bioRxiv preprint doi: https://doi.org/10.1101/681718; this version posted June 25, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Mitochondrial genome of non-photosynthetic mycoheterotrophic plant *Hypopitys monotropa*: unusual structure and extensive RNA editing

Victoria Y. Shtratnikova[1], Mikhail I. Schelkunov[2,3], Maria D. Logacheva[1,2,3]

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

Corresponding Author:
Maria D. Logacheva[1,2,3]
[1] Lomonosov Moscow State University, Leninskie Gory, GSP-1, Moscow 119991, Russia
[2] Bolshoy Karetny lane, 19/1, Moscow 127051, Russia
[3] Skolkovo Institute of Science and Technology, Nobel St. 3, Moscow 143026, Russia

Email address: maria.log@gmail.com

Abstract

The plants that lost the ability to photosynthesis (heterotrophic) are characterized by a number of changes at all levels of organization - morphological, physiological and genomic. Heterotrophic plants divide into two large categories – parasitic and mycoheterotrophic. The question of to what extent these changes are similar in these two categories is still open. Plastid genomes of non-photosynthetic plants are well characterized and they demonstrate similar patterns of reduction in both groups. In contrast, little is known about mitochondrial genomes of mycoheterotrophic plants. We report the structure of the mitochondrial genome of *Hypopitys monotropa*, a mycoheterotrophic member of Ericaceae, and the expression of mitochondrial 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*, its complete size is ~810 Kbp. We found 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 on parasitic non-photosynthetic plants. Transcriptome sequencing revealed trans-splicing of several genes and RNA editing in 33 genes of 38. 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

bioRxiv preprint doi: https://doi.org/10.1101/681718; this version posted June 25, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Hypopitys monotropa (Ericacae, Monotropoideae) is a non-photosynthetic plant gaining carbon from fungi that are ectomycorrhizal with tree roots (Bjorkman, 1960). In contrast to most other mycoheterotrophic plants, which are very rare and/or very narrowly distributed, Monotropoideae, including *H. monotropa*, are quite widespread, being associated with old-growth conifer forests. Thus, *H. monotropa* is used as a model system in studies of plant-mycorrhizal associations and developmental biology of mycoheterotrophic plants (e.g. (Olson, 1993, 1990)). Recent advances in DNA sequencing allow expanding the studies of mycoheterotrophs into genomics. By now, the most attention was focused on plastid genomes; in particular, complete plastid genomes of several individuals of *H. monotropa* were reported (Gruzdev et al., 2016; Logacheva et al., 2016). It shows high reduction and complete absence of genes of photosynthesis electron transport chain. Thus, the mycoheterotrophic lifestyle strongly influenced on plastids, but what about the mitochondrial genome?

In contrast to animals, where mitochondrial genomes are usually conserved in size and gene content across large taxonomic groups, in plants they are highly variable and may be very dissimilar even in closely related species. The size of the angiosperm mitogenome is ranging from 66 Kb in the hemiparasitic mistletoe *Viscum scurruloideum* (Skippington et al., 2015), 222 Kb in autotrophic *Brassica napus* (Handa, 2003) to more than 11 Mb in *Silene noctiflora* (Sloan et al., 2012). Despite their large sizes, the mitogenome gene content is rather stable. The number of genes is usually within the range between 50 and 60, encoding subunits of the oxidative phosphorylation chain complexes and proteins involved in the biogenesis of these complexes, as well as several ribosomal proteins (Gualberto et al., 2014), 15–21 tRNA genes and 3 rRNA (18S, 26S, 5S).

