Extensive plastome reduction and loss of photosynthesis genes in *Diphelypaea coccinea*, a holoparasitic plant of the family Orobanchaceae

Eugeny V. Gruzdev[1,2], Vitaly V. Kadnikov[1], Alexey V. Beletsky[1], Andrey V. Mardanov[1] and Nikolai V. Ravin[1,2]

[1] Institute of Bioengineering, Research Center of Biotechnology of the Russian Academy of Sciences, Moscow, Russia
[2] Moscow State University, Moscow, Russia

ABSTRACT

Background. Parasitic plants have the ability to obtain nutrients from their hosts and are less dependent on their own photosynthesis or completely lose this capacity. The reduction in plastid genome size and gene content in parasitic plants predominantly results from loss of photosynthetic genes. Plants from the family Orobanchaceae are used as models for studying plastid genome evolution in the transition from an autotrophic to parasitic lifestyle. *Diphelypaea* is a poorly studied genus of the Orobanchaceae, comprising two species of non-photosynthetic root holoparasites. In this study, we sequenced the plastid genome of *Diphelypaea coccinea* and compared it with other Orobanchaceae, to elucidate patterns of plastid genome evolution. In addition, we used plastid genome data to define the phylogenetic position of *Diphelypaea* spp.

Methods. The complete nucleotide sequence of the plastid genome of *D. coccinea* was obtained from total plant DNA, using pyrosequencing technology.

Results. The *D. coccinea* plastome is only 66,616 bp in length, and is highly rearranged; however, it retains a quadripartite structure. It contains only four rRNA genes, 25 tRNA genes and 25 protein-coding genes, being one of the most highly reduced plastomes among the parasitic Orobanchaceae. All genes related to photosynthesis, including the ATP synthase genes, had been lost, whereas most housekeeping genes remain intact. The plastome contains two divergent, but probably intact *clpP* genes. Intron loss had occurred in some protein-coding and tRNA genes. Phylogenetic analysis yielded a fully resolved tree for the Orobanchaceae, with *Diphelypaea* being a sister group to *Orobanche* sect. *Orobanche*.

Subjects Genomics, Plant Science
Keywords Parasitic plant, Plastid genome, Orobanchaceae, *Diphelypaea*

INTRODUCTION

About 1% of all angiosperm species can parasitize other flowering plants or mycorrhizal fungi (*Nickrent et al., 1997*). Facultative or obligatory hemiparasites still carry out photosynthesis to some extent, while holoparasites have completely lost the ability

to perform photosynthesis, and obtain nutrients from their host. Transition from an autotrophic to heterotrophic lifestyle is associated with a relaxation of selection pressure on photosynthesis-related genes, both in the nuclear and the plastid genomes. Most evident is a functional and physical reduction of the plastid genome (plastome), which correlates with a loss of genes encoding the photosynthetic machinery and related functions, increased substitution rates, and structural rearrangements (*DePamphilis & Palmer, 1990*; *Wolfe, Morden & Palmer, 1992*; *Barrett et al., 2014*). Since the plastome contains both photosynthesis-related and housekeeping genes, gene loss is not random and follows a particular pattern (*Barrett & Davis, 2012*; *Barrett et al., 2014*; *Graham, Lam & Merckx, 2017*). The NAD(P)H dehydrogenase (*ndh*) genes are usually lost first, followed by photosynthesis-related genes (*psa*, *psb*, *pet*, *rbcL* and *atp*) and plastid-encoded RNA polymerase. Housekeeping genes encoding rRNAs, ribosomal proteins and tRNAs are the last to be lost. Extensive studies of parasitic plants have revealed different levels of plastome degradation and their correlation with the types of parasitism, ranging from minimal in hemiparasitic members of Orobanchaceae (*Wicke et al., 2013*) and Viscaceae (*Petersen, Cuenca & Seberg, 2015*) to extreme in some holoparasitic species, such as *Pilostyles aethiopica* (*Bellot & Renner, 2015*), and even possibly complete loss of the plastome, as in *Rafflesia lagascae* (*Molina et al., 2014*).

The broomrape family, Orobanchaceae, is an ideal lineage to study plastid genome evolution in the course of transition to heterotrophy, since it comprises about 2000 hemiparasitic and holoparasitic species (*Bennett & Mathews, 2006*; *McNeal et al., 2013*), along with three autotrophic genera, *Lindenbergia*, *Rehmannia* and *Triaenophora*. Within Orobanchaceae, holoparasitism has evolved independently at least three times (*Young, Steiner & dePamphilis, 1999*; *McNeal et al., 2013*). Complete sequences of plastid genomes have been reported for *Lindenbergia philippensis* (*Wicke et al., 2013*), *Triaenophora shennongjiaensis* (*Xia & Wen, 2018*), six *Rehmannia* species (*Zeng et al., 2017*), and photosynthetic hemiparasitic species of the genera *Schwalbea* (*Wicke et al., 2013*), *Pedicularis* (*Cho et al., 2018*), *Castilleja* (*Fan et al., 2016*), *Aureolaria*, *Buchnera*, and *Striga* (*Frailey et al., 2018*), as well as for numerous holoparasites, including *Aphyllon* (*Myzorrhiza*), *Epifagus*, *Boulardia*, *Cistanche*, *Conopholis*, *Orobanche*, *Phelipanche* and *Lathraea* (*Wicke et al., 2013*; *Samigullin et al., 2016*; *Schneider et al., 2018*). Analysis of these plastid genomes has allowed reconstruction of the history of gene loss and genome reconfiguration in this family, in the course of transition to a holoparasitic lifestyle, revealing a limited set of commonly retained genes (*Wicke et al., 2013*; *Wicke et al., 2016*).

