# Mitochondrial genome of the nonphotosynthetic mycoheterotrophic plant *Hypopitys monotropa*, its structure, gene expression and RNA editing

Viktoria Yu Shtratnikova[1], Mikhail I. Schelkunov[2,3], Aleksey A. Penin[3] and Maria D. Logacheva[2]

[1] A. N. Belozersky Institute of Physico-Chemical Biology, Lomonosov Moscow State University, Moscow, Russia
[2] Skolkovo Institute of Science and Technology, Moscow, Russia
[3] Laboratory of Plant Genomics, Institute for Information Transmission Problems of the Russian Academy of Sciences, Moscow, Russia

## ABSTRACT

Heterotrophic plants—plants that have lost the ability to photosynthesize—are characterized by a number of changes at all levels of organization. Heterotrophic plants are divided into two large categories—parasitic and mycoheterotrophic (MHT). The question of to what extent such changes are similar in these two categories is still open. The plastid genomes of nonphotosynthetic plants are well characterized, and they exhibit similar patterns of reduction in the two groups. In contrast, little is known about the mitochondrial genomes of MHT plants. We report the structure of the mitochondrial genome of *Hypopitys monotropa*, a MHT member of Ericaceae, and the expression of its genes. In contrast to its highly reduced plastid genome, the mitochondrial genome of *H. monotropa* is larger than that of its photosynthetic relative *Vaccinium macrocarpon*, and its complete size is ~810 Kb. We observed an unusually long repeat-rich structure of the genome that suggests the existence of linear fragments. Despite this unique feature, the gene content of the *H. monotropa* mitogenome is typical of flowering plants. No acceleration of substitution rates is observed in mitochondrial genes, in contrast to previous observations in parasitic non-photosynthetic plants. Transcriptome sequencing revealed the trans-splicing of several genes and RNA editing in 33 of 38 genes. Notably, we did not find any traces of horizontal gene transfer from fungi, in contrast to plant parasites, which extensively integrate genetic material from their hosts.

## INTRODUCTION

*Hypopitys monotropa* (Ericacae, Monotropoideae) is a nonphotosynthetic plant that obtains carbon from fungi in ectomycorrhizal relationships with tree roots (*Bjorkman, 1960*). In contrast to most other mycoheterotrophic (MHT) plants, which are very rare

and/or very narrowly distributed, Monotropoideae, including *H. monotropa*, are quite widespread and are associated with old-growth conifer forests. Thus, *H. monotropa* is used as a model system in studies of plant-mycorrhizal associations and the developmental biology of MHT plants (*Olson, 1993*, *1990*). Recent advances in DNA sequencing have expanded research on mycoheterotrophs into genomics. To date, most attention has been focused on the plastid genomes of MHT plants, which are highly reduced in size and gene content, including a complete absence or pseudogenization of genes of the photosynthesis electron transport chain (for a review, see *Graham, Lam & Merckx, 2017*). Thus, the MHT lifestyle strongly affects plastids, but what about the mitochondrial genome?

In contrast to animals, in which mitochondrial genomes are usually conserved in size and gene content across large taxonomic groups, in plants, these genomes are highly variable and may be very dissimilar even in closely related species. The size of the angiosperm mitogenome ranges from 66 Kb in the hemiparasitic mistletoe *Viscum scurruloideum* (*Skippington et al., 2015*) and 222 Kb in the autotrophic *Brassica napus* (*Handa, 2003*) to more than 11 Mb in *Silene noctiflora*. Despite such huge variations in size, a large fraction of mitochondrial genes—those that encode the components of the oxidative phosphorylation chain complexes and proteins involved in the biogenesis of these complexes—are stable in content and show very low sequence divergence. More variation exists in the group of genes involved in translation, that is, ribosomal proteins and transfer RNAs (*Adams et al., 2002*; *Gualberto et al., 2014*), presumably due to the transfer to the nuclear genome that occurred in several plant lineages or the non-essentiality of these genes. In addition, many plant mitochondrial genomes carry open reading frames (ORFs) that potentially encode functional proteins (*Qiu et al., 2014*); such ORFs are highly lineage specific. Nonphotosynthetic plants are divided into two large groups—those that are parasitic on other plants and those that are mycoheterotropic. To date, only a few complete mitochondrial genomes of nonphotosynthetic plants have been characterized, and most of them belong to parasitic species. A comparative analysis of the mitogenomes of several parasitic, hemiparasitic and autotrophic Orobanchaceae (*Fan et al., 2016*) showed that the gene content does not depend on trophic specialization in the family range. The mitogenomes of two non-related lineages of parasitic plants, Rafflesiaceae and Cynomoriaceae, also do not show a reduction in gene content, and they provide an example of massive HGT from other plants (including but not limited to their hosts). This is not, however, a trait that is unique to parasitic plants—for example, in *Amborella trichopoda*, an autotrophic plant from the basal angiosperms, a large fraction of the mitochondrial genome was acquired from green algae, mosses, and other angiosperms (*Rice et al., 2013*). In contrast, in the hemiparasitic plant *V. scurruloideum*, the mitogenome is drastically reduced in size and gene content, and it lacks all nine *nad* genes and *matR* (*Skippington et al., 2015*). The mitogenomes of other *Viscum* species are not reduced in length but exhibit a reduced gene content, similar to that of *V. scurruloideum* (*Petersen et al., 2015*; but see *Skippington et al., 2017*). The available sampling is obviously insufficient even for parasitic plants; regarding mycoheterotrophs, the data are almost completely lacking. Only two mitochondrial genomes of MHT plants

Shtratnikova et al. (2020), *PeerJ*, DOI 10.7717/peerj.9309

2/23

have been characterized by date—those of the orchid *Gastrodia elata* (*Yuan et al., 2018*) and *Epirixanthes elongata* (Polygalaceae) (*Petersen et al., 2019*). The former was characterized as a part of a genome sequencing project and is not completely assembled, being represented by 19 contigs. The *G. elata* mitogenome is large (~1.3 Mb), while that of *E. elongata* is much smaller (~0,4 Mb). However, both of these studies lack a comparative analysis with autotrophic relatives that would allow to highlight the changes associated with mycoheterotrophy.

In this context, we set the following objectives: (1) to characterize the structure and gene content of the mitochondrial genome of *H. monotropa*, (2) to estimate the possibility of horizontal gene transfer (HGT) from fungi, and (3) to study mitochondrial gene expression and RNA editing.

The photosynthetic plant with a characterized mitochondrial genome phylogenetically closest to *H. monotropa* is cranberry, *V. macrocarpon* (*Fajardo et al., 2014*). In this study, we used *V. macrocarpon* as a basis for comparative analysis aimed at identifying the patterns (if any) of mitochondrial genome changes associated with mycoheterotrophy.

## MATERIALS AND METHODS

### Sample collection and sequencing

The procedures for sample collection, DNA and RNA library preparation for most datasets (DNA shotgun and mate-pair libraries and an inflorescence transcriptome library) and sequencing were described in the reports of *Logacheva et al. (2016)* and *Schelkunov, Penin & Logacheva (2018)*. In addition to the data generated in previous studies, we used four new datasets that represent the transcriptomes of the anthers and ovules of *H. monotropa*. The samples were collected in the same location as previous samples but in 2018. *H. monotropa* is not an endangered or threatened plant species; thus, no specific permissions were required for its collection. RNA was extracted using an RNEasy kit (Qiagen, Hilden, Germany) with the addition of the Plant RNA Isolation Aid reagent (Thermo Fisher, Waltham, MA, USA) to the lysis buffer. The removal of ribosomal RNA was performed using a Ribo-Zero plant leaf kit (Illumina, San Diego, CA, USA), and further sample preparation was performed using a NEBNext RNA library preparation kit (New England Biolabs, Ipswich, MA, USA). The libraries were sequenced on the HiSeq 4000 platform (Illumina, San Diego, CA, USA) in 150-nt paired-end mode at the Skoltech Genomics Core Facility. The reads were deposited in the NCBI Sequence Read Archive under BioProject PRJNA573526.

### Assembly

Read trimming and assembly were performed as described in the report of *Logacheva et al. (2016)*. The contig coverage was determined by mapping all reads in CLC Genomics Workbench v. 7.5.1 (https://www.qiagenbioinformatics.com/products/clc-genomics-workbench/), requiring at least 80% of a read's length to align with at least 98% sequence similarity. To find contigs corresponding to the mitochondrial genome, we performed BLASTN and TBLASTX alignment of the *V. macrocarpon* genome (GenBank accession NC_023338) against all contigs in the assembly. BLASTN and TBLASTX from the

BLAST 2.3.0+ suite (*Camacho et al., 2009*) were used, and the alignment was performed with a maximum allowed *e*-value of $10^{-5}$. Low-complexity sequence filters were switched off to avoid missing genes with an extremely high or low GC content. A contig that corresponded to the plastid genome of *H. monotropa* also aligned to the mitochondrial genome of *V. macrocarpon* because of the presence of inserts from the plastid genome in the mitochondrial genome. The complete sequence of the *H. monotropa* plastome is known from our previous study (*Logacheva et al., 2016*) and was excluded from further analyses. Contigs showing coverage of less than 10 or those that exhibited matches only to non-mitochondrial sequences when aligned to the NCBI nt database by BLASTN (*e*-value $10^{-3}$) were considered nuclear.

