Plastome of the mycoheterotrophic eudicot *Exacum paucisquama* (Gentianaceae) exhibits extensive gene loss and a highly expanded inverted repeat region

Zhanghai Li[1], Xiao Ma[1], Yi Wen[1,2], Sisi Chen[3], Yan Jiang[1,2] and Xiaohua Jin[1,4]

[1] State Key Laboratory of Systematic and Evolutionary Botany, Institute of Botany, Chinese Academy of Sciences, Beijing, China
[2] University of Chinese Academy of Sciences, Beijing, China
[3] Nanchang University, Nanchang, China
[4] Southeast Asia Biodiversity Research Institute, Chinese Academy of Sciences (CAS-SEABRI), Xishuangbanna, China

ABSTRACT

Mycoheterotrophic plants are highly specialized species able to acquire organic carbon from symbiotic fungi, with relaxed dependence on photosynthesis for carbon fixation. The relaxation of the functional constraint of photosynthesis and thereby the relaxed selective pressure on functional photosynthetic genes usually lead to substantial gene loss and a highly degraded plastid genome in heterotrophs. In this study, we sequenced and analyzed the plastome of the eudicot *Exacum paucisquama*, providing the first plastid genome of a mycoheterotroph in the family Gentianaceae to date. The *E. paucisquama* plastome was 44,028 bp in length, which is much smaller than the plastomes of autotrophic eudicots. Although the *E. paucisquama* plastome had a quadripartite structure, a distinct boundary shift was observed in comparison with the plastomes of other eudicots. We detected extensive gene loss and only 21 putative functional genes (15 protein-coding genes, four rRNA genes and two tRNA genes). Our results provide valuable information for comparative evolutionary analyses of plastomes of heterotrophic species belonging to different phylogenetic groups.

Subjects Evolutionary Studies, Plant Science
Keywords Mycoheterotrophy, *Exacum paucisquama*, Gentianaceae, Plastid genome, Gene loss, Inverted repeat region expansion

INTRODUCTION

The characteristic feature of green plants is the presence of chloroplasts, which convert inorganic carbon to organic carbon by photosynthesis. However, a small number of plants have evolved a more specific type of nutrient acquisition strategy, namely heterotrophy. Heterotrophic plants can be subdivided into parasitic and mycoheterotrophic plants which acquire organic carbon from plant hosts or symbiotic fungi, respectively (*Merckx et al., 2013a*). Heterotrophy has evolved independently in plants numerous times, enabling the investigation of the loss of photosynthetic capability on a broad phylogenetic

scale (*Barrett et al., 2014*; *Barrett, Wicke & Sass, 2018*; *Logacheva et al., 2014*; *Feng et al., 2016*; *Lim et al., 2016*; *Wicke et al., 2016*; *Barrett & Kennedy, 2018*; *Petersen et al., 2018*).

The relaxation of the functional constraint of photosynthesis results in the relaxed selective pressure on functional photosynthetic genes, which leads to the substantial gene loss and a highly degraded plastid genome in heterotrophs (*Yuan et al., 2018*). However, the reduction of the plastome size, gene loss and mutations are observed to varying degrees in heterotrophs (*Delannoy et al., 2011*; *Logacheva et al., 2014*; *Molina et al., 2014*; *Lam, Soto Gomez & Graham, 2015*; *Schelkunov et al., 2015*; *Feng et al., 2016*; *Lim et al., 2016*; *Naumann et al., 2016*; *Graham, Lam & Merckx, 2017*; *Barrett & Kennedy, 2018*; *Petersen et al., 2018*; *Kim et al., 2019*). General trends in gene loss observed in independently evolved heterotrophic lineages have provided a basis for the concept of a conserved core gene set, originally thought to include rRNA genes, ribosomal protein genes, several tRNA genes and four protein-coding genes (*clp*P, *acc*D, *ycf*1 and *ycf*2) (*Delannoy et al., 2011*; *Logacheva, Schelkunov & Penin, 2011*). Subsequently, *Barrett & Davis (2012)* proposed a model to describe the pattern of plastid gene loss in heterotrophs during plastome degradation (*Barrett & Davis, 2012*; *Barrett et al., 2014*). Based on recent studies, *Graham, Lam & Merckx (2017)* proposed a modified model, with broader windows for the retention of PEP, Rubisco and ATP synthase. Hence, plastid genome sequences of additional heterotrophic species from different phylogenetic clades are important to evaluate different scenarios of core gene retention in fully heterotrophic taxa.

The family Gentianaceae (order Gentianales) includes approximately 99 genera and 1,736 species. Of these, 25 species are putative full mycoheterotrophs, with at least four independent origins in *Voyria*, *Voyriella*, *Exacum* (including the formerly recognized genus *Cotylanthera*) and *Exochaenium* (*Merckx et al., 2013a*; *Struwe, 2014*). The mycoheterotrophic lifestyle has only been reported in Gentianaceae in the order Gentianales (*Merckx et al., 2013b*). However, plastid genomes of heterotrophic species in this family have not yet been characterized. Little is known about the modification of the plastid genome in the mycoheterotrophic Gentianales.