Non-photosynthetic plants divide into two large groups—those that are parasitic on other plants and mycoheterotropic. By now only few complete mitochondrial genomes of non-photosynthetic plants are characterized, and all but one of them are parasitic. A comparative analysis of mitogenomes of several parasitic, hemiparasitic and autotrophic Orobanchaceae (Fan et al., 2016) showed that the gene content does not depend on trophic specialization in family range. Mitogenomes of two non-related lineages of parasitic plants: Rafflesiaceae and Cynomoriaceae also do not show reduction in gene content; besides that, they are the example of massive HGT from other plants (including, but not only, their hosts). This is not, however, a trait unique for parasitic plants—e.g. *Amborella trichopoda*, an autotrophic plant from basal angiosperms, has in its mitochondrial genome a large fraction acquired from green algae, mosses, and other angiosperms (Rice et al., 2013). In contrast, in the hemiparasitic plant *Viscum scurruloideum* the mitogenome is drastically reduced in size and gene content, it lacks all nine *nad* genes, cytochrome biogenesis genes and *matR* (Skippington et al., 2015). Mitogenomes of other Viscum species are not reduced in length but have reduced gene content, similar to *V. scurruloideum* ((Petersen et al., 2015); but see (Skippington et al., 2017)). The sampling is obviously insufficient even for parasitic plants; regarding mycoheterotrophs the data are almost completely lacking. The only mitochondrial genome of a mycoheterotrophic plant characterized by date is that of orchid *Gastrodia elata* (Yuan et al., 2018). Taking this into account, we set the

bioRxiv preprint doi: https://doi.org/10.1101/681718; this version posted June 25, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

following objectives: 1) to characterize the structure and gene content of the mitochondrial genome of *H. monotropa* 2) to estimate if horizontal gene transfer (HGT) from fungi took place 3) to study the mitochondrial gene expression and RNA editing.

The photosynthethic plant with characterized mitochondrial genome phylogenetically closest to *H. monotropa* is cranberry, *Vaccinium macrocarpon* (Fajardo et al., 2014). Its mitogenome is ~460 Kb in length, contains 34 protein-coding genes, 3 ribosomal RNAs and 17 transfer RNAs. Compared to other autotrophic plants, *sdh3*, *rps2*, *rps3* and *rps11* are lost; *atp6* and *nad4L* genes are reported to be pseudogenised. An unusual trait of *V. macrocarpon* mitogenome is the presence of tRNA-Sec and a selenocysteine insertion sequence (SECIS) element, a stem-loop structure located in the 3′UTR (but see discussion below).

Materials & Methods

2.1. Sample collection and sequencing

Sample collection, DNA and RNA libraries preparation and sequencing were described in (Logacheva et al., 2016) and (Schelkunov et al., 2018). The reads are deposited in NCBI Sequence Read Archive under accession number PRJNA522958.

2.2. Assembly

Read trimming and assembly were made as described in (Logacheva et al., 2016). The contig coverages were 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 *Vaccinium macrocarpon* genome (GenBank accession NC_023338) against all contigs in the assembly. BLASTN and TBLASTX were those of BLAST 2.3.0+ suite (Camacho et al., 2009), the alignment was performed with the maximum allowed e-value of 10^{-5}. Low-complexity sequence filters were switched off in order not to miss genes with extremely high or low GC-content. A contig that corresponded to the plastid genome of *Hypopitys monotropa* also aligned to the mitochondrial genome of *Vaccinium macrocarpon* because of the presence of inserts from the plastid genome into 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 that had coverage close to the coverage of nuclear contigs or ones that produce matches only to non-mitochondrial sequences when aligned to NCBI NT database by BLASTN (e-value $10^{-3)}$, were considered nuclear.

