit relaxation of purifying selection either in *Neottia* or in *Rhizanthella*.

Key words: *Neottia nidus-avis*, Orchidaceae, plastid genome, mycoheterotrophy, gene loss, substitution rate.

Introduction

The information on chloroplast genome sequences is essential in different fields of plant biology: plant physiology, population genetics, phylogenetics, and evolution. As a consequence, the number of published chloroplast genome sequences greatly increased in past three years allowing an assessment of general trends of their evolution (reviewed in Wicke et al. 2011). In contrast, little is known about the structure of the plastid genome in nonphotosynthetic species. In higher plants, complete plastid genome sequences are available now for only three completely nonphotosynthetic species: two angiosperms—the parasitic plant *Epifagus virginiana* from Orobanchaceae (Wolfe et al. 1992) and mycoheterotrophic orchid, *Rhizanthella gardneri*

(Delannoy et al. 2011), and a mycoheterotrophic liverwort *Aneura mirabilis* (Wickett, Zhang, et al. 2008). While there is about 3,000 nonphotosynthetic plants representing more than ten families, this is obviously insufficient to infer general patterns of the evolution of plastome in the absence of photosynthetic activity. We have increased this set by sequencing the entire plastome of a mycoheterotrophic orchid, *Neottia nidus-avis*. *Neottia* belongs to the same family as *Rhizanthella* whose plastid genome was recently characterized (Delannoy et al. 2011) but the shift to heterotrophy occurred independently in these two lineages (Molvray et al. 2000). Thus, this information is useful for revealing both general features of nonphotosynthetic plant plastomes and for comparing the plastid genome structure under the

© The Author(s) 2011. 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/3.0/>), which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