In this study, we sequenced the complete plastome of *Exacum paucisquama* (Fig. 1), an achlorophyllous Gentianaceae species that parasitizes fungi to obtain nutrients (*Leake, 1994*; *Merckx & Freudenstein, 2010*). Our aims were to (1) explore the characteristics of the plastid genome of *E. paucisquama* and (2) compare the plastome of *E. paucisquama* with the plastomes of other heterotrophic plant lineages that have undergone independent losses of photosynthesis-related genes. The results of these analyses provide a more comprehensive understanding of plastome evolution in heterotrophic plants.

MATERIALS AND METHODS

Fresh stems and flowers were harvested from *E. paucisquama* plants growing in a field in Yunnan, China. Samples were dried in silica gel and preserved at −20 °C. Total DNA was isolated using a modified CTAB protocol (*Li et al., 2013*). DNA (>100 ng/ml) was sheared to fragments of 400–600 bp using Covaris M220. The NEBNext Ultra DNA Library Prep Kit was used to prepare DNA libraries for subsequent sequencing, according to the manufacturer's protocol. Paired-end sequencing with a read length of 150 bp was
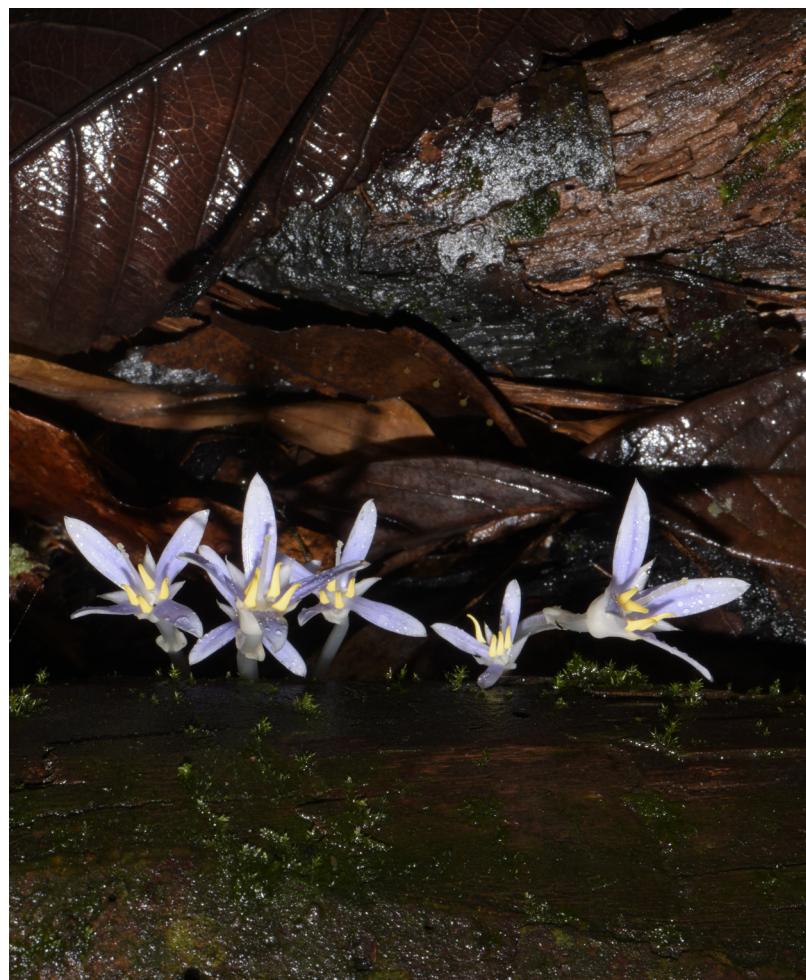
Figure 1 Image of the tiny, achlorophyllous eudicot *Exacum paucisquama* in its natural habitat. Photo: Xiao-Hua Jin.   Full-size ⊡ DOI: 10.7717/peerj.9157/fig-1

performed using the Illumina HiSeq 2500 platform at the Institute of Botany, Chinese Academy of Sciences.

To clean the raw sequence reads, quality control was performed following the methods of *Li et al. (2019)*. Plastomes were assembled from clean reads according to the methods of *Feng et al. (2016)* and *Li et al. (2019)*. In short, clean reads were mapped to the plastome of *Gentiana straminea* (GenBank, KJ657732) using Geneious v10.2.2 (http://www.geneious.com, last accessed 4 May 2019) to filter reads matching the reference genome (*Kearse et al., 2012*). De novo assemblies were constructed using VELVET with several *K*-mer values (*Zerbino & Birney, 2008*), and contigs from each assembly were merged and combined into scaffolds in Geneious. Then, using the scaffolds as references to filter the plastome reads from clean reads, the assembly steps were repeated to obtain the draft plastome. Additionally, NOVOPlasty v.2.7.1 (*Dierckxsens, Mardulyn & Smits, 2016*), which uses a reference sequence (in this study, *acc*D from the leafy, photosynthetic gentian *G. straminea*) as an initial seed, was used to build a draft plastome. Reads were mapped with high stringency to the draft plastomes, produced by both assemblies, using