The genus *Diphelypaea* (*Nicolson, 1975*), also known as *Phelypaea*, is phylogenetically close to *Orobanche* sect. *Orobanche* (*Schneeweiss et al., 2004a*) and comprises two species, *Diphelypaea coccinea* (M.Bieb.) Nicolson and *Diphelypaea tournefortii* (Desf.) Nicolson, which occur in the Caucasus, Crimea and Western Asia (Turkey and Iran). *D. coccinea* is an achlorophyllous obligately parasitic perennial herbaceous plant up to 30–50 cm in height. The stem is unbranched, red to reddish brown, ending in a single flower of a bright red color (Fig. 1). *D. coccinea* parasitizes the roots of plants of the genus *Psephellus*. In order to further explore plastome evolution in the course of transition to holoparasitism, and to clarify the phylogenetic position of *Diphelypaea*, we determined the complete sequence of

Figure 1 **General view of *Diphelypaea coccinea* plants.** Source credit: Vitaly V. Kadnikov.

Full-size 🖼 DOI: 10.7717/peerj.7830/fig-1

the plastid genome of *D. coccinea* and compared it to previously sequenced plastomes of Orobanchaceae.

MATERIALS & METHODS

DNA extraction, genome sequencing and sequence assembly

The above ground portion of a single *D. coccinea* plant growing in a mountain region near the town of Kislovodsk, North Caucasus, Russia (43°50′59.8″N, 42°38′38.7″E) was used for the extraction of total genomic DNA, by a CTAB-NaCl method (*Murray & Thompson, 1980*). The extracted DNA was sheared using a g-TUBE device (Covaris Ltd., Brighton, UK) to obtain an average fragment size of 8 kb. The sheared DNA was then electrophoresed on a 0.5% agarose gel, and a portion of gel containing fragments of 6–9 kb was excised. The

DNA was purified with a QIAquick Gel Extraction Kit (Qiagen, Hilden, Germany), and then used to prepare a long paired-end library according to the manufacturer's instructions (Roche, Risch-Rotkreuz, Switzerland). The library was sequenced with a Roche GS FLX Genome Sequencer, using the Titanium XL+ protocol. About 329 Mb of sequence, with an average read length of 414 nts was generated. *De novo* assembly was performed with a Newbler Assembler v.2.9 (454 Life Sciences, Branford, CT, USA) with default settings, which yielded six plastid DNA contigs with an average coverage of 56-fold, ordered in a single scaffold. These contigs were identified based on sequence similarity to plastid genomes of Orobanchaceae and high coverage. The complete plastid genome sequence was obtained through generation of appropriate polymerase chain reaction (PCR) fragments spanning the junctions of the contigs and their sequencing by the Sanger method, using an ABI PRISM 3730 analyzer (Applied Biosystems, Waltham, MA, USA). The list of primers is available in Table S1. Reads spanning junctions between single copy regions and inverted repeats were used to infer contiguous sequences. To verify the correct assembly of the reconstructed plastid genome, raw reads were mapped against the reconstructed sequence with GS Reference Mapper (454 Life Sciences, Branford, CT, USA).

The raw reads were deposited in the Sequence Read Archive (SRA) under the accession number SRR9665263. The sequence of the plastid genome of *D. coccinea* was submitted to GenBank under accession number MK922354.

Plastid genome annotation and analysis tools

Plastid genome annotation was performed using the Dual Organellar GenoMe Annotator (DOGMA; *Wyman, Jansen & Boore, 2004*), with further manual correction using similarity searches against previously annotated plastid genomes. The tRNAscan-SE server was also used to locate tRNA genes (*Lowe & Chan, 2016*). A circular map of the plastome was drawn using OrganellarGenomeDRAW software (*Lohse et al., 2013*).

Phylogenetic analysis was performed using concatenated nucleotide sequences of 17 conserved protein-coding genes (*matK*, *rpl14*, *rpl16*, *rpl2*, *rpl20*, *rpl33*, *rpl36*, *rps11*, *rps12*, *rps14*, *rps18*, *rps19*, *rps2*, *rps3*, *rps4*, *rps7*, *rps8*) from plastid genomes of 30 species of Orobanchaceae: *Aphyllon uniflorum* var. uniflorum (MH580290), *Aphyllon fasciculatum* (MH580292.1), *Aphyllon epigalium* subsp. epigalium (MH050785), *Aphyllon californicum* (syn. *Myzorrhiza californica*, NC_025651), *Boulardia latisquama* (HG514460), *Castilleja paramensis* (KT959111), *Cistanche deserticola* (KC128846), *Cistanche phelypaea* (HG515538), *Conopholis america* (HG514459), *Epifagus virginiana* (M81884), *Lathraea squamaria* (KM652488), *Lindenbergia philippensis* (HG530133), *Orobanche rapum-genistae* (KT387725), *Orobanche californica* (HG515539.2), *Orobanche cernua* var. cumana (KT387722), *Orobanche crenata* (HG515537), *Orobanche pancicii* (KT387724), *Orobanche austrohispanica* (KT387721), *Orobanche densiflora* (KT387723), *Orobanche gracilis* (HG803179), *Pedicularis ishidoyana* (KU170194), *Pedicularis hallaisanensis* (NC_037433), *Phelipanche purpurea* (HG515536), *Phelipanche ramosa* (HG803180), *Rehmannia piasezkii* (KX636160), *Rehmannia elata* (KX636161), *Rehmannia glutinosa* (KX636157), *Rehmannia solanifolia* (KX636159), *Schwalbea americana* (HG738866), and *Triaenophora shennongjiaensis* (MH071405). The plastid genome of *Nicotiana tabacum*

Table 1 Summary of genes identified in the *D. coccinea* plastome.