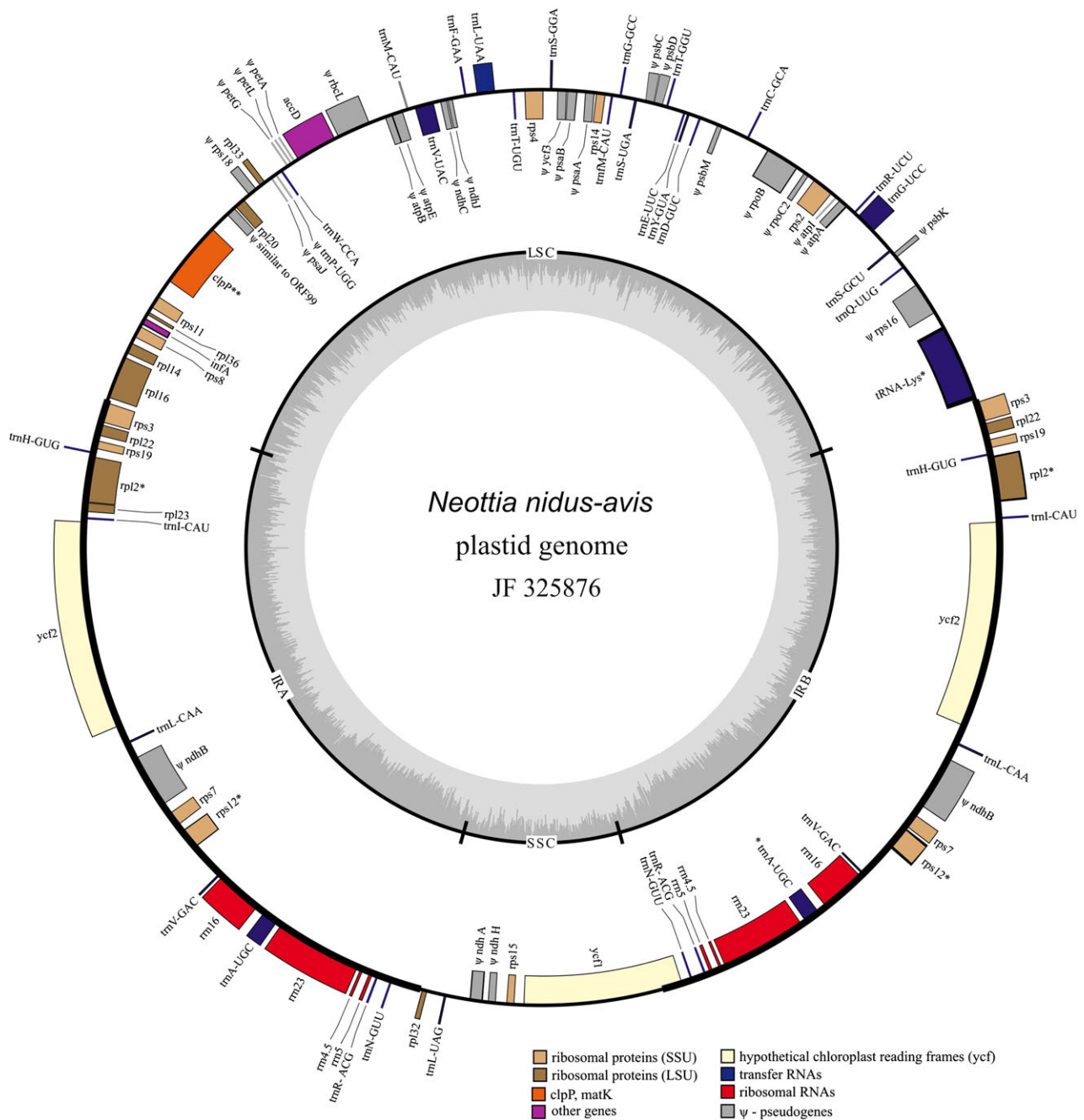


Fig. 1.—Circular map of the plastid genome of *Neottia nidus-avis*. Genes shown inside the circle are transcribed clockwise, those outside the circle are transcribed counterclockwise. Asterisks indicate intron-containing genes, dark gray bars inside the inner circle indicate guanine-cytosine content.

parallel loss of photosynthetic activity in a specific plant group, the family Orchidaceae.

The sequence of *Neottia* plastome was assembled from the partial genomic DNA sequence produced using high-throughput pyrosequencing technology (454 sequencing) complemented with Sanger sequencing. As expected, the plastid genome of *Neottia* is highly reduced in length (92 Kb compared with 146–149 in photosynthetic orchids)

and in gene content (fig. 1). All genes encoding photosystem I and II components are lost or pseudogenized. The same is true for the genes of the cytochrome *b₆f* components and photosystem I assembly proteins; *ccsA* (involved in c-type cytochrome synthesis) and *cemA* (chloroplast envelope membrane protein) genes are also lost. All *ndh* genes are completely lost or turned to pseudogenes. The latter does not however seem to be related to heterotrophic

way of life since it is characteristic for several photosynthetic angiosperms, including orchids (Chang et al. 2006; Wu et al. 2010; Blazier et al. 2011) and for gymnosperms (Braukmann et al. 2009). Also, *Neottia* plastome apparently lacks functional *matK* gene. It contains a region with high similarity to *matK* but its comparison with orthologous sequences from photosynthetic orchids reveals strong divergence of its 5' end (including the substitution in the start codon); this suggests that in *Neottia*, *matK* is a pseudogene. Despite this, the genes that are supposed to require *matK* protein activity for the splicing of their mRNAs (*rpl2*, *rps12*, *clpP*, *trnA-UGC*, *trnG-UCC*) have retained their introns. This is in contrast with the genus *Cuscuta* where *matK* loss observed in the species of subgenus *Grammica* is correlated with the loss of group Ila introns (McNeal et al. 2009). There are two possible explanations. First, we can imagine that *matK*-like region retained in *Neottia* is still functional—this is possible under the condition that alternative start codon is adopted or that multiple RNA editing events occur in the 5' end of this region. Second, the involvement of *matK* in the splicing of group Ila introns could be not as essential in orchids as it is *Cuscuta* (and, apparently, in most plants). This hypothesis is supported by the fact that *Rhizanthella*, another nonphotosynthetic orchid, totally lacks any *matK*-like regions but also retains group Ila introns and that *matK* is pseudogenized in *Corallorhiza trifida* (Freudenstein and Senyo 2008), a mycoheterotrophic species that partially retains photosynthetic activity.

The only class of genes that is unaffected by the reduction is the ribosomal RNA genes. All four rRNA genes characteristic for typical plant plastomes are present in *Neottia* and share very high similarity (96–99%) with their orthologs from other orchids. Genes encoding another essential component of the ribosome, the ribosomal proteins, are also mostly retained: *Neottia* plastome encodes a complete set of large subunit proteins genes (including *rpl23* and *rpl22* that are turned to pseudogenes in many photosynthetic species) and most small subunit proteins genes. The functionality of two small subunit proteins genes—*rps16* and *rps18*—is questionable. Though they are highly similar to the orthologous genes from photosynthetic orchids in silico translation reveals internal stop codons in both sequences. These stop codons are however in frame and thus are potential targets of RNA editing so additional experiments are required to confirm the nonfunctionality of *rps16* and *rps18*. In *Anthoceros formosae*, conversion of nonsense codons into sense is found in 52 genes, including *rps18* (Kugita et al. 2003). RNA editing system was shown to be active in *Rhizanthella* so it is presumably active in *Neottia* plastids too. Another possible target of RNA editing is the *rpl2* gene that has atypical start codon ACG. As for transfer RNA, we were able to find the sequences with high similarity for all tRNAs characteristic for plant plastomes with the exception of *trnL-GAU*. Most of them seem to be functional genes because