Geneious to verify assembly errors. The plastome was first annotated using Geneious and GeSeq (*Tillich et al., 2017*), and tRNA genes were further predicted using tRNAscan-SE (*Lowe & Chan, 2016*). Start and stop codons, exon and intron boundaries and putative non-functional pseudogenes were identified and adjusted by aligning the plastome to protein-coding, tRNA and rRNA gene sequences of *G. straminea*. The circular plastome map was visualized using OGDRAW v1.2 (*Lohse et al., 2013*).

The plastome of *E. paucisquama* was aligned with the plastomes of three species, *G. straminea*, *Halenia corniculata* (GenBank, MK606372), and *Swertia verticillifolia* (GenBank, MF795137), using the progressiveMAUVE (*Darling, Mau & Perna, 2010*) plugin for Geneious to identify syntenic blocks and thereby to detect genomic rearrangements. Putative functional genes (i.e., genes with open reading frames), pseudogenes (i.e., genes with interrupted open reading frames or nontriplet nucleotide indels), and physical gene losses were identified by comparisons to the plastome of the leafy, photosynthetic relative *G. straminea*. To compare the plastome size and functional gene content among fully heterotrophic species, plastome sequences of 17 heterotrophs were downloaded from the NCBI database (Table S1).

RESULTS AND DISCUSSION

A total of 6,455,059,021 clean reads, with an average length of 150 bp, were recovered from 7,147,180,924 raw reads. Among the clean reads, 32,399,763 (0.50%) corresponded to the plastome. The average coverage depths were 644× and 1,359× in single-copy (SC) regions and inverted repeat (IR) regions, respectively (Fig. S1). The assembled plastome of *E. paucisquama* has been deposited in NCBI GenBank under accession number MN067514. The complete plastome of *E. paucisquama* was 44,028 bp in length, with a quadripartite structure (Fig. 2). It had a basically high degree of collinearity with the genome of its autotrophic relatives (Fig. S2). The plastid genome of *E. paucisquama* contained 21 putative functional genes, including 13 ribosomal protein genes, *clp*P, *acc*D. four rRNA genes and two tRNA genes (*trn*E and *trn*fM). The total GC content of the *E. paucisquama* plastome was 37.1%, after removing one copy of the IR region (Table S1).

The IR region is hypothesized to stabilize the plastome (*Palmer & Thompson, 1982*) and is retained in most sequenced plastomes, including those of heterotrophic species with extensive gene loss (*Lim et al., 2016*; *Schelkunov et al., 2015*). Unlike the plastomes of autotrophs in Gentianaceae, more than 80% of the plastome of *E. paucisquama* was classified as IRs (2 × 17,622 bp), and the IR regions harbored most of the genes, starting with *rps*3 and ending with a part of *rrn*23/*rps*7 (Fig. 2). IRs have a major impact on the rate of plastome sequence evolution; in general, the rate of nucleotide substitution is several times lower in IR than in SC regions (*Zhu et al., 2016*), and genes translocated from the SC into the IR region show decelerated substitution rates in the fern (*Li et al., 2016*).