After the procedures described above, seven contigs remained, with lengths of 255, 5,651, 20,627, 102,868, 106,298, 234,240 and 309,413 bp and coverages of 37, 89, 93, 110, 96, 107 and 108 bp, respectively. The lower coverage of the smallest contig was presumably an artifact caused by its length. To understand their order in the mitochondrial genome, we mapped the mate-pair reads from the library with the largest insert size (8,705 ± 2,537 bp) to the contigs and investigated the mate-pair links. The mapping was performed by using CLC Genomics Workbench and 100% of a read's length was required to align to a contig with a sequence similarity of 100%, to minimize the amount of incorrectly mapped reads. Mate-pair links were visualized by using Circos 0.67–5 (*Krzywinski et al., 2009*) and investigated manually. In one contig, the mate-pair links spanned the entire length from the end to the start; thus, the contig corresponded to a circular chromosome. The other six contigs were connected into a linear sequence, whose structure is described in the Results section. Gaps between the contigs were closed by GapFiller (*Boetzer & Pirovano, 2012*), which was run with the following parameters: -m 30 -o 2 -r 0.8 -n 50 -d 100000 -t 100 -g 0 -T 4 -i 100000. The source code of GapFiller was rewritten to allow it to use Bowtie2 as the read mapper. Additionally, we tried to extend the edges of the linear chromosome by using GapFiller to check that there were no assembly forks (sites with two alternative extensions). The absence of such forks indicates that the linearity of the sequence is not a consequence of an assembly problem. Instead, during extension, the coverage gradually decreased to zero, implying a linear chromosome with lengths varying between different copies of the chromosome. To check whether the two chromosomes were assembled correctly, we mapped the reads from all three sequencing libraries by using CLC Genomics Workbench v. 7.5.1, requiring at least 90% of a read's length to align with at least 99% sequence similarity and checked that the coverage was approximately uniform along the whole length of the chromosomes. As observed during the analysis by GapFiller, we found that the coverage gradually decreased at the edges of the linear chromosome, indicating that different copies of the chromosome in a plant exhibit different lengths of the terminal regions. For the sake of precision, we elongated the terminal regions such that the coverage dropped to 0; thus, the length of the linear chromosome approximately corresponded to the maximal possible length among all copies of the chromosome in the sequenced plant. Additionally, we mapped the reads from the mate-pair library with the largest insert sizes, requiring 100% of a read's length to align with a sequence similarity of 100%, and we investigated the

distribution of mean insert sizes over each genomic position and the number of mate pairs than spanned each position. These values were approximately uniform along both chromosomes (Fig. S1), suggesting that the assembly was correct.

## Mitogenome annotation

Initial annotation was performed using Mitofy (*Alverson et al., 2010*) with further manual correction. To verify exon boundaries, we mapped RNA-seq reads to the sequences of the chromosomes by using CLC Genomics Workbench v. 7.5.1, requiring at least 80% of a read's length to map with at least 90% sequence similarity. The relaxed setting for the percentage of read length to be mapped allows the mapping reads to span splice junctions, and the relaxed setting for sequence similarity allows the mapping of reads to regions with a high density of RNA editing sites. The alignment was visualized in CLC Genomics Workbench 7.5.1.

To check if there were any protein-coding genes that were missed, we visually inspected the alignment to search for unannotated regions in the chromosomes with high coverage by RNA-seq reads. This search did not add new genes.

The search for genes encoding selenocysteine tRNAs (tRNA-Sec) was performed for *H. monotropa* and *V. macrocarpon* by using an online version of Secmarker 0.4 (*Santesmasses, Mariotti & Guigó, 2017*) with the default parameters. SECIS element prediction was performed by using an online version of SECISEARCH3 (*Mariotti et al., 2013*) that was current as of 5 May 2017 with the default parameters. Assembled and annotated sequences of *H. monotropa* mitochondrial genome are deposited in the NCBI Genbank under accession numbers MK990822 and MK990823.

Gene Ontology (GO) terms were assigned to the genes of the mitochondrial genome by the NetGO 2.0 server (*You et al., 2019*). Genes with their GO terms are provided in Table S2. GO terms with prediction scores <= 0.6 are considered of low confidence in the manual of NetGO and were therefore discarded.

## RNA editing analysis

To obtain information on RNA editing sites, we artificially spliced gene sequences and mapped RNA-seq reads to the sequences with the parameters indicated above for RNA-seq read mapping. Variant calling was performed by using CLC Genomics Workbench 7.5.1. We considered a site in a CDS to be edited if it was covered by at least 20 reads and at least 10% of the reads in the corresponding positions differed from the genomic sequence.

To estimate the fraction of RNA editing sites that change the amino acid sequences of the corresponding proteins, we used a measure that has been used in several previous studies (*Chen, 2013*; *Xu & Zhang, 2014*; *Zhu et al., 2014*) but has not yet been given a name. Here, we refer to this measure as dEN/dES. It is calculated similarly to dN/dS, but while dN/dS represents the ratio of nonsynonymous and synonymous substitutions per nonsynonymous and synonymous sites, dEN/dES ("E" stands for "editing" here) represents the ratio of nonsynonymous and synonymous RNA editing events per nonsynonymous and synonymous sites.

To compare the RNA editing sites in the mitochondrial mRNAs of *H. monotropa* with those in the mitochondrial mRNAs of other species, we used data on RNA editing in *Arabidopsis thaliana* (*Giege & Brennicke, 1999*) and *Oryza sativa* (*Notsu et al., 2002*). These data are incorporated in the annotations of *A. thaliana* and *O. sativa* in GenBank, with accession numbers NC_037304.1 and BA000029.3, respectively.

## Search for sequences transferred to the mitogenome

To identify sequences transferred from the plastome to the mitogenome (known as "MIPTs", which stands for MItochondrial Plastid Transfers), we aligned the plastid genes of *Camellia sinensis* to the sequences of the *H. monotropa* mitogenome. *Camellia sinensis* was chosen because its plastome contains a complete set of typical plastid genes and is phylogenetically closest to *H. monotropa* (as of May 2017) among all sequenced plants with such a complete set. We searched for transferred genes and not intergenic regions because we were interested in transfers that may have initially been functional. The proteins of *Camellia* were aligned to the sequences of the mitochondrial chromosomes by using TBLASTN with a maximum allowed *e*-value of $10^{-3}$. tRNA and rRNA coding genes were similarly aligned to the sequences of the chromosomes by using BLASTN. The matching regions of the chromosomes were then aligned by BLASTX to the NCBI nr database (for regions that matched *Camellia* proteins) and by BLASTN to the NCBI nt database (for regions that matched *Camellia* RNA coding genes). If the best matches in the database for a region in the mitogenome were to sequences belonging to plastomes, that region was considered a MIPT. To calculate the number of frameshift-inducing indels and nonsense mutations in the transferred regions, these regions were considered together with their 200 bp-long flanking sequences on both ends and aligned to homologous genes from *C. sinensis*. The alignment was performed by using BLASTN with the default parameters. The resultant alignments were inspected by eye.

To search for possible horizontal gene transfers from fungi, the mitogenome sequences were split into windows of 500 bps each, with a step size of 50 bps. These windows were aligned by BLASTX to the NCBI nr database and by BLASTN to the NCBI nt database. The maximum *e*-value allowed for the matches was $10^{-5}$. The regions that yielded significant matches to fungi were extracted and aligned back to the NCBI nr database to determine whether they were uniquely shared by *H. monotropa* and fungi.

The analysis of the origin of the *cox1* intron in *H. monotropa* was performed as follows. First, we aligned its sequence to the NCBI nt database by using BLASTN online with the default parameters, taking the best 100 matching sequences. To these sequences, we added the sequences of all *cox1* introns from Ericales that were not among the best 100 matches. Then, we aligned these sequences by using the MAFFT server at https://mafft.cbrc.jp/alignment/server/ (*Katoh, Rozewicki & Yamada, 2019*) with the default parameters. The unrooted phylogenetic tree of the *cox1* introns was built by using RAxML 8.2.12 (*Stamatakis, 2014*) with 20 starting maximum parsimony trees and the GTR+Gamma model. The required number of bootstrap pseudoreplicates was automatically determined by the extended majority rule consensus tree criterion (the "autoMRE" option).