Function	Genes[a]
Ribosomal proteins (large subunit)	**rpl2**[i], rpl14, **rpl16**, **rpl20**, rpl22, **rpl33**, **rpl36**
Ribosomal proteins (small subunit)	**rps2**, rps3, **rps4**, **rps7**, **rps8**, **rps11**, **rps12**[i], **rps14**, rps16 , **rps18**, rps19
Other protein-coding genes	infA, **matK** , **clpP** *(2 genes)*, accD, **ycf1**, **ycf2**
rRNAs	**rrn16, rrn23**[b], **rrn4.5, rrn5**
tRNAs	trnC-GCA, **trnD-GUC**, **trnE-UUC**, trnF-GAA, trnG-GCC, trnG-UCC[i], **trnH-GUG**, **trnI-CAU**, **trnL-CAA**, trnL-UAA[i], **trnL-UAG**, **trnM-CAU**, **trnfM-CAU**, **trnN-GUU**, **trnP-UGG**, **trnQ-UUG**, trnR-UCU, **trnS-GCU**, trnS-GGA, trnS-UGA, trnT-GGU, trnT-UGU, trnV-GAC, **trnW-CCA**, **trnY-GUA**
Pseudogenes	ψrpl32, ψrps15, ψtrnI-GAU [i], ψtrnfM-CAU, ψtrnM-CAU

Notes.
[a]Genes duplicated in inverted repeats were counted once, [i] denotes intron-containing genes, including trans-spliced *rps12*. Genes present in all Orobanchaceae species analysed in (*Wicke et al., 2016*) are shown in bold.
[b]*rrn23* gene contains 259-bp intervening sequence.

(Z00044) was used as an outgroup. Nucleotide sequences were extracted from GenBank, concatenated for each of the genomes and aligned using MAFFT v.7.055b (*Katoh & Standley, 2013*), with default parameters. Poorly aligned regions were excluded using trimAl v.1.2rev59 software with the -gappyout option (*Capella-Gutiérrez, Silla-Martínez & Gabaldón, 2009*). A maximum likelihood phylogenetic tree was constructed using PhyML v.3.3 (*Guindon et al., 2010*). The Hasegawa-Kishino-Yano nucleotide substitution model with the gamma model of rate heterogeneity (HKY+Γ) was selected using jModeltest v. 2.1.10 (*Posada, 2008*).

Verification of the presence of two copies of *clpP*)

To verify the presence of two copies of the *clpP* gene, we designed primer pairs flanking each copy (Table S1). Appropriate PCR fragments obtained using DNA samples extracted from two individual *D. coccinea* plants were analyzed by agarose gel electrophoresis (Fig. S1) and sequenced by the Sanger method.

RESULTS AND DISCUSSION

Plastid genome structure and gene content

The plastome of *D. coccinea* was assembled into a circular sequence of 66,616 bp from approximately 1.3 million paired-end reads (∼6 kb fragments). It has a typical quadripartite structure with a 37,964 bp large single copy (LSC) region, a 5,220 bp small single copy (SSC) region and a pair of inverted repeats (IRs), each of 11,716 bp (Fig. 2).

The *D. coccinea* plastome was predicted to contain 54 presumably intact unique genes (Table 1), which was fewer than its fully autotrophic relative *Lindenbergia philippensis* (113 genes), but comparable to that of holoparasitic Orobanchaceae (42–74; *Wicke et al., 2013*). Consistent with the inability to photosynthesize and the holoparasitic lifestyle of *D. coccinea*, its plastome lacks all genes coding for the NAD(P)H dehydrogenase complex and photosynthesis-related proteins. In particular, the *D. coccinea* plastome lacks the ATP synthase genes that are retained intact in most parasitic Orobanchaceae (*Wicke et al., 2013*). The genes for plastid-encoded RNA polymerase are also missing.
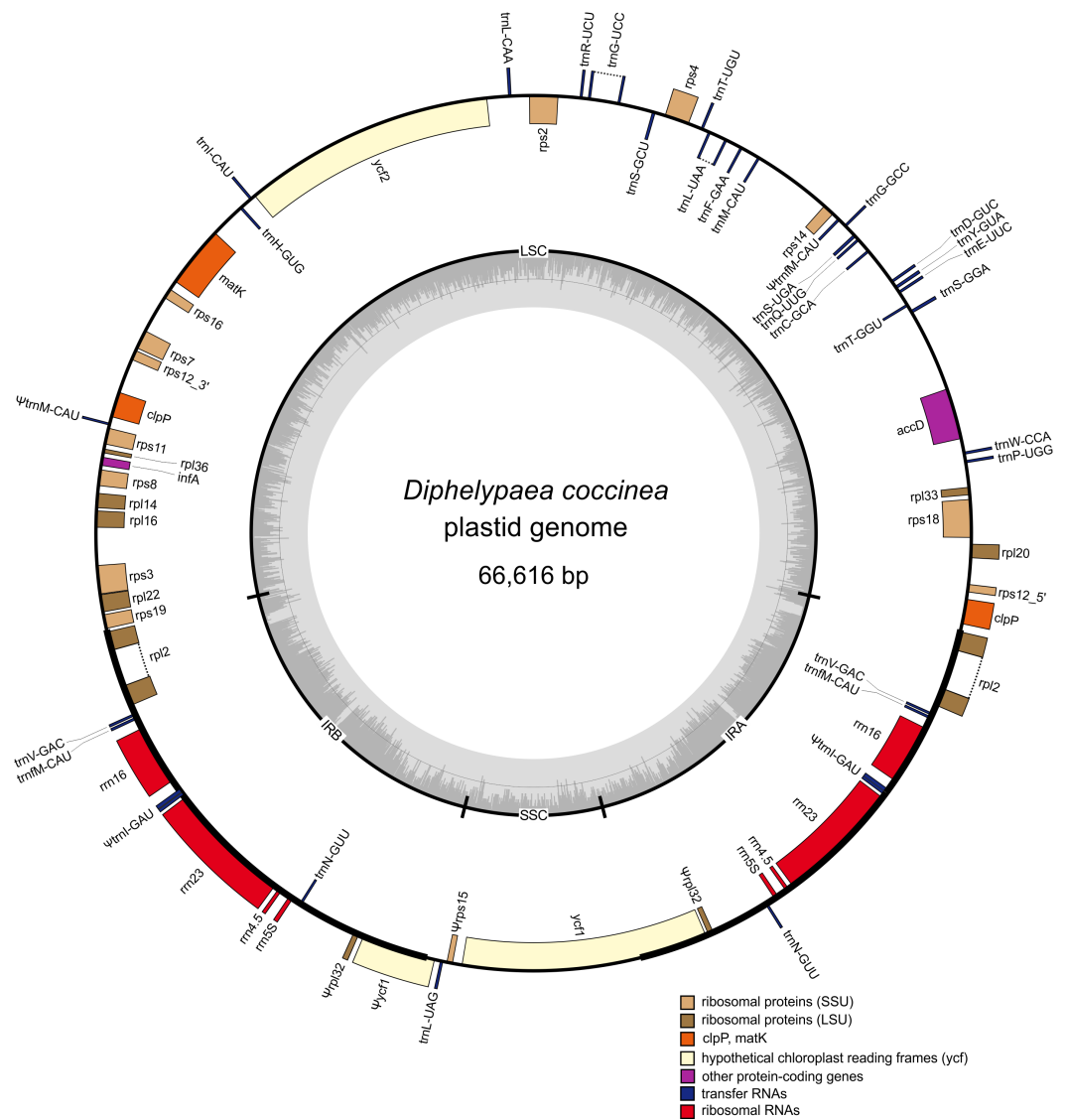
Figure 2 Gene map of the *D. coccinea* plastid genome. Genes shown inside the outer circle are transcribed clockwise, and those outside the circle are transcribed counter clockwise. Differential functional gene groups are color-coded. Pseudogenes are marked by ψ. GC content variation is shown in the middle circle.