The large single copy (LSC) region of the *E. paucisquama* plastome had a length of 6,651 bp and contained *rrn*16, the 3′ end of *rps*12, a part of *rrn*23, and a part of *rps*7. The small single copy (SSC) region was 2,133 bp in length and contained only *rpl*2, *rps*19,
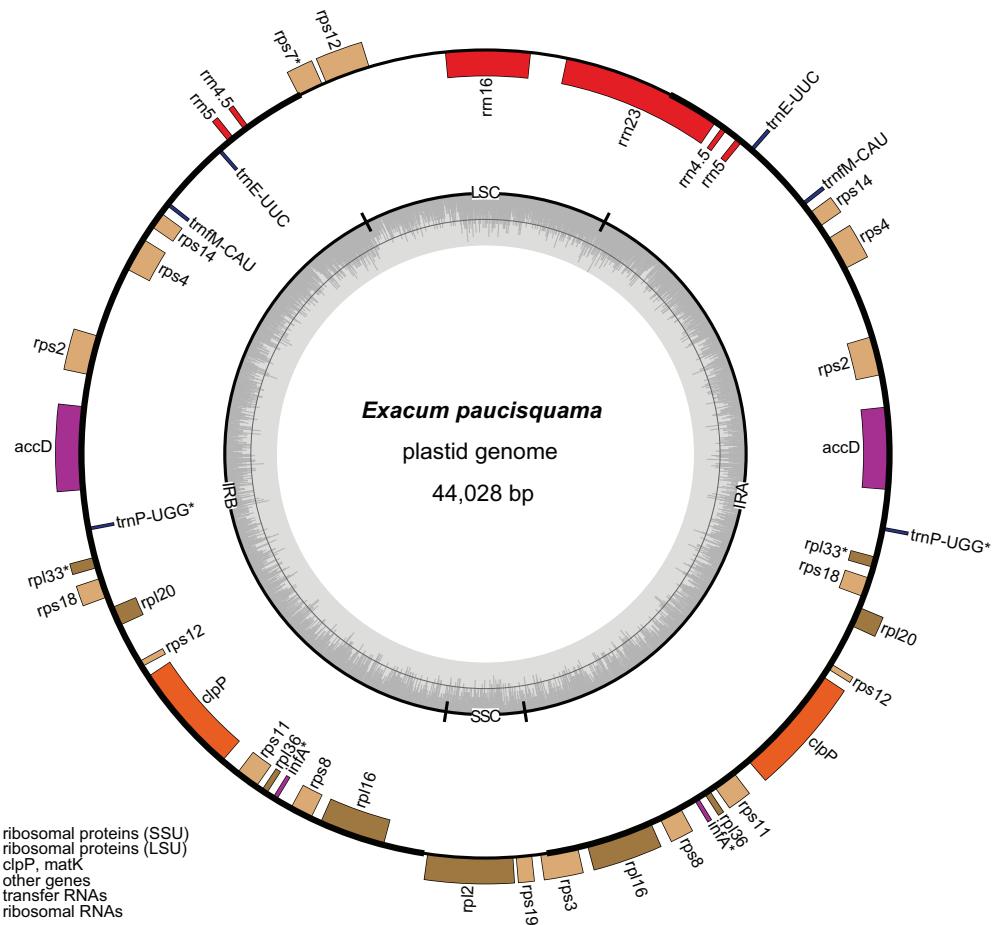
Figure 2 Circular map of the plastome of *E. paucisquama*. Asterisks (*) indicate pseudogenes. Thick lines indicate the extent of the inverted repeat regions (IRa and IRb), which separate the genome into small (SSC) and large (LSC) single copy regions. Genes drawn inside the circle are transcribed clockwise, while those outside of the circle are transcribed counter clockwise. Genes belonging to different functional groups are color coded. Dark gray in the inner circle corresponds to the GC content, while light gray corresponds to the AT content. Full-size ⬜ DOI: 10.7717/peerj.9157/fig-2

and a part of *rps*3. Generally, in plastid genomes, each IR region contains three to five tRNA genes, four *rrn* genes, two *rps* genes and *ycf*2, and the SSC region contains six *ndh* genes, *rpl*32, *psa*C, *rps*15, *ycf*1 and *trn*L^{UAG}. However, the *E. paucisquama* plastome showed a shift in the boundaries between the IR and SC regions. All of the genes usually present in the SSC region were translocated to the IR region, and some of the sequences usually contained in IR regions were located in the SC regions. The boundary shift led to the presence of a single copy of *rrn*16 and partial *rrn*23, in contrast to the duplicates in IR regions observed in most plastomes.

The extensive loss of plastid genes in *E. paucisquama* corresponds to the final stage of the plastome degradation model proposed by *Barrett & Davis (2012)*. Some "housekeeping" genes, such as *ycf*1, *ycf*2, *mat*K and many tRNA genes, present in most mycoheterotrophic species, were lost in the *E. paucisquama* plastome. The *E. paucisquama* plastome contains only two tRNA genes: *trn*fM^{CAU} and *trn*E^{UUC}. The loss of some

tRNA genes may be compensated for by the import of tRNAs from the cytosol (*Alkatib et al., 2012*) or by 'superwobbling' (*Rogalski, Karcher & Bock, 2008*). The gene product of *trn*E^{UUC} (glutamyl-tRNA) plays a secondary role in heme biosynthesis (*Jahn, Verkamp & So, 1992*) and may regulate the translation of the nuclear-encoded plastid RNA synthase (NEP) (*Hanaoka et al., 2005*). *Barbrook, Howe & Purton (2006)* proposed that the interaction of *trn*E^{UUC} with multiple enzymes involved in heme biosynthesis makes its replacement by a cytosolic product unlikely. *trn*fM^{CAU}, plastid-encoded formylmethionyl-tRNA, regulates the translation initiation in plastids and possibly in mitochondria (*Barbrook, Howe & Purton, 2006*). The indispensable nature of these two plastid-encoded tRNAs could explain the retention of plastomes in non-photosynthetic organisms; this hypothesis is referred to as the essential tRNA hypothesis (*Barbrook, Howe & Purton, 2006*). The *mat*K gene encodes the only plastid-encoded group IIa intron maturase, MATK (*Zoschke et al., 2010*). Although *mat*K is present in nearly all plant plastid genomes, it has been deleted from the plastome of *E. paucisquama*. However, the *E. paucisquama* plastome retained loci with group IIa introns, and at least two of these genes (*rpl*2 and *rps*12) are thought to be targeted by MATK (*Zoschke et al., 2010*). It is possible that an alternative splicing factor facilitates intron removal from their RNA transcripts, as proposed by *Delannoy et al. (2011)*, based on the observation that *rpl*2 was correctly spliced in the case of a *mat*K gene deletion in the *Rhizanthella gardneri* plastome.