After the procedures described above, 7 contigs remained, with lengths 255, 5651, 20627, 102868, 106298, 234240, 309413 bp and coverages 37, 89, 93, 110, 96, 107, 108 respectively. The lower coverage of the smallest contig is presumably an artifact caused by its length. To understand their order in a mitochondrial genome, we mapped mate-pair reads from the library with the bigger insert size (8705±2537 bp) to the contigs and investigated mate-pair links. The mapping was performed by CLC Genomics Workbench, 100% of a read' 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 Circos 0.67-5 (Krzywinski et al., 2009) and

bioRxiv preprint doi: https://doi.org/10.1101/681718; this version posted June 25, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

investigated manually. In one contig mate-pair links spanned from its end to the start and thus the contig corresponds to a circular chromosome. Other six contigs connect into a linear sequence, with a structure described in the Results section. Gaps between the contigs were closed by GapFiller (Boetzer and Pirovano, 2012) that was run with 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. To check whether the two chromosomes have been assembled correctly, we mapped reads from all three sequencing libraries by 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 is approximately uniform along the whole length of the chromosomes. We found that the coverage gradually decreases on the edges of the linear chromosome, thus indicating that different copies of the chromosome in a plant have different lengths of the terminal regions. For definiteness, we elongated the terminal regions such that the coverage drops to 0, thus the length of the linear chromosome approximately corresponds to the maximal possible length among all copies of the chromosome in the sequenced plant. Also, we mapped reads from the mate-pair library with the bigger insert size, requiring 100% of a read length to align with a sequence similarity of 100%, and investigated a distribution of mean insert sizes over each genomic position and a number of mate pairs than span over each position. These values were approximately uniform along both chromosomes (Supplementary Figure S1), thus suggesting that the assembly was correct.

2.3. 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 by CLC Genomics Workbench v. 7.5.1 to the sequences of the chromosomes, requiring at least 80% of a read's length to map with at least 90% sequence similarity. The relaxed setting for a percent of read length to be mapped allows mapping reads spanning the splice junctions and the relaxed setting for sequence similarity allows 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 are any unnoticed protein-coding genes, we visually inspected the alignment to search for unannotated regions in the chromosomes with high coverage by RNA-seq reads.

The search for genes encoding selenocysteine tRNAs (tRNA-Sec) was done for *Hypopitys monotropa* and *Vaccinium macrocarpon* by an online-version of Secmarker 0.4 (Santesmasses et al., 2017) with the default parameters. The SECIS element prediction was performed by an online-version of SECISEARCH3 (Mariotti et al., 2013), current as of 5 May 2017, with the default parameters.

2.4. RNA editing analysis

To obtain information on RNA editing sites, we artificially spliced gene sequences and mapped RNA-seq reads to them with the parameters stated above for RNA-seq reads mapping. The variant calling was performed by CLC Genomics Workbench 7.5.1. We considered a site in a CDS as edited if it was covered by at least 20 reads and at least 10% of the reads differ in this position from the genomic sequence.

bioRxiv preprint doi: https://doi.org/10.1101/681718; this version posted June 25, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

To estimate the fraction of RNA editing sites that change amino acid sequence of corresponding proteins, we use a measure that was already used in several previous studies (Chen, 2013; Xu and Zhang, 2014; Zhu et al., 2014), but hasn't yet got its own name. Here we call it dEN/dES. It is calculated similarly to dN/dS, but while dN/dS counts nonsynonymous and synonymous mutations, dEN/dES ("E" here stands for "editing") counts nonsynonymous and synonymous RNA editing events. If editing sites are distributed in a CDS irrespectively to whether they do or do not change amino acids, dEN/dES will be close to one.

2.5. Search for sequences transferred to the mitogenome

In order to find the sequences transferred from the plastome to the mitogenome (known as "MIPTs", which stands for MItochondrial Plastid Transfers), we aligned 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 complete set. Its proteins were aligned to the sequences of the mitochondrial chromosomes by TBLASTN with the maximum allowed e-value of 10^{-3}. tRNA and rRNA coding genes were similarly aligned to the sequences of the chromosomes by BLASTN. The matching regions in the chromosomes were then aligned by BLASTX to NCBI NR (for regions that matched to *Camellia* proteins) and by BLASTN to NCBI NT (for regions that matched to *Camellia* RNA coding genes). If, for a region in the mitogenome, the best matches in the database were to sequences belonging to plastomes, that region was considered a MIPT. To calculate the number of frameshifting indels and nonsense mutations in the transferred regions they were taken together with their 200 bp-long flanking sequences on both ends and aligned to homologous genes from *C. sinensis*. The alignment was done by 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 500 bps each, with a step size of 50 bps. These windows were aligned by BLASTX to NCBI NR and proteins of *Russula brevipes*. and also by BLASTN to NCBI NT and scaffolds of *Russula brevipes. Russula brevipes* BPL707 was obtained from the MycoCosm portal (Grigoriev et al., 2014), the assembly version was 1.0. The maximum e-value allowed for the matches was 10^{-5}.