they share high similarity with their orthologs from photosynthetic orchids (overall similarity of tRNA sequences between *Neottia* and *Phalaenopsis* is 0.95) and have conserved secondary structure typical for tRNAs. *trnV-UAC* and *trnP-UGG* are putative pseudogenes since they differ from their orthologs by multiple substitutions and indels. *infA* gene, which encodes translation initiation factor 1 is also present. In contrast to translation apparatus that seems to be almost unaffected, genes of the transcription machinery—those encoding plastid RNA polymerase subunits—are lost (*rpoA*, *rpoC1*) or pseudogenized (*rpoB*, *rpoC2*). Two genes involved in plastid metabolism—*accD* and *clpP*—are retained, as well as two large ORF encoding proteins of the unknown function—*ycf1* and *ycf2*. The retention of *ycf1* is the most enigmatic since it is presumably pseudogenized in photosynthetic orchids *Oncidium* and *Phalaenopsis*. The regions with strong similarity to *ycf1* are present in their plastomes but their in silico translation reveals multiple stop codons due to nontriplet indels. Pseudogenization of *ycf1*, as well as another large plastid ORF, *ycf2* occurs in grasses where a region similar to *ycf1* is present but highly reduced due to multiple deletions (Hiratsuka et al. 1989). The structure of *ycf1* in *Phalaenopsis* and *Oncidium* may represent a first stage of gene degradation. Another explanation of the frameshifts might be related to sequencing errors because *ycf1* is rich in homopolymer regions, which are prone to errors. Broader survey of *ycf1* integrity in orchids is required to see if this gene is indeed nonfunctional in either lineage of orchids.

Despite about one-third reduction in length, the overall structure of *Neottia* plastid genome is conserved and collinear with that of photosynthetic orchids. The only alteration of the gene order is the position of inverted repeat-large single copy border that resides in the *rps3-rpl16/rps3-trnK* spacer in *Neottia*. The expansions of the IR are well documented in different lineages of angiosperms; in particular in orchids IR was found to include complete sequence of *rps19*, *trnH*, and a part of *rpl22* (Wang et al. 2008). In *Neottia*, further expansion of the IR is observed—it includes *rpl22* and *rps3*. Such structure of J_{IR-LSC} is by now unique however it is possible that increased taxon sampling will reveal similar structures in other orchids, either photosynthetic or not.

To assess the selection constraint on plastid genes in *Neottia*, we performed the analysis of synonymous and nonsynonymous substitutions in protein-coding genes. *Rhizanthella*, another nonphotosynthetic orchid was also included in the analysis. The examination of dN/dS ratio for individual genes that are shared between *Neottia*, *Rhizanthella*, and photosynthetic orchids demonstrates that most of them are evolving under purifying selection in both nonphotosynthetic species (fig. 2). The only *Neottia* gene where dN/dS is greater than 1 is *rpl23*. The sequence of this gene is highly conserved between *Phalaenopsis* and *Oncidium*, differing by single substitution (nonsynonymous). *rpl23* is

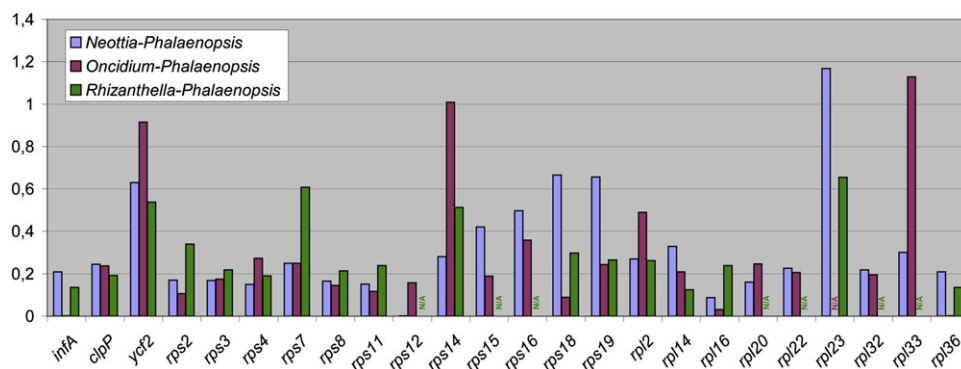


FIG. 2.—Pairwise dN/dS (the ratio of the number of nonsynonymous substitutions per nonsynonymous site to the number of synonymous substitutions per synonymous site) for all genes shared between photosynthetic orchids and *Neottia* and/or *Rhizanthella*. N/A indicates the cases where it was not possible to estimate the ratio due to the absence of synonymous substitutions or the absence of gene in the plastid genome.