In the *E. paucisquama* plastome, only two protein-coding genes, *acc*D and *clp*P. involved in functions other than translation, were retained, similar to other non-photosynthetic plant species with a highly reduced plastome (Fig. 3), such as *Epipogium* (*Schelkunov et al., 2015*). The *acc*D gene plays an important role in fatty acid biosynthesis (*Bryant et al., 2011*), and the loss of *acc*D from the plastome of heterotrophic plants is a rare event (*Lim et al., 2016*). The *clp*P gene encodes a subunit of the Clp protease (or ATP-dependent caseinolytic protease), involved in the regulation of protein turnover and processing and has also been linked to isoprenoid and tetrapyrrole biosynthesis and fibrillin (lipid-body stabilizing molecules) (*Kim et al., 2009*; *Stanne et al., 2009*). Notably, both of these genes participate in the regulation of lipid metabolism, supporting the role of lipids in sustaining colonization by mutualistic mycorrhizal and parasitic fungi (*Jiang et al., 2017*). In the *E. paucisquama* plastome, these two genes were duplicated along with the expansion of the IR region. Expanded IR has been observed in other mycoheterotrophs (*Naumann et al., 2016*; *Joyce et al., 2019*) and autotrophs (*Zhu et al., 2016*; *Sinn et al., 2018*), which leads to a decelerated substitution rates for the genes translocated from the SC into the IR region in fern (*Li et al., 2016*). Biased gene conversion, which means new mutations corrected back to ancestral states preferentially (*Birky & Walsh, 1992*), is hypothesized as the reason behind the lower rate of nucleotide substitution in the IR region (*Wu & Chaw, 2015*; *Li et al., 2016*; *Zhu et al., 2016*). The expansion of the IR region in *E. paucisquama* may be a mechanism for the prevention of immoderate mutations to retain the functionality of genes translocated to the IR region. The movement of genes in SC region to the IR region could be an advantageous move that is, selected for in *E. paucisquama*. This coincides with the view that the degradation of
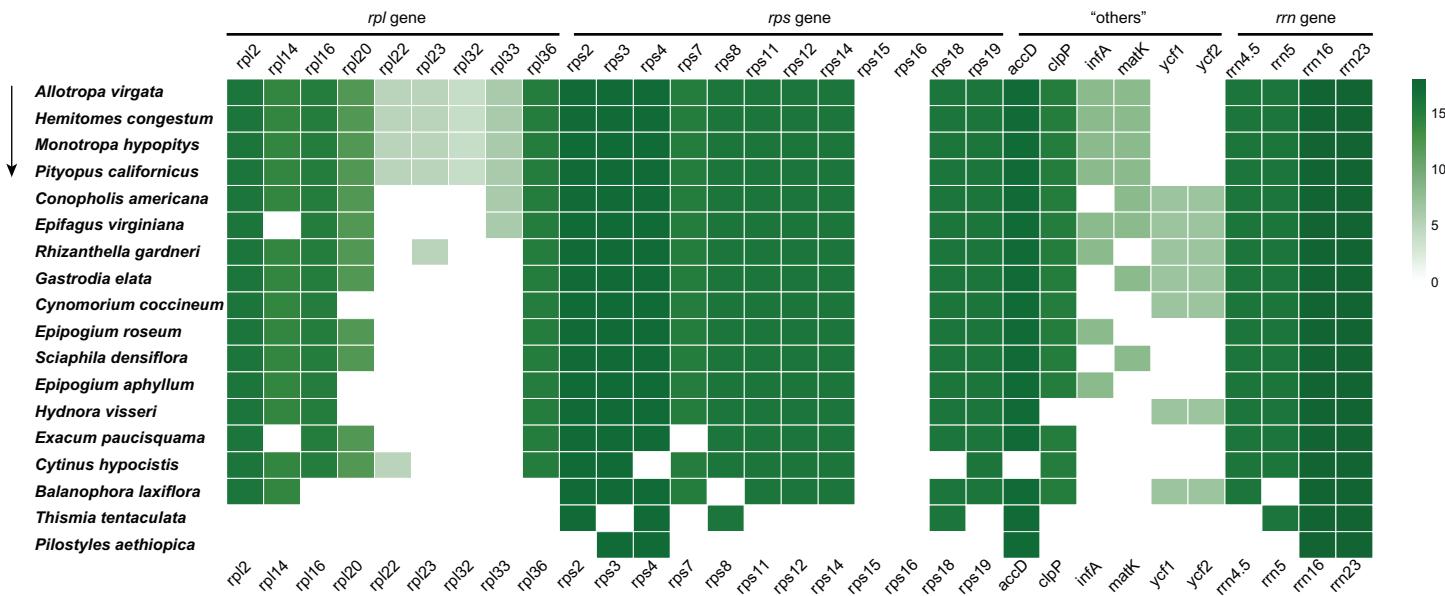
Figure 3 Summary of the *rpl*, *rps*, "others" (*acc*D, *clp*P, *inf*A, *mat*K, *ycf*1 and *ycf*2), and *rrn* genes in the plastome of *E. paucisquama* and 17 other fully heterotrophic species. The arrow indicates the degree of gene loss. Green boxes represent the retained putatively functional genes; a deeper green color indicates a higher number of species retaining the genes. Full-size 🖼 DOI: 10.7717/peerj.9157/fig-3