Most of the retained genes are involved in protein synthesis: four rRNA genes, 25 tRNA genes, seven genes coding for the small subunit ribosomal proteins and 11 for the large subunit ribosomal proteins. Most ribosomal protein genes usually found in plastid genomes of photosynthetic angiosperms were also in the plastome of *D. coccinea* (Table 1). Exceptions are *rps15* and *rpl32*, which are retained as truncated pseudogenes, and *rpl23,* which could not be identified. The loss of these genes has also been reported in other holoparasitic Orobanchaceae (*Wicke et al., 2013*). While the plastid genome of *D. coccinea* contains genes for 25 tRNA species (Table 1), it lacks the tRNA genes *trnA-UGC,*

trnI-GAU, *trnK-UUU*, *trnR-ACG* and *trnV-UAC*, which are usually present in the plastomes of photosynthetic flowering plants. The loss of essential tRNA genes has been observed in parasitic plant plastomes showing an advanced stage of degradation (*DePamphilis & Palmer, 1990*; *Delannoy et al., 2011*; *Wicke et al., 2013*; *Ravin et al., 2016*). Analysis of the plastomes of holoparasitic Orobanchaceae revealed that up to 13 of 30 conserved tRNA genes, including those mentioned above, could be lost or pseudogenized (*Wicke et al., 2013*).

Besides genes involved in protein synthesis, the *D. coccinea* plastid genome contains genes *infA*, *matK*, *accD*, *clpP*, *ycf1* and *ycf2*, all of which were lost in some lineages of angiosperms, both parasitic and autotrophic. The most frequently lost gene is *infA*, encoding a translation initiation factor, the loss of which has been described in at least 24 separate lineages of angiosperms (*Millen et al., 2001*). The AccD protein, the beta subunit of acetyl-CoA carboxylase involved in fatty acid synthesis and leaf development (*Kode et al., 2005*), is essential for plastome maintenance (*Krause, 2012*). The gene for this protein is preserved even in the plastomes of most parasitic plants, although loss from the plastome and functional relocation to the nucleus occurs in some photosynthetic species (*Rousseau-Gueutin et al., 2013*). Among the Orobanchaceae, it has a 5′ truncation in the hemiparasite *Schwalbea americana* and holoparasitic species *Phelipanche purpurea* and *Phelipanche ramosa* (*Wicke et al., 2013*). Although *ycf1* and *ycf2* are considered to be essential for plastid maintenance (*Drescher et al., 2000*), multiple instances of loss of these genes, without transfer to the nuclear genome, have been reported in plant plastids (e.g., *Wakasugi, Tsudzuki & M, 2001*; *Cai et al., 2008*), but not in parasitic Orobanchaceae (*Wicke et al., 2013*). The functionality of the *ycf1* and *ycf2* genes in the *D. coccinea* plastome might be questioned, since their deduced protein products have long repeat-containing internal insertions contrary to typical Ycf1 and Ycf2 proteins, e.g., from *N. tabacum*. However, these insertions did not interrupt the open reading frames, suggesting that the proteins could retain functionality.

An interesting finding was the detection of two likely functional copies of the *clpP* gene in the *D. coccinea* plastome. *clpP* encodes a proteolytic subunit of Clp protease involved in protein metabolism within the plastid (*Krause, 2012*), and was proposed to be essential, being present even in highly reduced plastomes of parasitic plants (*Delannoy et al., 2011*; *Ravin et al., 2016*). Among the Orobanchaceae, a presumably functional *clpP* was found in all species (*Wicke et al., 2013*). In the *D. coccinea* plastome, two copies of *clpP* with 86% nucleotide sequence identity were found (Fig. 1). This duplication of the *clpP* gene was confirmed by PCR for two individual *D. coccinea* plants (Fig. S1). Notably, both genes lacked the introns usually present in *clpP* of photosynthetic angiosperms, but often absent in parasitic species. Their deduced protein products showed 75% amino acid sequence identity, but <35% identity with other plastidial ClpP proteins. It should be noted, that fast evolution of ClpP has been observed in several parasitic and photosynthetic lineages (*Wicke et al., 2013*; *Sloan et al., 2014*). Both gene copies contained intact reading frames; therefore, they probably remain functional.

Reduction of the size and gene content of the *D. coccinea* plastome is also reflected in the loss of introns in the remaining genes. Introns are only present in *rpl2*, *rps12* (*trans*-spliced),

trnG-UCC and *trnL-UAA*, and appeared to be lost in *clpP*, *rpl16*, *rps16* and *rps12* (*cis*-spliced intron). The presence of *matK* correlates with the retention of a group IIA intron in the *rpl2* gene, which requires maturase activity for splicing (*Zoschke et al., 2010*).