2.6 Phylogenetic analysis.

Genes common for 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 by MAFFT (Katoh et al., 2017). The phylogenetic analysis was performed using RaXML (raxmlGUI v.1.3.1) with nucleotide sequences under GTR+gamma substitution model. In order to infer possible horizontal transfers all protein-coding genes found in *H. monotropa* mitogenome were aligned and analyzed in the same way as the concatenated gene set.

Results

3.1. Genome assembly and annotation

bioRxiv preprint doi: https://doi.org/10.1101/681718; this version posted June 25, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Seven contigs of the assembly were identified as mitochondrial based on read coverage and presence of typical mitochondrial genes. The mitochondrial contigs have a 70-80× coverage. After scaffolding and gap filling these contigs were assembled in two sequences. GC content of both sequences is ~ 45%. One of them is circular with a length of 106 Kb (Figure 1). It does not have any long repeats; the mapping of mate-pair reads shows the absence of pairs with abnormal insert sizes suggesting that this sequence is not a subject of recombination. The second fragment, the longer one (Figure 1), is ~704 Kbp long. It was assembled as linear and has a more complex structure. In particular, it contains several long repeats. As shown in Figure 2, three pairs of long direct repeats are 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 are also found: between the 260-262 Kb region and the 657-659 Kb region. A peculiar feature is a gradual decrease of coverage on both of its ends (Supplementary Figure S1). This suggests that copies of the mitogenome with different lengths of this repeat coexist in plant cells. 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 (Supplementary Figure S1) has no sharp drops, therefore it is unlikely that these chromosomes join frequently.

The circular chromosome contains only two full-length protein coding genes — *ccmFc* and *cox1* and two exons of *nad5* gene, while two other exons are located on linear fragment. Summary data for the annotation is presented in Table 1.

The gene content of the *H. monotropa* mitogenome is typical of flowering plants, with most genes encoding components of electron transfer chain and translation apparatus. Unlike parasitic plants, which show a certain degree of reduction, from limited to high (Fan et al., 2016; Skippington et al., 2015), the gene content in *H. monotropa* is not reduced. Surprisingly it is even larger than in *V. macrocarpon*, a close photosynthetic relative: *nad4L* and *atp6* are pseudogenized and *sdh3* and *rps3* are absent in *V. macrocarpon*. There are, however, two pseudogenes in *H. monotropa*: *rps14* and *rps19*. Both of them can be either present or absent in autotrophic as well as in non-photosynthetic plants (Petersen et al., 2015; Skippington et al., 2015).

Though being highly conserved in coding sequence, mitochondrial genes sometimes differ in 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, acquired by horizontal transfer events ((Cho et al., 1998; Sanchez-Puerta et al., 2011) but see (Cusimano et al., 2008). The *cox1* intron is highly overrepresented in the parasitic plants that have been examined to date, though 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* it is present. In other genes *H. monotropa* has the same intron content as *V. macrocarpon*. Three genes—*nad1*, *nad2*, *nad5* have trans-spliced transcripts. Notably, *nad5* exons are located in different chromosomes – exons 1 and 2 in the linear one and exons 3 and 4

bioRxiv preprint doi: https://doi.org/10.1101/681718; this version posted June 25, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