plastomes in mycoheterotrophic species presents a highly lineage-specific pattern (*Feng et al., 2016*).

CONCLUSION

We report the first plastid genome of mycoheterotrophic species in the family Gentianaceae sequenced to date. The *E. paucisquama* plastome showed extensive gene losses and contained only 21 putative functional genes (15 protein-coding genes, four rRNA genes and two tRNA genes). Some "housekeeping" genes, such as *ycf*1, *ycf*2, *mat*K and many tRNA genes, were lost in the *E. paucisquama* plastome. More than 80% of the plastome of *E. paucisquama* is IR regions, and these regions harbor most of the remaining genes. Our results provide valuable information for the comparative evolutionary analyses of plastomes of heterotrophic species belonging to different phylogenetic lineages.

ACKNOWLEDGEMENTS

ADDITIONAL INFORMATION AND DECLARATIONS

Funding

Competing Interests

The authors declare that they have no competing interests.

Author Contributions

- Zhanghai Li performed the experiments, analyzed the data, prepared figures and/or tables, authored or reviewed drafts of the paper, and approved the final draft.
- Xiao Ma performed the experiments, prepared figures and/or tables, and approved the final draft.
- Yi Wen analyzed the data, prepared figures and/or tables, and approved the final draft.
- Sisi Chen performed the experiments, prepared figures and/or tables, and approved the final draft.
- Yan Jiang performed the experiments, prepared figures and/or tables, and approved the final draft.
- Xiaohua Jin conceived and designed the experiments, authored or reviewed drafts of the paper, and approved the final draft.

Data Availability

The following information was supplied regarding data availability:
The complete plastid genome sequence of Exacum paucisquama is available as a Supplemental File and at GenBank: MN067514.

Supplemental Information

Supplemental information for this article can be found online at http://dx.doi.org/10.7717/peerj.9157#supplemental-information.

REFERENCES

Alkatib S, Fleischmann TT, Scharff LB, Bock R. 2012. Evolutionary constraints on the plastid tRNA set decoding methionine and isoleucine. *Nucleic Acids Research* **40(14)**:6713–6724 DOI 10.1093/nar/gks350.

Barbrook AC, Howe CJ, Purton S. 2006. Why are plastid genomes retained in non-photosynthetic organisms? *Trends in Plant Science* **11(2)**: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. *American Journal of Botany* **99(9)**: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(12)**:3095–3112 DOI 10.1093/molbev/msu252.

Barrett CF, Kennedy AH. 2018. Plastid genome degradation in the endangered, mycoheterotrophic, North American orchid *Hexalectris warnockii*. *Genome Biology and Evolution* **10(7)**:1657–1662 DOI 10.1093/gbe/evy107.

Barrett CF, Wicke S, Sass C. 2018. Dense infraspecific sampling reveals rapid and independent trajectories of plastome degradation in a heterotrophic orchid complex. *New Phytologist* **218(3)**:1192–1204 DOI 10.1111/nph.15072.

Birky CW, Walsh J. 1992. Biased gene conversion, copy number, and apparent mutation rate differences within chloroplast and bacterial genomes. *Genetics* **130**:677–683.

Bryant N, Lloyd J, Sweeney C, Myouga F, Meinke D. 2011. Identification of nuclear genes encoding chloroplast-localized proteins required for embryo development in *Arabidopsis*. *Plant Physiology* **155(4)**:1678–1689 DOI 10.1104/pp.110.168120.

Darling AE, Mau B, Perna NT. 2010. progressiveMauve: multiple genome alignment with gene gain, loss and rearrangement. *PLOS ONE* **5(6)**:e11147 DOI 10.1371/journal.pone.0011147.

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(7)**:2077–2086 DOI 10.1093/molbev/msr028.

Dierckxsens N, Mardulyn P, Smits G. 2016. NOVOPlasty: de novo assembly of organelle genomes from whole genome data. *Nucleic Acids Research* **45(4)**:e18 DOI 10.1093/nar/gkw955.

Feng YL, Wicke S, Li JW, Han Y, Lin CS, Li DZ, Zhou TT, Huang WC, Huang LQ, Jin XH. 2016. Lineage-specific reductions of plastid genomes in an orchid tribe with partially and fully mycoheterotrophic species. *Genome Biology and Evolution* **8(7)**:2164–2175 DOI 10.1093/gbe/evw144.