As in most angiosperms, the *D. coccinea* plastome contained an *rrn* gene cluster within an IR region. Although the gene order is typical (*rrn16-rrn23-rrn4-rrn5*), two tRNA genes located in the *rrn16-rrn23* spacer were lost (*trnA-UGC*) or truncated as a pseudogene (*trnI-GAU*). Interestingly, the *rrn23* gene contains a 259 bp intervening sequence, absent from any other plastidial *rrn23* sequence available in the GenBank database. A BLASTn search against GenBank found no sequences with high similarity to this insert. The insert occurred in the side hairpin of the H38 helix region of domain II of the *rrn23* gene (Fig. S2). Introns in *rrn23* genes have been found in the plastomes of the charophytes *Chlorella* (*Wakasugi et al., 1997*) and *Chlamydomonas* (*Turmel et al., 1993*), and the hornwort *Anthoceros formosae* (*Kugita et al., 2003*); however, they have not been reported in other land plants. The site of insertion in the *D. coccinea rrn23* gene did not match the positions of intron insertions in the charophytes and *Anthoceros formosae*. It is possible that this intervening sequence is not an intron but an insertion that is still compatible with the final structure of the ribosome. RNA-seq analysis would help to clarify this issue.

Structural rearrangements and duplications in the *D. coccinea* plastome

Plastid genomes of most angiosperms are highly conserved, not only in terms of overall quadripartite structure and gene content, but also in the gene order. Although deviation from a conserved gene order due to plastid genome rearrangements occurred in some photosynthetic lineages, e.g., cereals, geranium and clover (*Chumley et al., 2006*; *Cai et al., 2008*), numerous translocations, duplications, inversions and deletions are most frequently observed in the plastomes of parasitic species (*Wicke et al., 2013*). Among Orobanchaceae, the plastome of autotrophic *Lindenbergia philippensis* is colinear with that of tobacco, and limited deviations are observed in the hemiparasitic species *Schwalbea americana* (*Wicke et al., 2013*). In addition, no major rearrangements are observed in the plastome of *Lathraea squamaria*, belonging to the Rhinantheae clade of Orobanchaceae (*Samigullin et al., 2016*). However, gene deletions, duplications, inversions, shifts of IR boundaries and even complete loss of one IR are observed in the plastomes of holoparasitic species of the Orobancheae clade (*Wicke et al., 2013*; *Schneider et al., 2018*).

Comparison of the order of genes in the plastome of *D. coccinea* with the standard for angiosperms showed that the *D. coccinea* plastome, in addition to gene losses, experienced multiple rearrangements, including inversion, translocation and duplication of genes (Fig. 2). The gene order differs from that in other species of Orobanchaceae reported by *Wicke et al. (2016)*. However, the *D. coccinea* plastome retains the highly conserved S10 operon (*rpl2*, *rps19*, *rpl22*, *rps3*, *rpl16*, *rpl14*, *rps8*, *infA*, *rpl36*, *rps11*) and the *rrn* gene cluster. These operons are conserved in all plastomes of Orobanchaceae, but appear to be deconstructed in some of the most highly reduced genomes of parasitic plants (*Bellot & Renner, 2015*; *Ravin et al., 2016*).
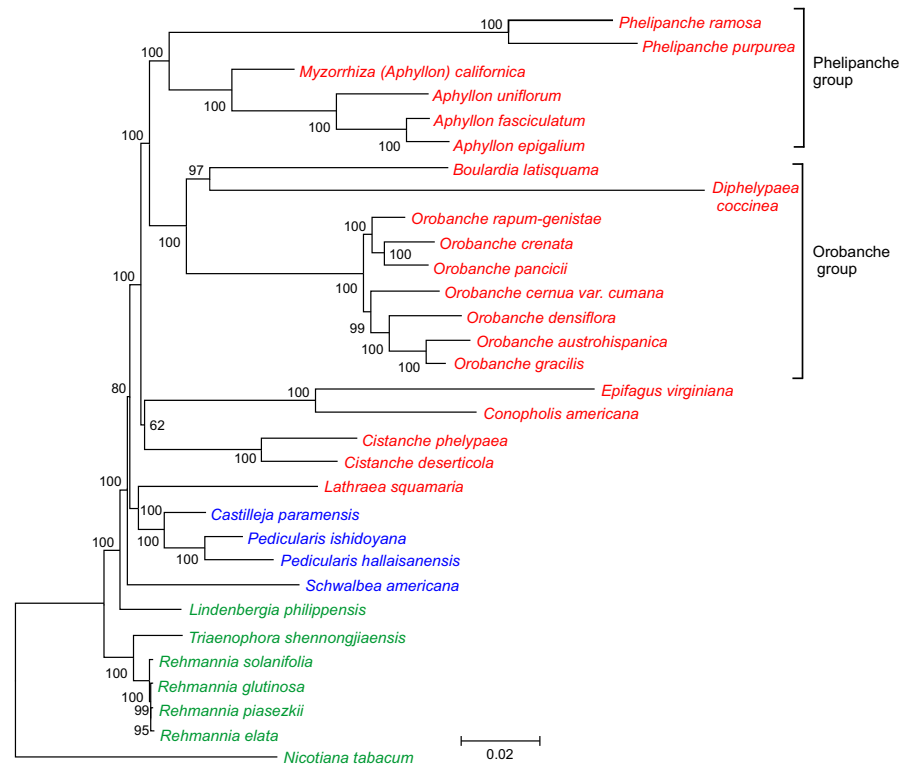
Figure 3 **Phylogenetic tree of 30 taxa of *Orobanchaceae*.** The tree was inferred by the maximum like-lihood approach. Bootstrap support values are provided at the nodes. The scale bar corresponds to 0.02 substitutions per site. Autotrophic taxa are shown in green, hemiparasites in blue and holoparasites in red.