a pseudogene in several photosynthetic plants, for example, Caryophyllales (Logacheva et al. 2008). High dN/dS in this gene may indicate on the relaxation of selection constraint as initial stage of pseudogenization. However, in *rps14* and *rpl33*, dN/dS is found to be increased in photosynthetic species. Such increase may be caused by the unreliability of dN/dS estimates for short genes (Haddrill et al. 2007) and thus not be biologically relevant. To get a generalized view on substitution rate in *Neottia* and *Rhizanthella*, we calculated substitution values for all genes combined in one sequence. This demonstrated that dN/dS ratios are very similar between two photosynthetic species and between photosynthetic and nonphotosynthetic ones. In contrast, the number of both synonymous and nonsynonymous substitutions differs greatly being about 2.5 times much in *Neottia* and 5 times in *Rhizanthella* (fig. 3). The apparent effect of higher substitution rate is often observed in annual plants when compared with perennials (e.g., Yue et al. 2010); however, both *Neottia* and *Rhizanthella* are perennials as well as *Phalaenopsis* and *Oncidium*. This suggests that this effect is related to their heterotrophic way of life and that plastid genes in nonphotosynthetic orchids have higher mutation rate but retain their functionality.

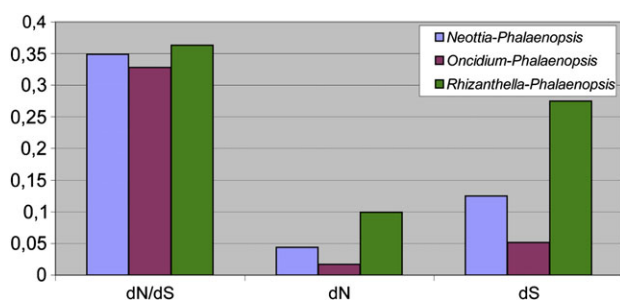


FIG. 3.— dN/dS ratio and dN and dS values for all genes shared between *Neottia*, *Rhizanthella*, *Oncidium*, and *Phalaenopsis*, combined in one sequence.

The broader comparison that includes all nonphotosynthetic higher plant plastomes—two orchids, parasitic dicot *Epifagus* and liverwort *Aneura*—reveals both parallelisms and dissimilarities in their structure. The plastome sizes differ almost twice—from about 108 Kb in *Aneura* to 59,190 bp in *Rhizanthella*. The degree of gene loss and pseudogenization also differs—in *Aneura*, the plastome is much less affected than in angiosperms. It retains all transfer RNA, RNA polymerase, and ribosomal protein genes and also many photosynthesis-related genes (Wickett, Zhang, et al. 2008) being similar in this respect to the species of *Cuscuta* that retain photosynthetic activity (Funk et al. 2007; McNeal et al. 2007). In strictly nonphotosynthetic angiosperms not a single photosynthesis-related gene is retained as intact reading frame, even in *Neottia* that is characterized by the least reduction in both size and gene content and is only 16 Kb less than in *Aneura*. In general, the structure and gene content of plastome are more similar within nonphotosynthetic angiosperms than in any of them compared with *Aneura*. It is well known that the activity of the plastid is not provided by plastid genome only but also is highly dependent on and coordinated with that of nuclear genome (reviewed in Taylor 1989 and Woodson and Chory 2008). The results from an ongoing genomic project of a liverwort, *Marchantia polymorpha*, indicate that the organization of the nuclear genomes in angiosperms and in liverworts differs greatly; in particular, the diversity of gene families is lower in *Marchantia* (Yamato et al. 2011). This and other differences in nuclear genome organization can account for the difference in the structure of plastid genomes in nonphotosynthetic angiosperms and nonphotosynthetic liverworts. Other possible explanation is the time past from the moment of switch to heterotrophy. Though there are no direct (paleontological) observations on this subject, we suggest that heterotrophy is the most recent in *A. mirabilis*—since another species of the genus *Aneura* are photosynthetic (Wickett, Fan, et al. 2008) and the most ancient in *Epifagus* which shares the holoparasitic way of life with related genera *Conopholis*