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.

Hanaoka M, Kanamaru K, Fujiwara M, Takahashi H, Tanaka K. 2005. Glutamyl-tRNA mediates a switch in RNA polymerase use during chloroplast biogenesis. *EMBO Reports* **6(6)**:545–550 DOI 10.1038/sj.embor.7400411.

Jahn D, Verkamp E, So D. 1992. Glutamyl-transfer RNA: a precursor of heme and chlorophyll biosynthesis. *Trends in Biochemical Sciences* **17(6)**:215–218 DOI 10.1016/0968-0004(92)90380-R.

Jiang YN, Wang WX, Xie QJ, Liu N, Liu LX, Wang DP, Zhang XW, Yang C, Chen XY, Tang DZ, Wang ET. 2017. Plants transfer lipids to sustain colonization by mutualistic mycorrhizal and parasitic fungi. *Science* **356(6343)**:1172–1175 DOI 10.1126/science.aam9970.

Joyce EM, Crayn DM, Lam VK, Gerelle WK, Graham SW, Nauheimer L. 2019. Evolution of *Geosiris* (Iridaceae): historical biogeography and plastid-genome evolution in a genus of non-photosynthetic tropical rainforest herbs disjunct across the Indian Ocean. *Australian Systematic Botany* **31(6)**:504–522 DOI 10.1071/SB18028.

Kearse M, Moir R, Wilson A, Stones-Havas S, Cheung M, Sturrock S, Buxton S, Cooper A, Markowitz S, Duran C, Thierer T, Ashton B, Meintjes P, Drummond A. 2012. Geneious basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics* **28(12)**:1647–1649 DOI 10.1093/bioinformatics/bts199.

Kim YK, Jo S, Cheon SH, Joo MJ, Hong JR, Kwak MH, Kim KJ. 2019. Extensive losses of photosynthesis genes in the plastome of a mycoheterotrophic orchid, *Cyrtosia septentrionalis* (Vanilloideae: Orchidaceae). *Genome Biology and Evolution* **11(2)**:565–571 DOI 10.1093/gbe/evz024.

Kim J, Rudella A, Rodriguez VR, Zybailov B, Olinares PDB, Van Wijk KJ. 2009. Subunits of the plastid ClpPR protease complex have differential contributions to embryogenesis, plastid biogenesis, and plant development in *Arabidopsis*. *Plant Cell* **21(6)**:1669–1692 DOI 10.1105/tpc.108.063784.

Lam VK, Soto Gomez M, Graham SW. 2015. The highly reduced plastome of mycoheterotrophic *Sciaphila* (Triuridaceae) is colinear with its green relatives and is under strong purifying selection. *Genome Biology and Evolution* **7(8)**:2220–2236 DOI 10.1093/gbe/evv134.

Leake JR. 1994. The biology of myco-heterotrophic ('saprophytic') plants. *New Phytologist* **127(2)**:171–216 DOI 10.1111/j.1469-8137.1994.tb04272.x.

Li FW, Kuo LY, Pryer KM, Rothfels CJ. 2016. Genes translocated into the plastid inverted repeat show decelerated substitution rates and elevated GC content. *Genome Biology and Evolution* **8(8)**:2452–2458 DOI 10.1093/gbe/evw167.

Li ZH, Ma X, Wang DY, Li YX, Wang CW, Jin XH. 2019. Evolution of plastid genomes of *Holcoglossum* (*Orchidaceae*) with recent radiation. *BMC Evolutionary Biology* **19(1)**:63 DOI 10.1186/s12862-019-1384-5.

Li JL, Wang S, Jing Y, Wang L, Zhou SL. 2013. A modified CTAB protocol for plant DNA extraction. *Chinese Bulletin of Botany* **48(1)**:72–78.

Lim GS, Barrett CF, Pang CC, Davis JI. 2016. Drastic reduction of plastome size in the mycoheterotrophic *Thismia tentaculata* relative to that of its autotrophic relative *Tacca chantrieri*. *American Journal of Botany* **103(6)**:1129–1137 DOI 10.3732/ajb.1600042.

Logacheva MD, Schelkunov MI, Nuraliev MS, Samigullin TH, Penin AA. 2014. The plastid genome of mycoheterotrophic monocot *Petrosavia stellaris* exhibits both gene losses and multiple rearrangements. *Genome Biology and Evolution* **6(1)**:238–246 DOI 10.1093/gbe/evu001.

Logacheva MD, Schelkunov MI, Penin AA. 2011. Sequencing and analysis of plastid genome in mycoheterotrophic orchid *Neottia nidus-avis*. *Genome Biology and Evolution* **3(9)**:1296–1303 DOI 10.1093/gbe/evr102.

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(W1)**: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.

Merckx VSFT, Freudenstein JV. 2010. Evolution of mycoheterotrophy in plants: a phylogenetic perspective. *New Phytologist* **185(3)**:605–609 DOI 10.1111/j.1469-8137.2009.03155.x.