in the circular. If these fragments are indeed independently replicated units, this is a rare case of interchromosome trans-splicing. In addition to the standard gene set we found a new ORF (ORF671) that hypothetically encodes a 671-aa protein, which does not have any significant similarity to other sequences. 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 typical of mitogenomes of autotrophic plant species. trnD-GUC in *H. monotropa* is split into two exons. 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) were reported in the mitogenome of *V. macrocarpon* (Fajardo et al., 2014). However, SecMarker, a novel specialized tool for the search of tRNA-Sec (Santesmasses et al., 2017) which has higher sensitivity and specificity than tools used by Fajardo and co-workers doesn't confirm the presence of a tRNA-Sec gene in the *V. macrocarpon* mitogenome. Thus, we suppose that the predictions of a tRNA-Sec gene and a SECIS element in the *V. macrocarpon* mitogenome are 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 presumable tRNA-Sec of *V. macrocarpon*. It is located upstream of *ccmC*, as well as one of the two presumable tRNA-Sec of *V. macrocarpon*. tRNA genes are known to have highly conserved sequences; in particular, the similarity of other *V. macrocarpon* and *H. monotropa* tRNA genes is 98-100%. This also suggests that the presumable tRNA-Sec of *V. macrocarpon* is not a functional gene but a pseudogene of some tRNA-coding gene.

Detailed data about genes, their intron-exon structure, expression and RNA-editing is presented in Supplementary Table 1.

3.2. Sequences transferred to the mitogenome

In the mitochondrial genome we found fragments with high similarity to plastid genes, both of those that are present in the *H. monotropa* plastome (*matK*, *rps2*, *rps4*) and those that were lost (*rpoB*, *C1*, *C2*, *psbC*, *ndhJ*, *B*, *ycf2*) (Supplementary Table 2). These fragments are derived from intracellular horizontal transfers from plastids (MIPTs). Such events are frequent in plant mitochondria (e.g. (Alverson et al., 2011; Grewe et al., 2014)). MIPTs in non-photosynthetic plants are of particular interest. If mitochondrial copies of plastid genes that were lost or pseudogenized in the plastome retain intact ORFs, potentially the reverse switch from heterotrophy to autotrophy is possible. Mitochondrial genome is in this case a “refugium” of plastid genes. The *Castilleja paramensis* mitogenome contains 55 full-length or nearly full-length plastid genes and only about half of them are obvious pseudogenes (Fan et al., 2016). By now the only example of functionality of mitochondrial genes transferred from plastids are tRNA genes of plastid origin recruited in the protein synthesis in mitochondria (Joyce and Gray, 1989). In *H. monotropa* all sequences that originated from plastid protein-coding genes do not represent intact ORFs, keeping in most cases less than 50% of the initial length and/or carrying multiple frameshift mutations (Supplementary table 2).

bioRxiv preprint doi: https://doi.org/10.1101/681718; this version posted June 25, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

3.3. Mitochondrial gene expression and RNA editing

In order to gain insights into mitochondrial gene expression and RNA editing and to refine the annotation we sequenced and assembled the transcriptome of *H. monotropa* (Logacheva et al., 2016; Schelkunov et al., 2018). We found expression of all annotated protein-coding genes. Minimal expression level (0.1) is observed for the hypothetical protein *ORF671*. Cytochrome *c* maturation factors are also expressed weakly (3 of 4 proteins have expression <1). Genes of Complex I have medium expression (1 — 6). The highest expression is observed for *sdh4*, *cox3*, *atp1* and *atp9*.