Full-size 🖼 DOI: 10.7717/peerj.7830/fig-3

Phylogenetic position of *D. coccinea*

Relatively little is known about the phylogenetic position of the genus *Diphelypaea*, and only two ∼600 bp-long nuclear internal transcribed spacer (ITS) sequences of *D. coccinea* and *D. tournefortii* are available in GenBank (accessed on May 3, 2019). Phylogenetic analysis of the nuclear ITS region revealed that the genus *Orobanche*, as defined by *Beck-Mannagetta (1930)*, divided into two genus-level groups, the *Orobanche* group (*Orobanche* sect. *Orobanche* and *Diphelypaea*) and the *Phelipanche* group (*Schneeweiss et al., 2004a*). This split is also supported by karyological features, since the chromosome base number in the *Orobanche*/*Diphelypaea* group is $x = 19$, while in the *Phelipanche* group it is $x = 12$ (*Schneeweiss et al., 2004b*). In the ITS phylogeny, two *Diphelypaea* species formed a basal lineage in the *Orobanche* group. However, these molecular phylogenetic implications were based on a limited sequence dataset and need to be clarified.

In this current study, we took advantage of availability of the complete plastid genome sequence of *D. coccinea* to define the phylogenetic position of this genus in the Orobanchaceae. Analysis of the concatenated nucleotide sequences of 17 conserved genes from 30 species of Orobanchaceae yielded a fully resolved phylogenetic tree (Fig. 3). *D. coccinea* appeared to be included in a cluster also comprising *Boulardia latisquama* and

species of *Orobanche* sect. *Orobanche*. By contrast, the *Phelipanche* group, comprising *Phelipanche* and *Aphyllon*, formed a distinct lineage.

CONCLUSIONS

Being only 66,616 bp in size and containing 54 presumably intact unique genes, the plastome of *D. coccinea* is one of the most highly reduced among the parasitic Orobanchaceae. Plastome rearrangements, gene duplications and the loss of introns are associated with gene loss and genome reduction. More pronounced gene loss has only been reported in the plastomes of *Conopholis americana* (45,673 bp, 42 genes), *Epifagus virginiana* (70,028 bp, 42 genes) and *Boulardia latisquama* (80,361 bp, 49 genes). In particular, the *D. coccinea* plastome lacks all genes of the photosynthetic apparatus, including ATP synthase genes that are retained intact in most Orobanchaceae. However, all 16 protein-coding genes, 14 tRNA genes and 4 rRNA genes commonly present in the plastid genomes of all hemi- and holoparasitic Orobanchaceae species (Table 1) are present in *D. coccinea*, suggesting that further gene loss is unlikely in this lineage. Phylogenetic analysis confirmed that *D. coccinea* belongs to the *Orobanche* group of the family Orobanchaceae.

ADDITIONAL INFORMATION AND DECLARATIONS

Competing Interests

The authors declare there are no competing interests.

Author Contributions

- Eugeny V. Gruzdev performed the experiments, analyzed the data, prepared figures and/or tables, authored or reviewed drafts of the paper, approved the final draft.
- Vitaly V. Kadnikov conceived and designed the experiments, performed the experiments, contributed reagents/materials/analysis tools, prepared figures and/or tables, approved the final draft.
- Alexey V. Beletsky performed the experiments, analyzed the data, contributed reagents/materials/analysis tools, authored or reviewed drafts of the paper, approved the final draft.

- Andrey V. Mardanov conceived and designed the experiments, performed the experiments, authored or reviewed drafts of the paper, approved the final draft.
- Nikolai V. Ravin conceived and designed the experiments, performed the experiments, analyzed the data, prepared figures and/or tables, authored or reviewed drafts of the paper, approved the final draft.

Data Availability

The following information was supplied regarding data availability:

The raw reads are available in the Sequence Read Archive (SRA): SRR9665263 and the sequence of the plastid genome of *D. coccinea* is available at GenBank: MK922354.

Supplemental Information

Supplemental information for this article can be found online at http://dx.doi.org/10.7717/peerj.7830#supplemental-information.

REFERENCES

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

Barrett CF, Freudenstein JV, Li J, Mayfield-Jones DR, Perez L, Pires JC, Santos C. 2014. Investigating the path of plastid genome degradation in an early-transitional clade of heterotrophic orchids, and implications for heterotrophic angiosperms. *Molecular Biology and Evolution* **31**:3095–3112 DOI 10.1093/molbev/msu252.

Beck-Mannagetta G. 1930. IV. 261, *Orobanchaceae*. In: Engler A, ed. *Das Pflanzenreich. Regni Vegetabilis Conspectus.* Leipzig: Wilhelm Engelmann, 1–348.

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 Biology and Evolution* **8(1)**:189–201 DOI 10.1093/gbe/evv251.

Bennett JR, Mathews S. 2006. Phylogeny of the parasitic plant family *Orobanchaceae* inferred from phytochrome A. *American Journal of Botany* **93**:1039–1051 DOI 10.3732/ajb.93.7.1039.

Cai Z, Guisinger M, Kim HG, Ruck E, Blazier JC, McMurtry V, Kuehl JV, Boore J, Jansen RK. 2008. Extensive reorganization of the plastid genome of *Trifolium subterraneum* (*Fabaceae*) is associated with numerous repeated sequences and novel DNA insertions. *Molecular Biology and Evolution* **67(6)**:696–704 DOI 10.1007/s00239-008-9180-7.

Capella-Gutiérrez S, Silla-Martínez JM, Gabaldón T. 2009. trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics* **25(15)**:1972–1973 DOI 10.1093/bioinformatics/btp348.

Cho W-B, Lee D-H, Choi I-S, Lee J-H. 2018. The complete chloroplast genome of hemiparasitic *Pedicularis hallaisanensis* (*Orobanchaceae*). *Mitochondrial DNA Part B—Resources* **3(1)**:235–236 DOI 10.1080/23802359.2018.1437820.