## Phylogenetic analysis

Genes common to the mitochondrial genomes of 25 seed plants (*atp1, atp4, atp8, atp9, ccmC, cob, cox1, cox2, cox3, matR, nad1, nad2, nad3, nad4, nad4L, nad5, nad6, nad7, nad8, nad9*) were used for the phylogenetic analysis. Their sequences were concatenated and aligned with MAFFT (*Katoh, Rozewicki & Yamada, 2019*). The phylogenetic analysis was performed using RaXML (raxmlGUI v.1.3.1) with nucleotide sequences under the GTR+Gamma substitution model with 1,000 bootstrap replicates. To infer possible horizontal transfers, all protein-coding genes found in the *H. monotropa* mitogenome were aligned and analyzed in the same way as the concatenated gene set.

# RESULTS

## Genome structure and gene content

Genome assembly resulted in two sequences. The GC content of both sequences is ~45%. One of the sequences is circular with a length of 106 Kb (Fig. 1). It does not exhibit any long repeats; the mapping of mate-pair reads showed an absence of pairs with abnormal insert sizes, suggesting that this sequence is not subject to recombination. The second, longer, fragment (Fig. 1) is ~704 Kb long. It was assembled into a linear sequence and exhibits a more complex structure. In particular, it contains several long repeats. As shown in Fig. 2, three pairs of long direct repeats were observed: the beginning of the chromosome and the 80–90 Kb region, the end of the chromosome and the 455–470 Kb region, the 415–420 Kb region and the 495–500 Kb region. Long inverted repeats were also found between the 260–262 kb region and the 657–659 kb region. A peculiar feature is a gradual decrease in the coverage on both of its ends (Fig. S1). This suggests that copies of the mitogenome with different lengths of this repeat coexist in plant cells. Repeats at the ends of the genome are observed in the linear mitochondrial genomes of several fungi, animals and protists (e.g., the works of *Janouškovec et al. (2013)* and *Kayal et al. (2012)*) and play an important role in their replication (*Nosek et al., 1998*). A similar mechanism could mediate the replication of the linear mitochondrial chromosome in *H. monotropa*. There are some mate-pair links between the two chromosomes, which suggests that they may recombine. However, the coverage of the smaller chromosome by mate-pair inserts (Fig. S1) presents no sharp drops; therefore, it is unlikely that recombination is frequent.

To check for the presence of sequences transferred from fungi, we performed a BLAST search against all fungal sequences available from the NCBI. Although a BLASTX search identified several regions exhibiting similarity to hypothetical proteins from the *Rhizophagus irregularis* (Glomeromycota) genome, the same regions show high similarity to other plants as well (Table S1). This indicates that they are not the result of recent HGT mediated by mycoheterotrophy but, rather, are a result of either the ancient integration of fungal sequences into the genome of the common ancestor of *Hypopitys* and other flowering plants or the parallel integration of mobile genetic elements (e.g., mitoviruses) in the fungal and plant genomes. At least one of these regions exhibits
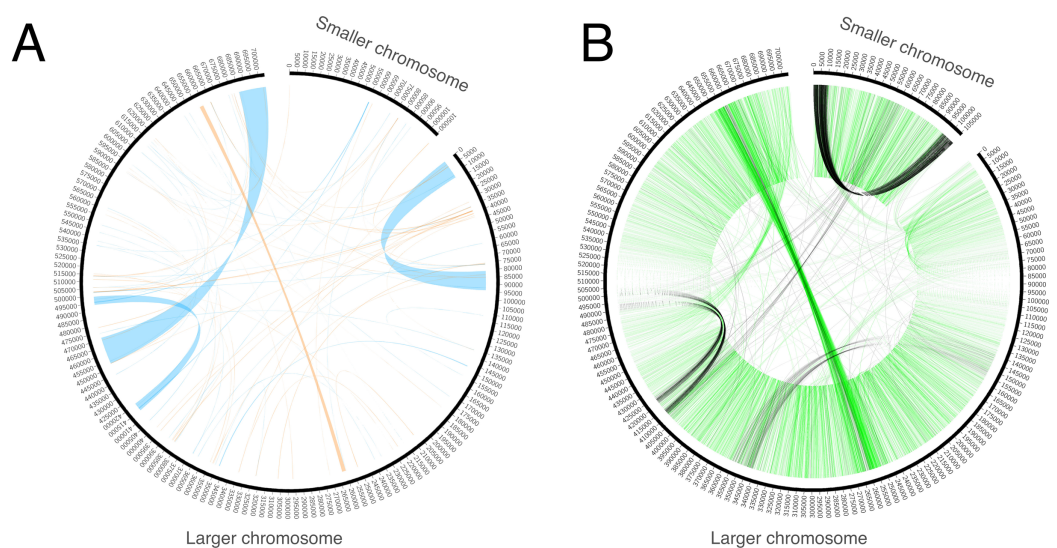
**Figure 1 Maps of the mitochondrial chromosomes of *Hypopitys monotropa*.** Trans-spliced introns are indicated by three colored lines—red in *nad1*, green in *nad2* and blue in *nad5*.

Full-size ◩ DOI: 10.7717/peerj.9309/fig-1

high similarity to an RNA-dependent RNA polymerase gene from a known plant mitovirus, supporting the latter explanation.

Regarding gene content, the circular chromosome contains only two full-length protein-coding genes—*ccmFc* and *cox1*—and three exons of the *nad5* gene, while two other exons are located in a linear fragment. Summary data for the annotation are presented in Table 1. All genes are supported by RNA-seq reads (Fig. S2; Table S2), and most regions with high coverage of RNA-seq reads correspond to annotated genes and their flanking regions (Fig. S3).

**Figure 2 Repeats and mate-pair links in the mitochondrial chromosomes of *Hypopitys monotropa*.**
(A) Repeats within and between the chromosomes. Direct repeats are connected by blue lines, inverted repeats are connected by orange ones. (B) "Improper" (with distance that exceeds expected length of the mate pair library or unexpected orientation) mate-pair links that indicate possible chromosome rearrangements. Read pairs with reads oriented in different directions (→ ← or ← →) are colored black and read pairs with reads oriented in the same direction (→ → or ← ←) are colored green. Only one of the two mate-pair libraries, that with the longer insert sizes (8,279 bp on average, standard deviation 2,583 bp), was used to build this diagram. A pair is considered improper if its reads are mapped not in the orientation → ←, or are mapped on different chromosomes, or are mapped in the orientation → ← but are separated by more than 20,000 bp. The green "torus" of closely situated reads in the orientations → → and ← ← comes from mate-pair reads improperly trimmed by NextClip, the tool for mate-pair reads' trimming that we used. Full-size 🖼 DOI: 10.7717/peerj.9309/fig-2

**Table 1 Summary data on the structure and annotation of *H. monotropa* mitogenome.**

| Chromosome | Accession number (NCBI) | Length | Protein-coding genes | rRNA-coding genes | tRNA-coding genes | Pseudogenes | Genes with introns | Genes with cis-splicing | Genes with trans-splicing |
|---|---|---|---|---|---|---|---|---|---|
| Larger (linear) | MK990822 | 704,088 | 35.5 | 3 | 17–18 | 2 | 8.5* | 6 | 2.5 |
| Smaller (circular) | MK990823 | 106,028 | 2.5 | 0 | 1 | 0 | 2.5* | 2 | 0.5 |
| Total | | | 38 | 3 | 18–19 | 1 | 11 | 8 | 3 |

**Note:**
* Fractional number of genes indicates that one gene has exons in both circular and linear fragments.

In general, the gene content of the *H. monotropa* mitogenome is typical of flowering plants. Surprisingly, it is even larger than that of *V. macrocarpon*, a close photosynthetic relative: *atp6* and *rps14* are pseudogenized and *sdh3* and *rps3* are absent in *V. macrocarpon*. There is only one pseudogene in *H. monotropa, rps14*. Pseudogenization or loss of *rps14* is very common in flowering plants ((*Figueroa et al., 1999*)). In addition to the standard gene set, we found a new ORF (ORF671) that hypothetically encodes a 671-aa protein. It shows low similarity to a gene that encodes a hypothetical protein in the mitochondrial genomes of many flowering plants (the closest

match is the hypothetical mitochondrial protein Salmi_Mp020 from *Salvia miltiorrhiza*). There are 4 genes encoding ribosomal proteins of the large subunit and 8 genes of the small subunit. Regarding the RNA component of the ribosome, all three ribosomal RNAs (26S, 18S, 5S) typical of plant mitochondrial genomes are present in *H. monotropa*. The set of tRNAs consists of 17 tRNAs that are typical of the mitogenomes of autotrophic plant species. There are no tRNAs of plastome origin except for tRNA-Gly-GCC. A selenocysteine tRNA (tRNA-Sec) gene and a sequence required for selenocysteine insertion during translation (SECIS element) have been reported in the mitogenome of *V. macrocarpon* (*Fajardo et al., 2014*). However, SecMarker, a novel specialized tool for searching for tRNA-Sec (*Santesmasses, Mariotti & Guigó, 2017*) that presents higher sensitivity and specificity than the tools used by Fajardo and coworkers, did not confirm the presence of a tRNA-Sec gene in the *V. macrocarpon* mitogenome. Thus, we infer that the predictions of a tRNA-Sec gene and a SECIS element in the *V. macrocarpon* mitogenome were false positives. In *H. monotropa*, we also did not find tRNA-Sec genes or SECIS elements. There is a region with moderate (78.8%) similarity to the presumed tRNA-Sec of *V. macrocarpon*. It is located upstream of *ccmC* as well as one of the two presumed tRNA-Sec sequences of *V. macrocarpon*. tRNA genes are known to exhibit highly conserved sequences; in particular, the similarity of other *V. macrocarpon* and *H. monotropa* tRNA genes is 98–100%. This suggests that the presumed tRNA-Sec of *V. macrocarpon* is not a functional gene but a pseudogene of a tRNA-coding gene.