RNA editing is a phenomenon widely observed in mitochondrial transcripts. Levels of editing and sites of editing differ a lot from species to species and from gene to gene. In *H. monotropa* we identified 509 RNA-editing sites (Supplementary Table 1), only positions with editing level >10% and coverage >20 are included) in 33 genes from 38. Most of them are C-U. Maximal level of RNA-editing (measured as the number of sites per 100 bp of a gene's CDS) is observed for *ccmB*. Genes of Complex I demonstrate different levels of RNA-editing from 0.2 of *nad5* to 5.6 of *nad4L*. Level of RNA-editing in genes of Complex II, III is weak, is not over 1.1. At last, *atp1*, *rpl16*, *rps12*, *13*, *ORF671* have zero level of editing. Some genes with high expression have low level of RNA-editing (*atp1*, *rps12*), and *vice versa* (cytochrome *c* maturation factors) but generally there is no such tendency. Median dEN/dES across all genes is 1,23. There are no edited stop codons, but there are several stop codons that are introduced by editing (Supplementary Table 3).

3.4. Phylogenetic analysis.

As mentioned above, plant mitochondrial genomes are prone to HGT. It is often detected by incongruence of phylogenetic trees based on different genome regions (Bergthorsson et al., 2004), (Cusimano and Renner, 2019). We performed phylogenetic analysis of the individual mitochondrial genes and of the combined set and found that the topologies are identical. This evidences that no genes were acquired via HGT. The combined tree of all mitochondrial genes shared across 25 seed plant species shows topology similar to that of based on nuclear and plastid genes, with monocots representing monophyletic group and eudicots divided into two large groups – asterids and rosids. *H. monotropa* is sister to *V. macrocarpon*, and both are with the asterids, as expected. Notably, *H. monotropa* genes do not demonstrate any increase in the substitution rates. The same is true for the another mycoheterotrophic plant, *Petrosavia stellaris* (data from (Logacheva et al., 2014)). Earlier, parasitic plants were reported to have elevated substitution rates in all three genomes (Bromham et al., 2013); however, recent study encompassing broader sampling of parasitic plants shows that this is not a universal phenomenon (Zervas et al., 2019).

Discussion

4.1 Mitogenome structure

All available evidence suggests that *H. monotropa* mitogenome has a linear fragment. While linear plasmids are found in mitochondrial genomes of several plants (e.g. (H. et al., 2002)), the linear fragment of the *H. monotropa* mitogenome lacks characteristic features of these plasmids:

bioRxiv preprint doi: https://doi.org/10.1101/681718; this version posted June 25, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

a terminal inverted repeat, the small size (not over 11 Kbp) and genes of RNA and DNA polymerases (reviewed in (Handa, 2008)). A linear structure is usual for mitochondrial genomes of fungi and protists (Nosek et al., 1995), (Janouškovec et al., 2013). As well as other mycoheteroprophic plants, *H. monotropa* lives in intimate symbiosis with fungi (Min et al., 2012). One would hypothesize that that linear fragment could be the result of either a contamination or HGT from fungi. However this is unlikely, for the following reasons: 1) DNA was isolated from inflorescences while mycorrhiza exists 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 4) all genes annotated on the linear fragments are typical plant mitochondrial genes. The fact that single circular molecule ("master circle") is an oversimplified representation of the plant mitochondrial genomes and that they rather exist in vivo as a mix of circular, linear and branched forms is not novel (see, e.g. (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* where internal repeats are also present but their location, as well as the distribution of mate pair links does not allow to reconstitute the master circle (see Figure 2). This suggests that the diversity of organization of plant mitochondrial genomes could be even greater than reported recently (Kozik et al., 2019).

A high level of convergence is observed in the gene set of plastid genomes of non-photosynthetic plants, whatever parasitic or mycoheterotrophic. They are characterized by a certain degree of reduction, which usually correlates with the time of transition to heterotrophy (e.g. (Samigullin et al., 2016)) and follows the general gene loss model (Barrett et al., 2014). In contrast, mitogenomes of heterotrophic plants are very diverse in terms of structure, size and gene content. In *H. monotropa* the total size of the mitogenome is 810 Kbp, almost twice as large as that of *V. macrocarpon*. However this expansion is unlikely to be associated with the heterotrophy. Large mitogenomes are known for heterotrophic plants, in particular for the mycoheterotrophic orchid *Gastrodia elata* (~ 1.3 Mbp) (Yuan et al., 2018), and the parasitic *Cynomorium* (1 Mbp) (Bellot et al., 2016) and *Lophophytum mirabile* (Balanophoraceae) (~820 Kbp) (in two latter cases the size is shaped by the fragments horizontally transferred from its host – see below). On the other extreme are highly miniaturized mitogenomes of *V. scurruloideum* (Skippington et al., 2015).