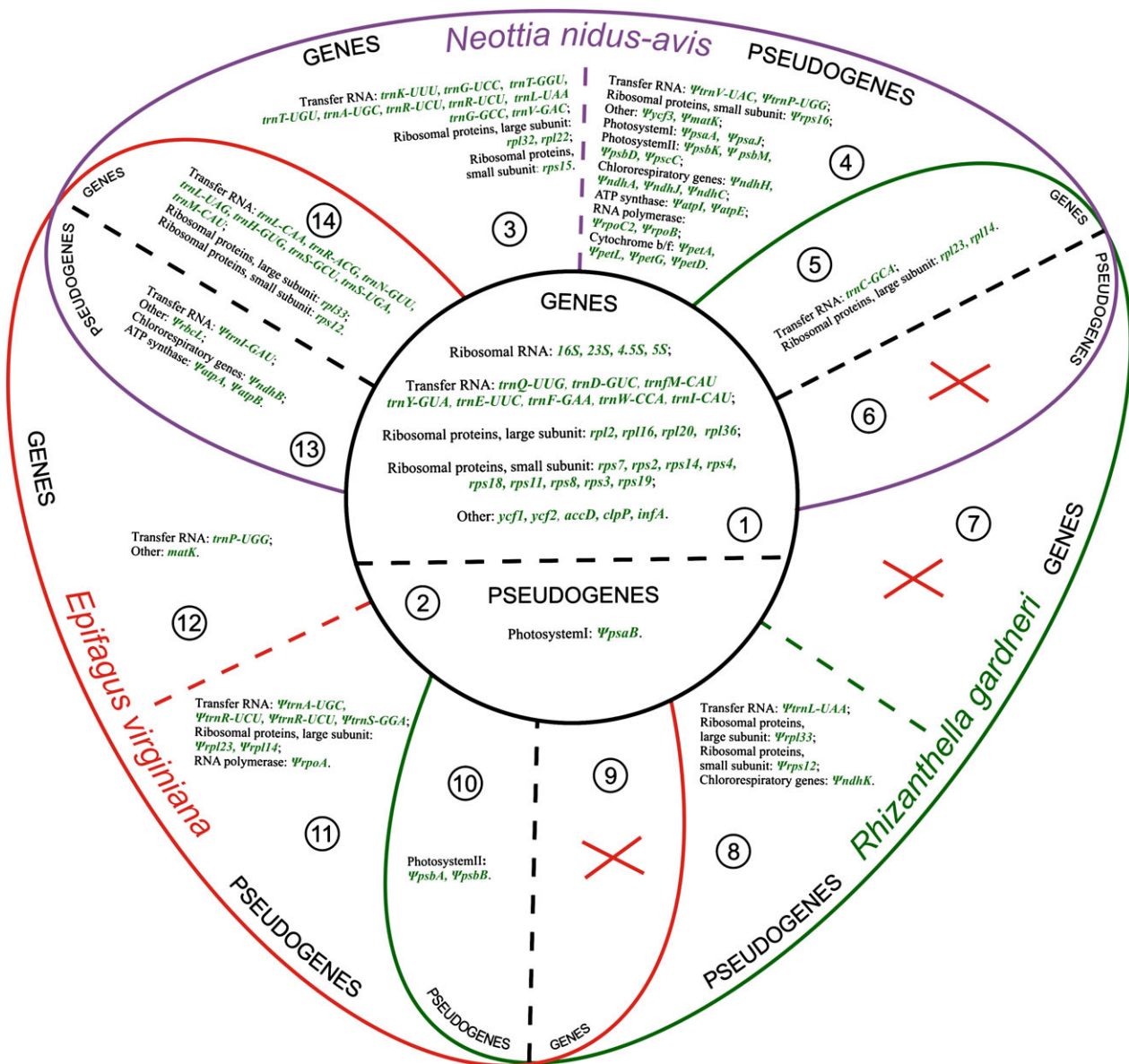


Fig. 4.—Gene and pseudogene content in the plastomes of nonphotosynthetic angiosperms, numbers indicate different groups of genes/pseudogenes according to their presence in the plastomes: (1) genes that are present in all plastid genomes of nonphotosynthetic angiosperms, (2) pseudogenes that are present in all plastid genomes of nonphotosynthetic angiosperms, (3) genes that are unique for *Neottia*, (4) pseudogenes that are unique to *Neottia*, (5) genes present only in *Neottia* and *Rhizanthella*, (6) pseudogenes present only in *Neottia* and *Rhizanthella* (no), (7) genes unique to *Rhizanthella* (no), (8) pseudogenes unique to *Rhizanthella*, (9) genes present only in *Rhizanthella* and *Epifagus* (no), (10) pseudogenes present only in *Rhizanthella* and *Epifagus*, (11) pseudogenes unique to *Epifagus*, (12) genes unique to *Epifagus*, (13) pseudogenes present only in *Epifagus* and *Neottia*, and (14) genes present only in *Epifagus* and *Neottia*.

and *Orobanche* (dePamphilis et al. 1997). *Neottia* and *Rhizanthella* represent genera that include only nonphotosynthetic species but their most closely related genera are completely or partially photosynthetic. Moreover, in some systematic treatments, *Neottia* is merged with photosynthetic genus *Listera* (Chase et al. 2003) that suggests recent switch to heterotrophy, similar to *Aneura*. In terms of gene content, *Neottia* is more similar to *Epifagus* than to *Rhizanthella* (fig. 4) or *Aneura*. Phylogenetically *Neottia* is close to

Rhizanthella thus the dissimilarity of their plastid genomes is unlikely to be related to major dissimilarities in the organization of their nuclear genomes but rather to lineage-specific variations in rate and pattern of plastid genome evolution. This phenomenon is well known with regard to single nucleotide substitutions (reviewed in Muse 2000), and it is likely to be applicable to length mutations as well.