Despite being highly conserved in coding sequences, mitochondrial genes sometimes differ in their intron content. For example, *cox2* may consist of three exons (*D. carota*, *V. vinifera*), two exons (*C. paramensis*, *A. thaliana*), or a single exon (*V. macrocarpon*). In many angiosperm lineages, the *cox1* gene contains a group I intron, presumably acquired via multiple horizontal transfer events (*Cho et al., 1998*; *Sanchez-Puerta et al., 2008*, *2011*), although there is an opposing hypothesis indicating a single HGT event and multiple losses (*Cusimano, Zhang & Renner, 2008*). The *cox1* intron is highly overrepresented in the parasitic plants that have been examined to date (*Barkman et al., 2007*), although the hypothesis that parasitism may serve as a mediator of horizontal intron transfer is not supported by phylogenetic analysis (*Fan et al., 2016*). *V. macrocarpon* lacks an intron in *cox1*, while in *H. monotropa*, an intron is present. Phylogenetic analysis indicates that the intron found in *H. monotropa* has the same origin as that in *Pyrola secunda* (Fig. S4) and was thus presumably vertically inherited from the common ancestor of *Pyrola* and *Hypopitys*. In other genes, *H. monotropa* exhibits the same intron content as *V. macrocarpon*. Three genes—*nad1*, *nad2*, and *nad5*—have trans-spliced transcripts. Notably, the *nad5* exons are located in different chromosomes—exons 1, 2 and 3 are located in the linear chromosome and exons 4 and 5 in the circular one—suggesting a case of interchromosome trans-splicing, which has been reported in other plants that have multichromosomal mitogenomes (*Lloyd Evans et al., 2019*; *Sloan et al., 2012*).

Detailed data on the genes and their intron-exon structure and expression are presented in Table S2.

Shtratnikova et al. (2020), *PeerJ*, DOI 10.7717/peerj.9309

10/23

## Sequences transferred to the mitogenome

In the mitochondrial genome, we found fragments with high similarity to plastid genes, including both genes that are present in the *H. monotropa* plastome (*matK, rps2, rps4*) and genes that have been lost (*rpoB, C1, C2, psbC, ndhJ, B, ycf2*) (Table S3). These fragments were derived from horizontal transfers from plastids (MIPTs; presumably intracellular). Such events are frequent in plant mitochondria (for example, see the works of *Alverson et al. (2011)* and *Grewe et al. (2014)*). In rarer cases, interspecies horizontal transfer from the mitochondrial genomes of other plants harboring inserts from the plastid genome may be a source of MIPTs, which is referred to as foreign MIPTs (*Gandini & Sanchez-Puerta, 2017*). MIPTs in nonphotosynthetic plants are of particular interest. If mitochondrial copies of plastid genes that were lost or pseudogenized in the plastome retain intact ORFs, the reverse switch from heterotrophy to autotrophy is potentially possible. The mitochondrial genome in this case acts as a "refugium" of plastid genes. The *Castilleja paramensis* mitogenome contains 55 full-length or nearly full-length plastid genes, and only approximately half of them are obvious pseudogenes (*Fan et al., 2016*). Currently, the only example of functional mitochondrial genes transferred from plastids is provided by the tRNA genes of plastid origin recruited for protein synthesis in mitochondria (*Joyce & Gray, 1989*). In *H. monotropa*, none of the sequences that originated from plastid protein-coding genes present intact ORFs, retaining less than 50% of the initial length and/or carrying multiple frameshift mutations in most cases (Table S3).

## Mitochondrial gene expression and RNA editing

To gain insights into mitochondrial gene expression and RNA editing and refine the sequence annotation, we sequenced and assembled the transcriptome of *H. monotropa* (*Logacheva et al., 2016*; *Schelkunov, Penin & Logacheva, 2018*). We observed the expression of all annotated protein-coding genes (hereafter, the expression level is defined as the number of mapped read pairs divided by gene length). A minimal expression level (0.1) is observed for the hypothetical protein *ORF671*. Cytochrome *c* maturation factors are also expressed at low levels (3 of 4 genes have expression <1). The genes of Complex I exhibit intermediate expression (1–6). The highest expression is observed for *sdh4, cox3, atp1* and *atp9* (Table S2).

RNA editing is a phenomenon that is widely observed in mitochondrial transcripts. The number and position of editing sites differ greatly from species to species and from gene to gene. In *H. monotropa*, we identified 545 RNA editing sites present in at least one RNA-seq sample (Table S4) under an editing-level threshold of 10% and a coverage threshold of 20×. Most of these editing events are observed in all samples (Fig. 3), and most of them are C–U events. However, we found at least one A–G editing event, which was present in all samples at the level of ~20% (see "Discussion"). The transcripts of 33 genes out of 38 are edited; the maximal RNA editing density (measured as the number of sites per 100 bp) was observed for *ccmB*. The genes of Complex I exhibit different RNA editing densities, ranging from 0.2 for *nad5* to 5.6 for *nad4L*. In *nad4L*, RNA editing restores the start codon; presumably, the same is true for *V. macrocarpon*, in which *nad4L*
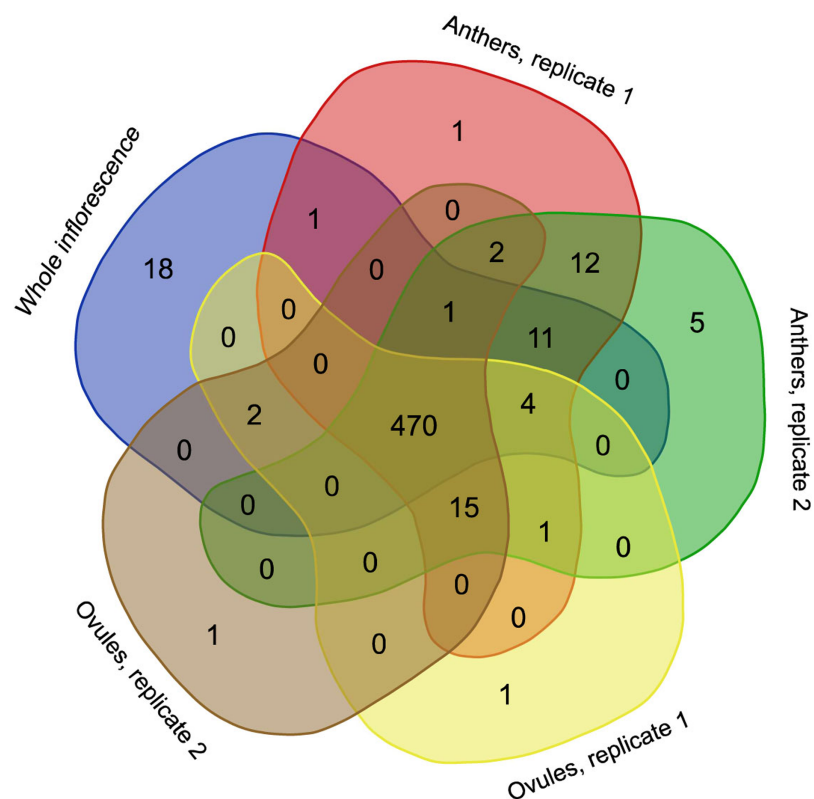
**Figure 3 Venn diagram representing the occurrence of RNA editing events in different RNA-seq samples.** Full-size ⊠ DOI: 10.7717/peerj.9309/fig-3