4.2. Horisontal gene transfer

HGT is a phenomenon very common in parasitic plants (Yang et al., 2016). HGT from host plants into the mitochondrial genome was shown for Rafflesiaceae (Xi et al., 2013), Orobanchaceae (Yang et al., 2016), *Cynomorium* (Cynomoriaceae) (Bellot et al., 2016), *Lophophytum mirabile* (Balanophoraceae). In the latter case horizontally transferred homologs replaced almost all native mitochondrial genes (Sanchez-Puerta et al., 2017). In contrast, there are no traces of HGT in *H. monotropa* mitogenome. Nuclear and mitochondrial genomes of a mycoheterotrophic orchid *Gastrodia elata* were recently characterized (Yuan et al., 2018) and there were also no observations of HGT. Despite that mycoheterotrophic plants are usually regarded alongside the parasitic plants that fed on other plants, the interactions between plant and

bioRxiv preprint doi: https://doi.org/10.1101/681718; this version posted June 25, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

its host 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 itself. Such connections are similar in many aspects to graft junctions and can be the route 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* was shown (Kim et al., 2014). The transfer of RNA from host is hypothesized to mediate HGT into parasite genome. In contrast, such transfer is not known for mycorrhizal symbiosis. This emphasizes that despite similar –heterotrophic – strategy, plant parasites and MHT plants are fundamentally different in terms of the interaction with their hosts and, potentially, in many other features that stem from this. Thus the knowledge on MHT plant biology should be obtained using MHT models, not by the simple transfer of knowledge from the plant parasites. Many MHT plants are rare endangered plants with very small distribution ranges; in contrast, *H. monotropa* is widespread and is thus prospective as a model MHT plant.

4.3. RNA editing

Extensive RNA sequencing allowed us to infer RNA editing in mitochondrial transcripts. By now the data on RNA-editing in mitochondria for non-photosynthetic plants are scarce. C-to-U RNA editing in seven genes (*atp1*, *atp4*, *atp6*, *cox2*, *nad1*, *rps4*, and *rps12*) was found in *R. cantleyi* (Xi et al., 2013). 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; most editing events are of C-to-U type, usual for plant mitochondria. In many cases RNA editing has clear functional role (e.g. restoration of typical start codon in plastid genes *rpl2* and *psbL* (Kudla et al., 1992). However in *H. monotropa* dEN/dES value in mitochondrial transcripts is 1.23 (does not significantly differ from 1, p-value of 0.060) which indicates that many of editing events are synonymous and thus are unlikely to have functional significance. More detailed examination of RNA editing in mycoheterotrophic plants, including its dynamics in different organs and developmental stages is required to highlight potentially functional events.

Conclusions

Non-photosynthetic plants represent ~1% of plant diversity and are an excellent model for the study of convergent evolution. Until recently the genomics of non-photosynthetic plants was focused on plant parasites; a usual assumption is that mycoheterotrophs - the plants that paraistise on fungi - have basically the same patterns of genome evolution as plant parasites. In order to test this hypothesis and to expand our knowledge on mycoheterotrophic plants we characterized mitochondrial genome of *H. monotropa*. Also, using RNA-seq we performed a genome-wide analysis of gene expression and RNA editing. We showed that the mitogenome structure in *H. monotropa* is highly unusual: it includes small circular fragment and a large linear fragment that has on its ends multiple repeats that, presumably, function as telomeres. Further studies that include characterization of mitogenomes of other Ericaceae and in vivo analysis of *H. monotropa* mitochondria are required to investigate the details of the evolution, replication and functioning of such unusual mitogenome. The gene set is similar to that of autotrophic