Among about 100 of genes encoded in typical plastid genome, only 29 are shared in all the plastid genomes of

nonphotosynthetic plants. This core set of genes constitutes of ribosomal and transfer RNA and ribosomal protein genes, translation initiation factor (*infA*), protease subunit (*clpP*), acetyl-CoA carboxylase subunit (*accD*) genes, and two large ORF of unknown function—*ycf1* and *ycf2* (fig. 4). This is in contrast with the situation observed in experimental systems of transition to heterotrophy. Calli cultured for a long time on sucrose-rich media also undergo multiple deletions in plastid DNA and extensive gene losses. However, their size and location drastically differ in different calli samples and do not seem to converge to a gene set shared by all samples. Moreover, the loss of circular structure typical for plastid DNAs, is reported (Harada et al. 1992; Abe et al. 2002). The stability of plastid genomes of nonphotosynthetic plants—the conservation of typical quadripartite structure, the existence of shared gene set, and the evidence of purifying selection acting on these genes—suggests that the evolution of plastid genomes in nonphotosynthetic plants is dictated by the constraints that are obviously different from those acting on photosynthetic plant plastomes but are not less strong.

Orchidaceae is a perfect model system for deeper study of the evolution of nonphotosynthetic plants' plastomes because it contains many cases of independent transition to nonphotosynthetic lifestyle (Molvray et al. 2000), from rather ancient occurring at the level of tribe to extremely recent occurring in certain individuals within the population (Tranchida-Lombardo et al. 2010). Moreover, many photosynthetic orchids are able to use fungi as secondary carbon source being mixotrophic rather than fully autotrophic (Bidartondo et al. 2004; Abadie et al. 2006). Mixotrophy is thought to be the preadaptation that mediates the transition to completely heterotrophic lifestyle (Selosse and Roy 2009). It is interesting to infer if any of the changes in the plastome characteristic for strictly nonphotosynthetic plants (e.g., gene loss, increase of substitution rate) are observed in mixotrophic species. We expect that further sampling of the plastid genome sequences from the species of Orchidaceae representing different stages and different times of transition to mycoheterotrophy will provide valuable information about the evolution of plastomes in nonphotosynthetic plants.

Materials and Methods

Total DNA was extracted from the above-ground part of three *Neottia* plants using NucleoSpin Plant II kit (Macherey-Nagel, Germany). DNA (10 μ g) was used for sequencing with Roche Genome Sequencer FLX system using the Titanium kit (454 Life Sciences). Sequencing was performed at the University of Illinois at Urbana-Champaign, W.M. Keck Center for Comparative and Functional Genomics. The sample was run on a half of a picotiter plate. The sequencing resulted in 590640 reads with an average read length of

528 bp. The output from the sequencing system in Standard Flowgram Format was converted to FASTA. Then a BLAST database was made from the resulting FASTA file and a plastome of *Oncidium*, the photosynthetic orchid was queried against it. All reads with e-value lower than 10^{-10} were used for de novo assembly using MIRA assembler ver. 3.2.0 (Chevreux et al. 1999). The assembly resulted in 23 contigs with length more than 1,000 nucleotides. Two largest contigs (11,141 and 49,963 nucleotides) were used as a basis for the generation of draft version of the *Neottia* plastome. They were aligned with *Oncidium* plastome using the whole VISTA genome alignment tool (Frazer et al. 2004). Then Sanger sequencing was used to fill the gaps. To check the accuracy of the assembly and to correct possible 454 sequencing errors associated with homopolymer runs, several regions were sequenced by Sanger sequencing using the primers designed on the base of the assembly (supplementary table S1, Supplementary Material online). Average coverage for regions derived from 454 assembly is assessed as 29.8 \times ; for regions sequenced by Sanger sequencing it is 2 \times . Polymerase chain reaction (PCR) amplification was performed on Biometra T300 thermal cycler using Encyclo PCR kit (Evrogen, Russia). PCR conditions were as follows: initial denaturation 3 min at 94 $^{\circ}$ C, then 35 cycles of 15 s at 94 $^{\circ}$ C, 25 s at 59 $^{\circ}$ C, and 1–5 min (depending on the expected length of the product) at 72 $^{\circ}$ C. Sanger sequencing was performed in the interinstitutional sequencing center at Engelhardt Institute of Molecular Biology (Moscow, Russia) using ABI PRISM BigDye Terminator kit v. 3.1 with following analysis on ABI PRISM 3730 genetic analyzer (Applied Biosystems). Initial annotation was produced using DOGMA (Wyman et al. 2004). Then manual correction and adjustment, that included alignment of every *Oncidium* and *Phalaenopsis* gene with *Neottia* plastome sequence was performed. The regions with similarity to known protein-coding genes but lacking intact ORF were classified as pseudogenes. To detect tRNAs pseudogenes, each tRNA-like sequence was analyzed using tRNAscan-SE (Lowe and Eddy 1997) and by comparison with its putative orthologs from *Oncidium* and *Phalaenopsis*. Those sequences that lack typical tRNA folding and/or differ from their orthologs by multiple indels or substitutions were considered to be pseudogenes. The map of *Neottia* plastome was visualized using OGDRAW online tool (Lohse et al. 2007) with further manual correction. Assembled and corrected sequence of *Neottia* plastome was deposited in the GenBank under accession number JF325876.