has been (mis) annotated as a pseudogene. The RNA editing density in the genes of Complexes II and III is low, not exceeding 1.1. Finally, in *atp1*, *rpl16*, *rps12*, *13* and *ORF671*, we did not find any editing events. Several genes with high expression present low levels of RNA editing (*atp1*, *rps12*) and *vise versa* (cytochrome *c* maturation factors), but there is generally no such tendency. The median dEN/dES value across all genes is 1.23. There are no edited stop codons (which is expected because most of the editing events are C–U events and stop codons lack C; however, the observation of non-C–U editing potentially enables this), but there are several stop codons that are introduced by editing (Table S5). To compare RNA editing in *H. monotropa* with that in other species, we used data from *A. thaliana* and *O. sativa*, the model plants in which RNA editing has been thoroughly characterized (*Giege & Brennicke, 1999*; *Notsu et al., 2002*). Overall, there are 842 positions that are edited in at least one species among the protein-coding genes common to these three species (Table S6). Only 151 (17.9%) of these positions are edited in all three species. Notably, 37.9% of the positions are that are edited in at least one species contain a T in the species that do not exhibit editing in this position, while only 6.1% contain an A or G. This is congruent with the results of (*Edera, Gandini & Sanchez-Puerta, 2018*), who found that the main pattern of the loss of RNA editing is the replacement of editing sites with thymines. A total of 337 (38.1%) positions contain unedited cytosines.
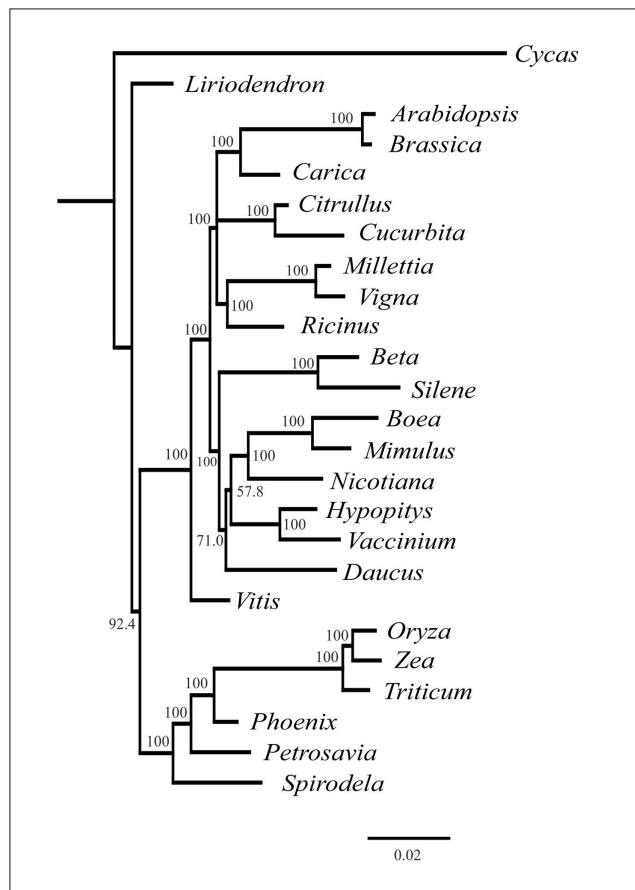
**Figure 4 Phylogenetic tree based on the maximum likelihood analysis of nucleotide sequences of the 20 genes set.** Values above nodes indicate bootstrap support. Branch lengths are proportional to the number of substitutions.                    Full-size ⬜ DOI: 10.7717/peerj.9309/fig-4

## Phylogenetic analysis

As mentioned above, plant mitochondrial genomes are prone to HGT. It is often detected from the incongruence of phylogenetic trees based on different genome regions (*Bergthorsson et al., 2004*; *Cusimano & Renner, 2019*). We performed phylogenetic analysis of individual mitochondrial genes and found that the topologies were similar with regard to the placement of *H. monotropa*—it is always placed together with *V. macrocarpon* (excluding the cases of unresolved nodes) (see Fig. S5). This evidence indicates that no genes were acquired via HGT. The combined tree of all mitochondrial genes shared across 25 seed plant species shows a topology similar to that based on nuclear and plastid genes, with monocots representing a monophyletic group and eudicots divided into two large groups—asterids and rosids. *H. monotropa* is sister to *V. macrocarpon*, and both species are within the asterids, as expected (Fig. 4). Notably, *H. monotropa* genes do not exhibit any increase in substitution rates. The same is true for another MHT plant, *Petrosavia stellaris* (data from (*Logacheva et al., 2014*)). Parasitic plants were previously reported to exhibit elevated substitution rates in all three genomes (*Bromham, Cowman & Lanfear, 2013*); however, a recent study involving broader

sampling of parasitic plants shows that this is not a universal phenomenon (*Zervas, Petersen & Seberg, 2019*).

## DISCUSSION

### Mitogenome structure

All available evidence suggests that the *H. monotropa* mitogenome contains a linear fragment. While linear plasmids are found in the mitochondrial genomes of several plants (e.g., in the work of *Handa, Itani & Sato (2002)*), the linear fragment of the *H. monotropa* mitogenome lacks characteristic features of these plasmids such as a terminal inverted repeat, small size (not over 11 Kb) and genes of RNA and DNA polymerases (reviewed in *Handa (2008)*). A linear structure is typical for the mitochondrial genomes of fungi and protists (*Janouškovec et al., 2013*; *Nosek et al., 1995*). Similar to other mycoheteroprophic plants, *H. monotropa* lives in intimate symbiosis with fungi (*Min et al., 2012*). One might hypothesize that the linear fragment could be the result of either contamination or HGT from fungi. However, these explanations are unlikely to be correct for the following reasons: (1) DNA was isolated from inflorescences, while mycorrhizae exist only in roots; (2) all potential fungal hosts of *H. monotropa* with a known mitogenome possess a circular chromosome and no linear plasmids; (3) there are no fragments with similarity to known fungal genomes; and (4) all genes annotated in the linear fragment are typical plant mitochondrial genes. The observation that a single circular molecule ("master circle") is an oversimplified representation of the plant mitochondrial genome and that these genomes instead exist in vivo as a mixture of circular, linear and branched forms is not novel (see, for example, the report of *Sloan (2013)*). However, a circular structure can usually be observed at the level of sequence assembly due to the presence of multiple repeats. This is not the case for *H. monotropa*, in which internal repeats are also present, but their location and the distribution of mate-pair links do not allow us to reconstitute the master circle (see Fig. 2). This suggests that the diversity of the organization of plant mitochondrial genomes could be even greater than that reported recently (*Kozik et al., 2019*). An alternative explanation for the observed linear structure is that this chromosome contains a very long (longer than the size of mate-pair inserts) highly GC-rich region that cannot be sequenced using Illumina technology. Third-generation sequencing technologies, particularly the Pacific Bioscience SMRT platform, are able to handle such regions (*Loomis et al., 2013*). However, we consider this explanation to be unlikely for two reasons: first, plant mitochondrial genomes are universally biased towards AT nucleotides, and second, we would have observed GC-rich regions at the ends of the linear fragment in this case, which were not found.

A high level of convergence is observed in the gene set of plastid genomes of non-photosynthetic plants, regardless of their parasitic or MHT status. They are characterized by a certain degree of reduction, which usually correlates with the timing of the transition to heterotrophy (for example, see the report of *Samigullin et al. (2016)*) and follows the general gene loss model (*Barrett et al., 2014*). In contrast, the mitogenomes of heterotrophic plants are very diverse in terms of their structure, size and gene content. In *H. monotropa*, the total size of the mitogenome is 810 Kb, which is almost twice as

large as that of *V. macrocarpon*. However, this expansion is unlikely to be associated with heterotrophy. Large mitogenomes are known from heterotrophic plants, particularly from the MHT orchid *Gastrodia elata* (~1.3 Mbp) (*Yuan et al., 2018*) and parasitic *Cynomorium* (1 Mbp) (*Bellot et al., 2016*) and *Lophophytum mirabile* (Balanophoraceae) (~820 Kb) (in the latter case, the size is shaped by fragments horizontally transferred from the host—see below). The other extreme is represented by the highly miniaturized mitogenomes of *V. scurruloideum* (*Skippington et al., 2015*).