Merckx VSFT, Freudenstein JV, Kissling J, Christenhusz MJM, Stotler RE, Crandall-Stotler B, Wickett NJ, Rudall PJ, De Kamer HM, Maas PJM. 2013a. Taxonomy and classification. In: Merckx VSFT, ed. *Mycoheterotrophy: The Biology of Plants Living on Fungi*. New York: Springer, 19–101.

Merckx VSFT, Mennes CB, Peay KG, Geml J. 2013b. Evolution and diversification. In: Merckx VSFT, ed. *Mycoheterotrophy: The Biology of Plants Living on Fungi*. New York: Springer, 215–244.

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 C-I, 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(4)**:793–803 DOI 10.1093/molbev/msu051.

Naumann J, Der JP, Wafula EK, Jones SS, Wagner ST, Honaas LA, Ralph PE, Bolin JF, Maass E, Neinhuis C, Wanke S, DePamphilis CW. 2016. Detecting and characterizing the highly divergent plastid genome of the nonphotosynthetic parasitic plant *Hydnora visseri* (Hydnoraceae). *Genome Biology and Evolution* **8(2)**:345–363 DOI 10.1093/gbe/evv256.

Palmer JD, Thompson WF. 1982. Chloroplast DNA rearrangements are more frequent when a large inverted repeat sequence is lost. *Cell* **29(2)**:537–550 DOI 10.1016/0092-8674(82)90170-2.

Petersen G, Zervas A, Pedersen HAE, Seberg O. 2018. Contracted genes and dwarfed plastome in mycoheterotrophic *Sciaphila thaidanica* (Triuridaceae, Pandanales). *Genome Biology and Evolution* **10(3)**:976–981 DOI 10.1093/gbe/evy064.

Rogalski M, Karcher D, Bock R. 2008. Superwobbling facilitates translation with reduced tRNA sets. *Nature Structural & Molecular Biology* **15(2)**:192–198 DOI 10.1038/nsmb.1370.

Schelkunov MI, Shtratnikova VY, Nuraliev MS, Selosse MA, Penin AA, Logacheva MD. 2015. Exploring the limits for reduction of plastid genomes: a case study of the mycoheterotrophic orchids *Epipogium aphyllum* and *Epipogium roseum*. *Genome Biology and Evolution* **7(4)**:1179–1191 DOI 10.1093/gbe/evv019.

Sinn BT, Sedmak DD, Kelly LM, Freudenstein JV. 2018. Total duplication of the small single copy region in the angiosperm plastome: rearrangement and inverted repeat instability in *Asarum*. *American Journal of Botany* **105(1)**:71–84 DOI 10.1002/ajb2.1001.

Stanne TM, Sjögren LL, Koussevitzky S, Clarke AK. 2009. Identification of new protein substrates for the chloroplast ATP-dependent Clp protease supports its constitutive role in *Arabidopsis*. *Biochemical Journal* **417(1)**:257–269 DOI 10.1042/BJ20081146.

Struwe L. 2014. Classification and evolution of the family gentianaceae. In: Rybczynski JJ, Davey MR, Mikula A, eds. *The Gentianaceae-Characterization and Ecology*. Vol. 1. Berlin: Springer, 13–35.

Tillich M, Lehwark P, Pellizzer T, Ulbricht-Jones ES, Fischer A, Bock R, Greiner S. 2017. GeSeq—versatile and accurate annotation of organelle genomes. *Nucleic Acids Research* **45(W1)**:W6–W11 DOI 10.1093/nar/gkx391.

Wicke S, Muller 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.

Wu CS, Chaw SM. 2015. Evolutionary stasis in cycad plastomes and the first case of plastome GC-biased gene conversion. *Genome Biology and Evolution* **7(7)**:2000–2009 DOI 10.1093/gbe/evv125.

Yuan Y, Jin X, Liu J, Zhao X, Zhou J, Wang X, Wang D, Lai C, Xu W, Huang J, Zha L, Liu D, Ma X, Wang L, Zhou M, Jiang Z, Meng H, Peng H, Liang Y, Li R, Jiang C, Zhao Y, Nan T, Jin Y, Zhan Z, Yang J, Jiang W, Huang L. 2018. The *Gastrodia elata* genome provides insights into plant adaptation to heterotrophy. *Nature Communications* **9(1)**:1615 DOI 10.1038/s41467-018-03423-5.

Zerbino DR, Birney E. 2008. Velvet: algorithms for de novo short read assembly using de Bruijn graphs. *Genome Research* **18(5)**:821–829 DOI 10.1101/gr.074492.107.

Zhu AD, Guo WH, Gupta S, Fan WS, Mower JP. 2016. Evolutionary dynamics of the plastid inverted repeat: the effects of expansion, contraction, and loss on substitution rates. *New Phytologist* **209(4)**:1747–1756 DOI 10.1111/nph.13743.

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.