Chumley TW, Palmer JD, Mower JP, Fourcade HM, Calie PJ, Boore JL, Jansen RK. 2006. The complete chloroplast genome sequence of *Pelargonium x hortorum*: organization and evolution of the largest and most highly rearranged chloroplast genome of land plants. *Molecular Biology and Evolution* **23**(**11**):2175–2190 DOI 10.1093/molbev/msl089.

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.

DePamphilis CW, Palmer JD. 1990. Loss of photosynthetic and chlororespiratory genes from the plastid genome of a parasitic flowering plant. *Nature* **348**:337–339 DOI 10.1038/348337a0.

Drescher A, Ruf S, Calsa Jr T, Carrer H, Bock R. 2000. The two largest chloroplast genome-encoded open reading frames of higher plants are essential genes. *Plant Journal* **22**(**2**):97–104 DOI 10.1046/j.1365-313x.2000.00722.x.

Fan W, Zhu A, Kozaczek M, Shah N, Pabon-Mora N, Gonzalez F, Mower JP. 2016. Limited mitogenomic degradation in response to a parasitic lifestyle in Orobanchaceae. *Scientific Reports* **6**:36285 DOI 10.1038/srep36285.

Frailey DC, Chaluvadi SR, Vaughn JN, Coatney CG, Bennetzen JL. 2018. Gene loss and genome rearrangement in the plastids of five hemiparasites in the family Orobanchaceae. *BMC Plant Biology* **18**(**1**):30 DOI 10.1186/s12870-018-1249-x.

Graham SW, Lam VK, Merckx VS. 2017. Plastomes on the edge: the evolutionary breakdown of mycoheterotroph plastid genomes. *New Phytologist* **214**(**1**):48–55 DOI 10.1111/nph.14398.

Guindon S, Dufayard JF, Lefort V, Anisimova M, Hordijk W, Gascuel O. 2010. New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Systematic Biology* **59**(**3**):307–321 DOI 10.1093/sysbio/syq010.

Katoh K, Standley DM. 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Molecular Biology and Evolution* **30**:772–780 DOI 10.1093/molbev/mst010.

Kode V, Mudd EA, Iamtham S, Day A. 2005. The tobacco plastid *accD* gene is essential and is required for leaf development. *Plant Journal* **44**:237–244 DOI 10.1111/j.1365-313X.2005.02533.x.

Krause K. 2012. Plastid genomes of parasitic plants: a trail of reductions and losses. In: Bullerwell CE, ed. *Organelle genetics.* Berlin, Heidelberg: Springer-Verlag, 79–103.

Kugita M, Kaneko A, Yamamoto Y, Takeya Y, Matsumoto T, Yoshinaga K. 2003. The complete nucleotide sequence of the hornwort (*Anthoceros formosae*) chloroplast genome: insight into the earliest land plants. *Nucleic Acids Research* **31**(**2**):716–721 DOI 10.1093/nar/gkg155.

Lohse M, Drechsel O, Kahlau S, Bock R. 2013. OrganellarGenomeDRAW—a suite of tools for generating physical maps of plastid and mitochondrial genomes and visualizing expression data sets. *Nucleic Acids Research* **41**(**Web Server issue**):W575–W581 DOI 10.1093/nar/gkt289.

Lowe TM, Chan PP. 2016. tRNAscan-SE On-line: integrating search and context for analysis of transfer RNA genes. *Nucleic Acids Research* **44**(**W1**):W54–W57 DOI 10.1093/nar/gkw413.

McNeal JR, Bennett JR, Wolfe AD, Mathews S. 2013. Phylogeny and origins of holoparasitism in *Orobanchaceae*. *American Journal of Botany* **100**:971–983 DOI 10.3732/ajb.1200448.

Millen RS, Olmstead RG, Adams KL, Palmer JD, Lao NT, Heggie L, Kavanagh TA, Hibberd JM, Gray JC, Morden CW, Calie PJ, Jermiin LS, Wolfe KH. 2001. Many parallel losses of *infA* from chloroplast DNA during angiosperm evolution with multiple independent transfers to the nucleus. *The Plant Cell* **13**(**3**):645–658 DOI 10.1105/tpc.13.3.645.

Molina J, Hazzouri KM, Nickrent D, Geisler M, Meyer RS, Pentony MM, Flowers JM, Pelser P, Barcelona J, Inovejas SA, Uy I, Yuan W, Wilkins O, Michel CI, Locklear S, Concepcion GP, Purugganan MD. 2014. Possible loss of the chloroplast genome in the parasitic flowering plant *Rafflesia lagascae* (Rafflesiaceae). *Molecular Biology and Evolution* **31**:793–803 DOI 10.1093/molbev/msu051.

Murray MG, Thompson WF. 1980. Rapid isolation of high molecular weight plant DNA. *Nucleic Acids Research* **8**:4321–4325 DOI 10.1093/nar/8.19.4321.

Nickrent D, Du RJ, Colwell AE, Wolfe AD, Young ND, Steine KE, DePamphilis CW. 1997. Molecular phylogenetic and evolutionary studies of parasitic plants. In: Soltis DE, Soltis PS, Doyle JJ, eds. *Molecular systematics of plants II. DNA sequencing.* Boston: Kluwer Academic Publishers, 211–241.

Nicolson DH. 1975. *Diphelypaea* (Orobanchaceae), nom. nov. and other cauterizations on a nomenclatural hydra. *Taxon* **24**(**5/6**):651–657 DOI 10.2307/1220740.

Petersen G, Cuenca A, Seberg O. 2015. Plastome evolution in hemiparasitic mistletoes. *Genome Biology and Evolution* **7**(**9**):2520–2532 DOI 10.1093/gbe/evv165.

Posada D. 2008. jModelTest: phylogenetic model averaging. *Molecular Biology and Evolution* **25**(**7**):1253–1256 DOI 10.1093/molbev/msn083.