## Horizontal gene transfer

HGT is a very common phenomenon in parasitic plants (*Yang et al., 2016*). HGT from host plants into the mitochondrial genome has been shown in Rafflesiaceae (*Xi et al., 2013*), Orobanchaceae (*Yang et al., 2016*), *Cynomorium* (Cynomoriaceae) (*Bellot et al., 2016*), and *Lophophytum mirabile* (Balanophoraceae). In the last case, horizontally transferred homologs have replaced almost all native mitochondrial genes (*Sanchez-Puerta et al., 2017*, *2019*). In contrast, there are no traces of HGT in the *H. monotropa* mitogenome. The nuclear and mitochondrial genomes of the MHT orchid *Gastrodia elata* were recently characterized (*Yuan et al., 2018*), as was the mitochondrial genome of the MHT dicot *Epirixanthes elongata* (*Petersen et al., 2019*), and no HGT was observed in these cases. Although MHT plants are usually regarded alongside the parasitic plants that feed on other plants, the interactions between plants and their hosts are very different in these two cases. Parasitic plants develop specialized structures that integrate into the vascular system of a host plant and channel the flow of nutrients from the host to themselves. Such connections are similar in many aspects to graft junctions and can be the route of transport not only for nutrients but also for high-molecular-weight compounds, including proteins and nucleic acids. The bidirectional transfer of nucleic acids through haustoria in the parasitic plant *Cuscuta pentagona* has been shown (*Kim et al., 2014*). The transfer of RNA from the host is hypothesized to mediate HGT into the parasite genome. In contrast, such transfer is not known to occur in mycorrhizal symbioses. This emphasizes that despite similar heterotrophic strategies, plant parasites and MHT plants are fundamentally different in terms of the interaction with their hosts and, potentially, many other features that stem from this interaction. This calls for the increased sampling of MHT plants in genomic projects and for the development of new model systems representing MHT plants. Many of these species are rare endangered plants with very small distribution ranges; in contrast, *H. monotropa* is widespread and is thus presents potential as a model MHT plant.

## RNA editing

RNA editing is an important characteristic of mitochondrial gene expression. It varies greatly in plants from complete absence in the bryophyte *Marchantia polymorpha* to the occurrence of several hundred or even thousands of editing events (for a review, see *Ichinose & Sugita (2016)*). RNA editing in nonphotosynthetic plants is of special interest because many proteins are involved in RNA editing in both plastids and mitochondria (for example, see *Bentolila et al. (2012)*). Thus, the reduction of the plastid genome and

the coordinated loss of nuclear genes involved in editing in plastids can also influence mitochondrial RNA editing. Currently, data on RNA editing in mitochondria for nonphotosynthetic plants are scarce. C-to-U RNA editing was found in seven genes (*atp1*, *atp4*, *atp6*, *cox2*, *nad1*, *rps4* and *rps12*) in *R. cantleyi* (*Xi et al., 2013*); and Barkman and coauthors demonstrated editing in the putatively horizontally transferred *atp1* gene in a wide range of plant species (*Barkman et al., 2007*). In *V. scurruloideum*, C-to-U editing was predicted computationally for the nine protein-coding genes (*Skippington et al., 2015*). In *H. monotropa*, we observed editing in the majority of genes; almost all editing events were of the C-to-U type, which is typical of plant mitochondria. The vast majority of the editing events were found in all samples (Fig. 3; Table S4). We observed a single non-C-U editing site, as A–G editing was observed in the *rps19* gene transcript. This editing was represented in all samples (inflorescence, ovules and anthers) at the level of 17–23% of reads (Table S4). Considering that there was no coverage peak at this position, it is unlikely that this was a result of read mismapping. We hypothesize that the A–G conversion that we observed actually represent not A-to-G editing but adenosine-to-inosine editing (inosine reads as G in sequencing data). This type of editing is typical of animals and fungi (*Cattenoz et al., 2013*; *Liu et al., 2016*) but has not previously been found in plants. In many cases, RNA editing plays a clear functional role (e.g., restoration of typical start codons in the plastid genes *rpl2* and *psbL* (*Kudla et al., 1992*)). We found one such case in *H. monotropa* (*nad4L*). A more detailed examination of RNA editing in MHT plants, including its dynamics in different organs and developmental stages, is required to highlight potentially functional events.

## CONCLUSIONS

Nonphotosynthetic plants represent ~1% of plant diversity and provide excellent models for the study of convergent evolution. Until recently, the genomic research in non-photosynthetic plants was focused on plant parasites; a common assumption is that mycoheterotrophs—plants that parasitize fungi—exhibit basically the same patterns of genome evolution as plant parasites. To test this hypothesis and to expand our knowledge of MHT plants, we characterized the mitochondrial genome of *H. monotropa*. Additionally, using RNA-seq, we performed a genome-wide analysis of gene expression and RNA editing. We showed that the mitogenome structure of *H. monotropa* is highly unusual: it includes a small circular fragment and a large linear fragment with multiple repeats at its ends that presumably function as telomeres. Further studies that include the characterization of mitogenomes of other Ericaceae and the in vivo analysis of *H. monotropa* mitochondria are required to investigate the details of the evolution, replication and functioning of such unusual mitogenomes. The gene set is similar to that of autotrophic plants. All protein-coding genes are expressed, and in most of these genes (33 out of 38), we observed editing of the transcripts. The intergenic regions of the mitogenome carry multiple sequences of plastid origin, including sequences of photosynthesis-related genes that are absent in the *H. monotropa* plastome. We did not find any traces of HGT from fungal hosts in the *H. monotropa* mitogenome or an increase in

nucleotide substitution rates. These new data highlight the contrast between MHT and parasitic plants and emphasize the need for new model species representing MHT plants.

## ADDITIONAL INFORMATION AND DECLARATIONS

### Competing Interests

The authors declare that they have no competing interests.

### Author Contributions

- Viktoria Yu Shtratnikova analyzed the data, prepared figures and/or tables, authored or reviewed drafts of the paper, and approved the final draft.
- Mikhail I Schelkunov analyzed the data, prepared figures and/or tables, authored or reviewed drafts of the paper, and approved the final draft.
- Aleksey A. Penin conceived and designed the experiments, performed the experiments, prepared figures and/or tables, and approved the final draft.
- Maria D. Logacheva conceived and designed the experiments, performed the experiments, analyzed the data, prepared figures and/or tables, authored or reviewed drafts of the paper, and approved the final draft.

### Field Study Permissions

The following information was supplied relating to field study approvals (i.e., approving body and any reference numbers):

*Hypopitys monotropa* is not an endangered or threatened plant species and thus according to Russian laws and local rules of our university no specific permissions were required for collection. The land is public so no permission is applicable.

### DNA Deposition

The following information was supplied regarding the deposition of DNA sequences:

The sequences are available at GenBank: MK990822 and MK990823 (BioProject: PRJNA522958 and PRJNA573526).

## Data Availability

The following information was supplied regarding data availability:

Raw data is available at NCBI, Bioproject PRJNA522958.

The code of the scripts written by the authors of this article and used for the analyses is deposited on Figshare. The link is DOI 10.6084/m9.figshare.8247023.

## Supplemental Information

Supplemental information for this article can be found online at http://dx.doi.org/10.7717/peerj.9309#supplemental-information.

## REFERENCES

**Adams KL, Qiu Y-L, Stoutemyer M, Palmer JD. 2002.** Punctuated evolution of mitochondrial gene content: high and variable rates of mitochondrial gene loss and transfer to the nucleus during angiosperm evolution. *Proceedings of the National Academy of Sciences of the United States of America* **99(15)**:9905–9912 DOI 10.1073/pnas.042694899.

**Alverson AJ, Rice DW, Dickinson S, Barry K, Palmer JD. 2011.** Origins and recombination of the bacterial-sized multichromosomal mitochondrial genome of cucumber. *Plant Cell* **23(7)**:2499–2513 DOI 10.1105/tpc.111.087189.

**Alverson AJ, Wei X, Rice DW, Stern DB, Barry K, Palmer JD. 2010.** Insights into the evolution of mitochondrial genome size from complete sequences of Citrullus lanatus and *Cucurbita pepo* (Cucurbitaceae). *Molecular Biology and Evolution* **27(6)**:1436–1448 DOI 10.1093/molbev/msq029.

**Barkman TJ, McNeal JR, Lim S-H, Coat G, Croom HB, Young ND, DePamphilis CW. 2007.** Mitochondrial DNA suggests at least 11 origins of parasitism in angiosperms and reveals genomic chimerism in parasitic plants. *BMC Evolutionary Biology* **7(1)**:248 DOI 10.1186/1471-2148-7-248.

**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.

**Bellot S, Cusimano N, Luo S, Sun G, Zarre S, Gröger A, Temsch E, Renner SS. 2016.** Assembled plastid and mitochondrial genomes, as well as nuclear genes, place the parasite family Cynomoriaceae in the Saxifragales. *Genome Biology and Evolution* **8(7)**:2214–2230 DOI 10.1093/gbe/evw147.

**Bentolila S, Heller WP, Sun T, Babina AM, Friso G, Van Wijk KJ, Hanson MR. 2012.** RIP1, a member of an *Arabidopsis* protein family, interacts with the protein RARE1 and broadly affects RNA editing. *Proceedings of the National Academy of Sciences of the United States of America* **109(22)**:E1453–E1461 DOI 10.1073/pnas.1121465109.

**Bergthorsson U, Richardson AO, Young GJ, Goertzen LR, Palmer JD. 2004.** Massive horizontal transfer of mitochondrial genes from diverse land plant donors to the basal angiosperm Amborella. *Proceedings of the National Academy of Sciences of the United States of America* **101(51)**:17747–17752 DOI 10.1073/pnas.0408336102.