For dN/dS calculation, all protein-coding plastid genes that are shared between *Neottia*, *Rhizanthella*, *Oncidium*, and *Phalaenopsis* were included in the analysis. Alignment was performed using ClustalW (Thompson et al. 1994); dN/dS ratios were calculated using codeml program from the package PAML 4.3 (Yang 2007). In the alignments of several genes, nontriplet indels or in-frame stop codons were found

near 3'-end. In these cases, only the part of sequence before presumable indel or stop codon was used for dN/dS calculation.

Supplementary Material

Supplementary table S1 is available at *Genome Biology and Evolution* online (<http://www.gbe.oxfordjournals.org/>).

Acknowledgments

We would like to thank Dmitry Sokoloff and Margarita Remizova (Moscow State University, Russia) for providing plant material for the study; Alvaro Hernandez (University of Illinois at Urbana-Champaign) for 454 sequencing; Viktoria Shtratnikova (Center of Innovation and Technology of Biologically Active Compounds, Russia); Georgiy Bazykin (Institute for Information Transmission Problems, Russia) for helpful discussion; and three anonymous reviewers for the suggestions and comments on the manuscript. This work was supported by the Russian Foundation for Basic Research (project no. 11-04-02031) and Ministry of Education and Science of Russian Federation (state contract no.1054 and project 11.G34.31.0008).

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.* 84(9):1462–1477.
- Abe T, li N, Togashi A, Sasahara T. 2002. Large deletions in chloroplast DNA of rice calli after long-term culture. *J Plant Physiol.* 159:917–923.
- Bidartondo MI, Burghardt B, Gebauer G, Bruns TD, Read DJ. 2004. Changing partners in the dark: isotopic and molecular evidence of ectomycorrhizal liaisons between forest orchids and trees. *Proc Biol Sci.* 271(1550):1799–1806.
- Blazier CJ, Guisinger MM, Jansen RK. 2011. Recent loss of plastid-encoded *ndh* genes within *Erodium* (Geraniaceae). *Plant Mol Biol.* 76(3–5):263–272.
- Braukmann TW, Kuzmina M, Stefanović S. 2009. Loss of all plastid *ndh* genes in Gnetales and conifers: extent and evolutionary significance for the seed plant phylogeny. *Curr Genet.* 55(3):323–337.
- Chang CC, 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.
- Chase MW, Freudenstein JF, Cameron KM. 2003. DNA data and Orchidaceae systematics: a new phylogenetic classification. In: Dixon KW, Kell SP, Barrett RL, Cribb PJ, editors. *Orchid conservation*. Kota Kinabalu (Malaysia): Natural History Publications. p. 69–89.
- Chevreur B, Wetter T, Suhai S. 1999. Genome sequence assembly using trace signals and additional sequence information. *Computer Science and Biology: Proceedings of the German Conference on Bioinformatics*. Vol. 99; 1999 Oct 4–6; Hanover (Germany).
- Delannoy E, Fujii S, des Francs CC, Brundrett M, Small I. 2011. Rampant gene loss in the underground orchid *Rhizanthella gardneri* highlights evolutionary constraints on plastid genomes. *Mol Biol Evol.* 28(7):2077–2086.
- dePamphilis CW, Young ND, Wolfe AD. 1997. Evolution of plastid gene *rps2* in a lineage of hemiparasitic and holoparasitic plants: many losses of photosynthesis and complex patterns of rate variation. *Proc Natl Acad Sci U S A.* 94(14):7367–7372.
- Frazer KA, Pachter L, Poliakov A, Rubin EM, Dubchak I. 2004. VISTA: computational tools for comparative genomics. *Nucleic Acids Res.* 32:W273–W279.
- Freudenstein JV, Senyo DM. 2008. Relationships and evolution of *matK* in a group of leafless orchids (*Corallorhiza* and *Corallorhizinae*; Orchidaceae: Epidendroideae). *Am J Bot.* 95:498–505.
- Funk HT, Berg S, Krupinska K, Maier UG, Krause K. 2007. Complete DNA sequences of the plastid genomes of two parasitic flowering plant species, *Cuscuta reflexa* and *Cuscuta gronovii*. *BMC Plant Biol.* 7:45.
- Hadrill PR, Halligan DL, Tomaras D, Charlesworth B. 2007. Reduced efficacy of selection in regions of the *Drosophila* genome that lack crossing over. *Genome Biol.* 8(2):R18.
- Harada T, Ishikawa R, Niizeki M, Saito K. 1992. Pollen-derived rice calli that have large deletions in plastid DNA do not require protein synthesis in plastids for growth. *Mol Gen Genet.* 233(1–2):145–150.
- Hiratsuka J, et al. 1989. The complete sequence of the rice (*Oryza sativa*) chloroplast genome: intermolecular recombination between distinct tRNA genes accounts for a major plastid DNA inversion during the evolution of the cereals. *Mol Gen Genet.* 217(2–3):185–194.
- Kugita M, Yamamoto Y, Fujikawa T, Matsumoto T, Yoshinaga K. 2003. RNA editing in hornwort chloroplasts makes more than half the genes functional. *Nucleic Acids Res.* 31(9):2417–2423.
- Lohse M, Drechsel O, Bock R. 2007. OrganellarGenomeDRAW (OGDRAW): a tool for the easy generation of high-quality custom graphical maps of plastid and mitochondrial genomes. *Curr Genet.* 52:267–274.
- Logacheva MD, Samigullin TH, Dhingra A, Penin AA. 2008. Comparative chloroplast genomics and phylogenetics of *Fagopyrum esculentum* ssp. *ancestrale*—a wild ancestor of cultivated buckwheat. *BMC Plant Biol.* 8:59.
- Lowe TM, Eddy SR. 1997. tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. *Nucleic Acids Res.* 25:955–964.
- 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.
- McNeal JR, Kuehl JV, Boore JL, Leebens-Mack J, dePamphilis CW. 2009. Parallel loss of plastid introns and their maturase in the genus *Cuscuta*. *PLoS One* 4(6):e5982.
- Molvray M, Kores PJ, Chase MW. 2000. Polyphyly of mycoheterotrophic orchids and functional influences of floral and molecular characters. In: Wilson KL, Morrison DA, editors. *Monocots: systematics and evolution*. Collingwood (Australia): CSIRO Publishing. p. 441–448.
- Muse SV. 2000. Examining rates and patterns of nucleotide substitution in plants. *Plant Mol Biol.* 42 (1): 25–43.
- Selosse MA, Roy M. 2009. Green plants that feed on fungi: facts and questions about mixotrophy. *Trends Plant Sci.* 14(2):64–70.
- Taylor WC. 1989. Regulatory interactions between nuclear and plastid genomes. *Annu Rev Plant Physiol Plant Mol Biol.* 40:211–233.
- Thompson JD, Higgins DG, Gibson TJ. 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* 22(22):4673–4680.