Ravin NV, Gruzdev EV, Beletsky AV, Mazur AM, Prokhortchouk EB, Filyushin MA, Kochieva EZ, Kadnikov VV, Mardanov AV, Skryabin KG. 2016. The loss of photosynthetic pathways in the plastid and nuclear genomes of the non-photosynthetic mycoheterotrophic eudicot *Monotropa hypopitys*. *BMC Plant Biology* **16**(**Suppl 3**):238 DOI 10.1186/s12870-016-0929-7.

Rousseau-Gueutin M, Huang X, Higginson E, Ayliffe M, Day A, Timmis JN. 2013. Potential functional replacement of the plastidic acetyl-CoA carboxylase subunit (accD) gene by recent transfers to the nucleus in some angiosperm lineages. *Plant Physiology* **161**(**4**):1918–1929 DOI 10.1104/pp.113.214528.

Samigullin TH, Logacheva MD, Penin AA, Vallejo-Roman CM. 2016. Complete plastid genome of the recent holoparasite *Lathraea squamaria* reveals earliest

stages of plastome reduction in *Orobanchaceae*. *PLOS ONE* **11(3)**:e0150718
DOI 10.1371/journal.pone.0150718.

Schneeweiss GM, Colwell A, Park JM, Jang CG, Stuessy TF. 2004a. Phylogeny of
holoparasitic *Orobanche* (*Orobanchaceae*) inferred from nuclear ITS sequences.
Molecular Phylogenetics and Evolution **30(2)**:465–478
DOI 10.1016/S1055-7903(03)00210-0.

Schneeweiss GM, Palomeque T, Colwell AE, Weiss-Schneeweiss H. 2004b. Chromo-
some numbers and karyotype evolution in holoparasitic *Orobanche* (*Orobanchaceae*)
and related genera. *American Journal of Botany* **91(3)**:439–448
DOI 10.3732/ajb.91.3.439.

Schneider AC, Chun H, Stefanovic S, Baldwin BG. 2018. Punctuated plastome reduction
and host-parasite horizontal gene transfer in the holoparasitic plant genus *Aphyllon*.
Proceedings of the Royal Society B: Biological Sciences **285(1887)**:Article 20181535
DOI 10.1098/rspb.2018.1535.

Sloan DB, Triant DA, Forrester NJ, Bergner LM, Wu M, Taylor DR. 2014. A recur-
ring syndrome of accelerated plastid genome evolution in the angiosperm tribe
Sileneae (Caryophyllaceae). *Molecular Phylogenetics and Evolution* **72**:82–89
DOI 10.1016/j.ympev.2013.12.004.

Turmel M, Gutell RR, Mercier JP, Otis C, Lemieux C. 1993. Analysis of the chloroplast
large subunit ribosomal RNA gene from 17 *Chlamydomonas* taxa. Three internal
transcribed spacers and 12 group I intron insertion sites. *Journal of Molecular Biology*
232:446–467 DOI 10.1006/jmbi.1993.1402.

**Wakasugi T, Nagai T, Kapoor M, Sugita M, Ito M, Ito S, Tsudzuki J, Nakashima K,
Tsudzuki T, Suzuki Y, Hamada A, Ohta T, Inamura A, Yoshinaga K, Sugiura M.
1997.** Complete nucleotide sequence of the chloroplast genome from the green alga
Chlorella vulgaris: the existence of genes possibly involved in chloroplast division.
Proceedings of the National Academy of Sciences of the United States of America
94:5967–5972 DOI 10.1073/pnas.94.11.5967.

Wakasugi T, Tsudzuki T, M Sugiura. 2001. The genomics of land plant chloroplasts:
gene content and alteration of genomic information by RNA editing. *Photosynthesis
Research* **70(1)**:107–118 DOI 10.1023/A:1013892009589.

Wicke S, Müller KF, DePamphilis CW, Quandt D, Bellot S, Schneeweiss GM. 2016.
Mechanistic model of evolutionary rate variation en route to a nonphotosynthetic
lifestyle in plants. *Proceedings of the National Academy of Sciences of the United States
of America* **113(32)**:9045–9050 DOI 10.1073/pnas.1607576113.

**Wicke S, Müller KF, DEPamphilis CW, Quandt D, Wickett NJ, Zhang Y, Renner SS,
Schneeweiss GM. 2013.** Mechanisms of functional and physical genome reduction in
photosynthetic and nonphotosynthetic parasitic plants of the broomrape family. *The
Plant Cell* **25(10)**:3711–3725 DOI 10.1105/tpc.113.113373.

Wolfe KH, Morden CW, Palmer JD. 1992. Function and evolution of a minimal
plastid genome from a nonphotosynthetic parasitic plant. *Proceedings of the
National Academy of Sciences of the United States of America* **89**:10648–10652
DOI 10.1073/pnas.89.22.10648.

Wyman SK, Jansen RK, Boore JL. 2004. Automatic annotation of organellar genomes with DOGMA. *Bioinformatics* **20**:3252–3255 DOI 10.1093/bioinformatics/bth352.

Xia Z, Wen J. 2018. The complete chloroplast genome of the endangered species *Triaenophora shennongjiaensis* (*Orobanchaceae* s.l.). *Mitochondrial DNA Part B—Resources* **3(2)**:506–507 DOI 10.1080/23802359.2018.1467242.

Young ND, Steiner KE, DePamphilis CW. 1999. The evolution of parasitism in *Scrophulariaceae/Orobanchaceae*: plastid gene sequences refute an evolutionary transition series. *Annals of the Missouri Botanical Garden* **86**:876–893 DOI 10.2307/2666173.

Zeng SY, Zhou T, Han K, Yang Y, Liu Z-L. 2017. The complete chloroplast genome sequences of six *Rehmannia* species. *Gene* **8(3)**:E103 DOI 10.3390/genes8030103.

Zoschke R, Nakamura M, Liere K, Sugiura M, Börner T, Schmitz-Linneweber C. 2010. An organellar maturase associates with multiple group II introns. *Proceedings of the National Academy of Sciences of the United States of America* **107(7)**:3245–3250 DOI 10.1073/pnas.0909400107.