**Bjorkman E. 1960.** *Monotropa hypopitys L.*—an epiparasite on tree roots. *Physiologia Plantarum* **13(2)**:308–327 DOI 10.1111/j.1399-3054.1960.tb08034.x.

**Boetzer M, Pirovano W. 2012.** Toward almost closed genomes with GapFiller. *Genome Biology* **13(6)**:R56 DOI 10.1186/gb-2012-13-6-r56.

**Bromham L, Cowman PF, Lanfear R. 2013.** Parasitic plants have increased rates of molecular evolution across all three genomes. *BMC Evolutionary Biology* **13(1)**:126 DOI 10.1186/1471-2148-13-126.

**Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, Madden TL. 2009.** BLAST+: architecture and applications. *BMC Bioinformatics* **10(1)**:421 DOI 10.1186/1471-2105-10-421.

**Cattenoz PB, Taft RJ, Westhof E, Mattick JS. 2013.** Transcriptome-wide identification of A & I RNA editing sites by inosine specific cleavage. *RNA* **19(2)**:257–270 DOI 10.1261/rna.036202.112.

**Chen L. 2013.** Characterization and comparison of human nuclear and cytosolic editomes. *Proceedings of the National Academy of Sciences* **110(29)**:E2741–E2747 DOI 10.1073/pnas.1218884110.

**Cho Y, Qiu YL, Kuhlman P, Palmer JD. 1998.** Explosive invasion of plant mitochondria by a group I intron. *Proceedings of the National Academy of Sciences* **95(24)**:14244–14249 DOI 10.1073/pnas.95.24.14244.

**Cusimano N, Renner SS. 2019.** Sequential horizontal gene transfers from different hosts in a widespread Eurasian parasitic plant, *Cynomorium coccineum*. *American Journal of Botany* **106(5)**:679–689 DOI 10.1002/ajb2.1286.

**Cusimano N, Zhang L-B, Renner SS. 2008.** Reevaluation of the *cox1* group I intron in Araceae and angiosperms indicates a history dominated by loss rather than horizontal transfer. *Molecular Biology and Evolution* **25(2)**:265–276 DOI 10.1093/molbev/msm241.

**Edera AA, Gandini CL, Sanchez-Puerta MV. 2018.** Towards a comprehensive picture of C-to-U RNA editing sites in angiosperm mitochondria. *Plant Molecular Biology* **97(3)**:215–231 DOI 10.1007/s11103-018-0734-9.

**Fajardo D, Schlautman B, Steffan S, Polashock J, Vorsa N, Zalapa J. 2014.** The American cranberry mitochondrial genome reveals the presence of selenocysteine (tRNA-Sec and SECIS) insertion machinery in land plants. *Gene* **536(2)**:336–343 DOI 10.1016/j.gene.2013.11.104.

**Fan W, Zhu A, Kozaczek M, Shah N, Pabón-Mora N, González F, Mower JP. 2016.** Limited mitogenomic degradation in response to a parasitic lifestyle in Orobanchaceae. *Scientific Reports* **6(1)**:36285 DOI 10.1038/srep36285.

**Figueroa P, Gómez I, Carmona R, Holuigue L, Araya A, Jordana X. 1999.** The gene for mitochondrial ribosomal protein S14 has been transferred to the nucleus in Arabidopsis thaliana. *Molecular and General Genetics MGG* **262(1)**:139–144 DOI 10.1007/s004380051068.

**Gandini CL, Sanchez-Puerta MV. 2017.** Foreign plastid sequences in plant mitochondria are frequently acquired via mitochondrion-to-mitochondrion horizontal transfer. *Scientific Reports* **7(1)**:123 DOI 10.1038/srep43402.

**Giege P, Brennicke A. 1999.** RNA editing in Arabidopsis mitochondria effects 441 C to U changes in ORFs. *Proceedings of the National Academy of Sciences of the United States of America* **96(26)**:15324–15329 DOI 10.1073/pnas.96.26.15324.

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

**Grewe F, Edger PP, Keren I, Sultan L, Pires JC, Ostersetzer-Biran O, Mower JP. 2014.** Comparative analysis of 11 Brassicales mitochondrial genomes and the mitochondrial transcriptome of *Brassica oleracea*. *Mitochondrion* **19**:135–143 DOI 10.1016/j.mito.2014.05.008.

**Gualberto JM, Mileshina D, Wallet C, Niazi AK, Weber-Lotfi F, Dietrich A. 2014.** The plant mitochondrial genome: dynamics and maintenance. *Biochimie* **100**:107–120 DOI 10.1016/j.biochi.2013.09.016.

**Handa H. 2003.** The complete nucleotide sequence and RNA editing content of the mitochondrial genome of rapeseed (*Brassica napus* L.): comparative analysis of the mitochondrial genomes of rapeseed and *Arabidopsis thaliana*. *Nucleic Acids Research* **31(20)**:5907–5916 DOI 10.1093/nar/gkg795.

**Handa H. 2008.** Linear plasmids in plant mitochondria: peaceful coexistences or malicious invasions? *Mitochondrion* **8(1)**:15–25 DOI 10.1016/j.mito.2007.10.002.

**Handa H, Itani K, Sato H. 2002.** Structural features and expression analysis of a linear mitochondrial plasmid in rapeseed (*Brassica napus* L.). *Molecular Genetics and Genomics* **267(6)**:797–805 DOI 10.1007/s00438-002-0711-4.

**Ichinose M, Sugita M. 2016.** RNA editing and its molecular mechanism in plant organelles. *Genes* **8(1)**:5 DOI 10.3390/genes8010005.

**Janouškovec J, Liu S-L, Martone PT, Carré W, Leblanc C, Collén J, Keeling PJ. 2013.** Evolution of red algal plastid genomes: ancient architectures, introns, horizontal gene transfer, and taxonomic utility of plastid markers. *PLOS ONE* **8(3)**:e59001 DOI 10.1371/journal.pone.0059001.

**Joyce PB, Gray MW. 1989.** Chloroplast-like transfer RNA genes expressed in wheat mitochondria. *Nucleic Acids Research* **17(14)**:5461–5476 DOI 10.1093/nar/17.14.5461.

**Katoh K, Rozewicki J, Yamada KD. 2019.** MAFFT online service: multiple sequence alignment, interactive sequence choice and visualization. *Briefings in Bioinformatics* **20(4)**:1160–1166 DOI 10.1093/bib/bbx108.

**Kayal E, Bentlage B, Collins AG, Kayal M, Pirro S, Lavrov DV. 2012.** Evolution of linear mitochondrial genomes in medusozoan Cnidarians. *Genome Biology and Evolution* **4(1)**:1–12 DOI 10.1093/gbe/evr123.

**Kim G, LeBlanc ML, Wafula EK, DePamphilis CW, Westwood JH. 2014.** Genomic-scale exchange of mRNA between a parasitic plant and its hosts. *Science* **345(6198)**:808–811 DOI 10.1126/science.1253122.

**Kozik A, Rowan B, Lavelle D, Berke L, Schranz ME, Michelmore RW, Christensen AC. 2019.** The alternative reality of plant mitochondrial DNA. *bioRxiv* **12(5)**:e0177606 DOI 10.1101/564278.

**Krzywinski M, Schein J, Birol I, Connors J, Gascoyne R, Horsman D, Jones SJ, Marra MA. 2009.** Circos: an information aesthetic for comparative genomics. *Genome Research* **19(9)**:1639–1645 DOI 10.1101/gr.092759.109.

**Kudla J, Igloi GL, Metzlaff M, Hagemann R, Kössel H. 1992.** RNA editing in tobacco chloroplasts leads to the formation of a translatable psbL mRNA by a C to U substitution within the initiation codon. *EMBO Journal* **11(3)**:1099–1103 DOI 10.1002/j.1460-2075.1992.tb05149.x.

**Liu H, Wang Q, He Y, Chen L, Hao C, Jiang C, Li Y, Dai Y, Kang Z, Xu J-R. 2016.** Genome-wide A-to-I RNA editing in fungi independent of ADAR enzymes. *Genome Research* **26(4)**:499–509 DOI 10.1101/gr.199877.115.

**Lloyd Evans D, Hlongwane TT, Joshi SV, Riaño Pachón DM. 2019.** The sugarcane mitochondrial genome: assembly, phylogenetics and transcriptomics. *PeerJ* **7(8)**:e7558 DOI 10.7717/peerj.7558.

**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, Shtratnikova VY, Matveeva MV, Penin AA. 2016.** Comparative analysis of plastid genomes of non-photosynthetic Ericaceae and their photosynthetic relatives. *Scientific Reports* **6(1)**:30042 DOI 10.1038/srep30042.