- Tranchida-Lombardo V, et al. 2010. Spatial repartition and genetic relationship of green and albino individuals in mixed populations of *Cephalanthera* orchids. *Plant Biol (Stuttg)*. 12(4):659–667.
- Wang RJ, et al. 2008. Dynamics and evolution of the inverted repeat-large single copy junctions in the chloroplast genomes of monocots. *BMC Evol Biol*. 8:36.
- Wicke S, Schneeweiss GM, dePamphilis CW, Müller KF, Quandt D. 2011. Evolution of the plastid chromosome in land plants: gene content, gene order, gene function. *Plant Mol Biol*. 76(3–5):273–297.
- Wickett NJ, Fan Y, Lewis PO, Goffinet B. 2008. Distribution and evolution of pseudogenes, gene losses, and a gene rearrangement in the plastid genome of the nonphotosynthetic liverwort, *Aneura mirabilis* (Metzgeriales, Jungermanniopsida). *J Mol Evol*. 67(1):111–122.
- Wickett NJ, et al. 2008. Functional gene losses occur with minimal size reduction in the plastid genome of the parasitic liverwort *Aneura mirabilis*. *Mol Biol Evol*. 25(2):393–401.
- 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(22):10648–10652.
- Woodson JD, Chory J. 2008. Coordination of gene expression between organellar and nuclear genomes. *Nat Rev Genet*. 9(5):383–395.
- Wu FH, 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.
- Wyman SK, Jansen RK, Boore JL. 2004. Automatic annotation of organellar genomes with DOGMA. *Bioinformatics*. 20:3252–3255.
- Yamato KT, Ishizaki K, Kohchi T, Floyd S, Bowman J. 2011. The genome of *Marchantia polymorpha* L. Abstracts of the XVIII International Botanical Congress; 2011 Jul 23–30; Melbourne (Australia). p. 107. Available from: http://www.abc2011.com/downloads/IBC2011_Abstract_Book.pdf.
- Yang Z. 2007. PAML 4: phylogenetic analysis by maximum likelihood. *Mol Biol Evol*. 24:1586–1591.
- Yue JX, et al. 2010. Genome-wide investigation reveals high evolutionary rates in annual model plants. *BMC Plant Biol*. 10:242.

Associate editor: Shu-Miaw Chaw