**Loomis EW, Eid JS, Peluso P, Yin J, Hickey L, Rank D, McCalmon S, Hagerman RJ, Tassone F, Hagerman PJ. 2013.** Sequencing the unsequenceable: expanded CGG-repeat alleles of the fragile X gene. *Genome Research* **23(1)**:121–128 DOI 10.1101/gr.141705.112.

**Mariotti M, Lobanov AV, Guigo R, Gladyshev VN. 2013.** SECISearch3 and Seblastian: new tools for prediction of SECIS elements and selenoproteins. *Nucleic Acids Research* **41(15)**:e149 DOI 10.1093/nar/gkt550.

**Min S, Chang-Qin Z, Yong-Peng M, Welti S, Moreau P-A, Selosse M-A. 2012.** Mycorrhizal features and fungal partners of four mycoheterotrophic Monotropoideae (Ericaceae) species from Yunnan, China. *Symbiosis* **57(1)**:1–13 DOI 10.1007/s13199-012-0180-4.

**Nosek J, Dinouël N, Kovac L, Fukuhara H. 1995.** Linear mitochondrial DNAs from yeasts: telomeres with large tandem repetitions. *Molecular and General Genetics MGG* **247(1)**:61–72 DOI 10.1007/BF00425822.

**Nosek J, Tomáška L, Fukuhara H, Suyama Y, Kováč L. 1998.** Linear mitochondrial genomes: 30 years down the line. *Trends in Genetics* **14(5)**:184–188 DOI 10.1016/S0168-9525(98)01443-7.

**Notsu Y, Masood S, Nishikawa T, Kubo N, Akiduki G, Nakazono M, Hirai A, Kadowaki K. 2002.** The complete sequence of the rice (*Oryza sativa* L.) mitochondrial genome: frequent DNA sequence acquisition and loss during the evolution of flowering plants. *Molecular Genetics and Genomics* **268(4)**:434–445 DOI 10.1007/s00438-002-0767-1.

**Olson AR. 1990.** Observation on the floral shoots of *Monotropa hypopitys* (Monotropaceae). *Rhodora* **92**:54–56.

**Olson AR. 1993.** Patterns of embryo and endosperm formation in *Monotropa hypopitys* (Monotropaceae) from North America and Sweden. *American Journal of Botany* **80(7)**:839–846 DOI 10.1002/j.1537-2197.1993.tb15299.x.

**Petersen G, Cuenca A, Møller IM, Seberg O. 2015.** Massive gene loss in mistletoe (Viscum, Viscaceae) mitochondria. *Scientific Reports* **5(1)**:525 DOI 10.1038/srep17588.

**Petersen G, Darby H, Lam VKY, Pedersen H, Merckx VSFT, Zervas A, Seberg O, Graham SW. 2019.** Mycoheterotrophic Epirixanthes (Polygalaceae) has a typical angiosperm mitogenome but unorthodox plastid genomes. *Annals of Botany* **124(5)**:791–807 DOI 10.1093/aob/mcz114.

**Qiu Y, Filipenko SJ, Darracq A, Adams KL. 2014.** Expression of a transferred nuclear gene in a mitochondrial genome. *Current Plant Biology* **1**:68–72 DOI 10.1016/j.cpb.2014.08.002.

**Rice DW, Alverson AJ, Richardson AO, Young GJ, Sanchez-Puerta MV, Munzinger J, Barry K, Boore JL, Zhang Y, DePamphilis CW, Knox EB, Palmer JD. 2013.** Horizontal transfer of entire genomes via mitochondrial fusion in the angiosperm *Amborella*. *Science* **342(6165)**:1468–1473 DOI 10.1126/science.1246275.

**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.

**Sanchez-Puerta MV, Abbona CC, Zhuo S, Tepe EJ, Bohs L, Olmstead RG, Palmer JD. 2011.** Multiple recent horizontal transfers of the *cox1* intron in Solanaceae and extended co-conversion of flanking exons. *BMC Evolutionary Biology* **11(1)**:197 DOI 10.1186/1471-2148-11-277.

**Sanchez-Puerta MV, Cho Y, Mower JP, Alverson AJ, Palmer JD. 2008.** Frequent, phylogenetically local horizontal transfer of the cox1 group I intron in flowering plant mitochondria. *Molecular Biology and Evolution* **25(8)**:1762–1777 DOI 10.1093/molbev/msn129.

**Sanchez-Puerta MV, Edera A, Gandini CL, Williams AV, Howell KA, Nevill PG, Small I. 2019.** Genome-scale transfer of mitochondrial DNA from legume hosts to the holoparasite Lophophytum mirabile (Balanophoraceae). *Molecular Phylogenetics and Evolution* **132**:243–250 DOI 10.1016/j.ympev.2018.12.006.

**Sanchez-Puerta MV, García LE, Wohlfeiler J, Ceriotti LF. 2017.** Unparalleled replacement of native mitochondrial genes by foreign homologs in a holoparasitic plant. *New Phytologist* **214(1)**:376–387 DOI 10.1111/nph.14361.

**Santesmasses D, Mariotti M, Guigó R. 2017.** Computational identification of the selenocysteine tRNA (tRNASec) in genomes. *PLOS Computational Biology* **13(2)**:e1005383 DOI 10.1371/journal.pcbi.1005383.

**Schelkunov MI, Penin AA, Logacheva MD. 2018.** RNA-seq highlights parallel and contrasting patterns in the evolution of the nuclear genome of fully mycoheterotrophic plants. *BMC Genomics* **19(1)**:1255 DOI 10.1186/s12864-018-4968-3.

**Skippington E, Barkman TJ, Rice DW, Palmer JD. 2015.** Miniaturized mitogenome of the parasitic plant *Viscum scurruloideum* is extremely divergent and dynamic and has lost all *nad* genes. *Proceedings of the National Academy of Sciences of the United States of America* **112(27)**:E3515–E3524 DOI 10.1073/pnas.1504491112.

**Skippington E, Barkman TJ, Rice DW, Palmer JD. 2017.** Comparative mitogenomics indicates respiratory competence in parasitic Viscum despite loss of complex I and extreme sequence divergence, and reveals horizontal gene transfer and remarkable variation in genome size. *BMC Plant Biology* **17(1)**:1115 DOI 10.1186/s12870-017-0992-8.

**Sloan DB. 2013.** One ring to rule them all? Genome sequencing provides new insights into the 'master circle' model of plant mitochondrial DNA structure. *New Phytologist* **200(4)**:978–985 DOI 10.1111/nph.12395.

**Sloan DB, Alverson AJ, Wu M, Palmer JD, Taylor DR. 2012.** Recent acceleration of plastid sequence and structural evolution coincides with extreme mitochondrial divergence in the angiosperm genus Silene. *Genome Biology and Evolution* **4(3)**:294–306 DOI 10.1093/gbe/evs006.

**Stamatakis A. 2014.** RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* **30(9)**:1312–1313 DOI 10.1093/bioinformatics/btu033.

**Xi Z, Wang Y, Bradley RK, Sugumaran M, Marx CJ, Rest JS. 2013.** Massive mitochondrial gene transfer in a parasitic flowering plant clade. *PLOS Genetics* **9(2)**:e1003265 DOI 10.1371/journal.pgen.1003265.

**Xu G, Zhang J. 2014.** Human coding RNA editing is generally nonadaptive. *Proceedings of the National Academy of Sciences of the United States of America* **111(10)**:3769–3774 DOI 10.1073/pnas.1321745111.

**Yang Z, Zhang Y, Wafula EK, Honaas LA, Ralph PE, Jones S, Clarke CR, Liu S, Su C, Zhang H, Altman NS, Schuster SC, Timko MP, Yoder JI, Westwood JH, DePamphilis CW. 2016.** Horizontal gene transfer is more frequent with increased heterotrophy and contributes to parasite adaptation. *Proceedings of the National Academy of Sciences of the United States of America* **113(45)**:E7010–E7019 DOI 10.1073/pnas.1608765113.

**You R, Yao S, Xiong Y, Huang X, Sun F, Mamitsuka H, Zhu S. 2019.** NetGO: improving large-scale protein function prediction with massive network information. *Nucleic Acids Research* **47(W1)**:W379–W387 DOI 10.1093/nar/gkz388.

**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)**:1129 DOI 10.1038/s41467-018-03423-5.

**Zervas A, Petersen G, Seberg O. 2019.** Mitochondrial genome evolution in parasitic plants. *BMC Evolutionary Biology* **19(1)**:227 DOI 10.1186/s12862-019-1401-8.

**Zhu Y, Luo H, Zhang X, Song J, Sun C, Ji A, Xu J, Chen S. 2014.** Abundant and selective RNA-editing events in the medicinal mushroom Ganoderma lucidum. *Genetics* **196(4)**:1047–1057 DOI 10.1534/genetics.114.161414.