bioRxiv preprint doi: https://doi.org/10.1101/681718; this version posted June 25, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

plants. All protein-coding genes are expressed and for the most of them (33 out of 38) we found editing of the transcripts. The intergenic regions of the mitogenome carry multiple sequences of plastid origin, including those of photosynthesis-related genes that are absent in *H. monotropa* plastome. We did not found neither any traces of HGT from fungal hosts in *H. monotropa* mitogenome nor the increase of nucleotide substitution rates. These new data highlight the contrast between mycoheterotrophic and parasitic plants and emphasize the need of the new model objects representing mycoheterotrophic plants.

Funding statement.

This work was supported by the Russian Science Foundation (project #17-14-01315, gene expression analysis) and budgetary subsidy to IITP RAS (project # 0053-2019-0005, genome analysis).

Acknowledgements

The authors thank Alexey Kondrashov (University of Michigan, Ann Arbor) for providing plant material and Aleksey Penin for helpful discussion.

Tables.

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

Supplementary Table S1. Characteristics of mitochondrial genes of *H. monotropa*

Supplementary Table S2. Plastid genes in mitogenome of *H. monotropa*

Supplementary Table S3. Genes with stop codons that are introduced by editing.

Figures.

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*.

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" mate-pair links 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 (8279 bp on average, standard deviation 2583 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 20000 bp. Closely situated reads in the orientations → → and ← ← that make a green "torus" across the genome (1.8% of all mapped mate pair reads) likely represent the artifact of library preparation.

bioRxiv preprint doi: https://doi.org/10.1101/681718; this version posted June 25, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Figure 3. 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.

Supplementary Figure S1. Read mapping characteristics along the mitochondrial chromosomes of *Hypopitys monotropa*.

(A) Average insert size between mate-pair reads spanning over different genomic positions. The nearly uniform insert size distribution suggests that there are no misassemblies involving large deletions or insertions. The fluctuations at the ends of the larger chromosome result from small numbers of reads mapping to the very ends of the chromosome, which is linear.

(B) Number of mate pair fragments covering each of the chromosomes' positions. The absence of positions with zero coverage suggests that there are no misassemblies involving genome fragments' rearrangements. The drops on the ends of the larger chromosome follow from its linearity.

(C) Coverage of the chromosomes by reads of all three sequencing libraries: the paired-end library and both mate-pair libraries. The coverage, though fluctuating, never reaches zero, thus suggesting the absence of misassemblies. Coverage at the ends of the smaller (circular) chromosome abruptly drops approximately sixfold, due to difficulty of mapping reads part of which map to the end and part to the beginning of the contig. In the larger (linear) chromosome, the gradual drops of coverage near the edges and near the positions 90 000 bp and 450 000 bp are due to the varying repeat copy number (see discussion in the main text).

References

Alverson, A.J., Rice, D.W., Dickinson, S., Barry, K., Palmer, J.D., 2011. Origins and recombination of the bacterial-sized multichromosomal mitochondrial genome of cucumber. Plant Cell 23, 2499–2513. https://doi.org/10.1105/tpc.111.087189

Alverson, A.J., Wei, X., Rice, D.W., Stern, D.B., Barry, K., Palmer, J.D., 2010. Insights into the evolution of mitochondrial genome size from complete sequences of *Citrullus lanatus<i/i> and Cucurbita pepo (Cucurbitaceae). Mol. Biol. Evol. 27, 1436–1448. https://doi.org/10.1093/molbev/msq029*

Barrett, C.F., Freudenstein, J.V., Li, J., Mayfield-Jones, D.R., Perez, L., Pires, J.C., Santos, C., 2014. Investigating the path of plastid genome degradation in an early-transitional clade of heterotrophic orchids, and implications for heterotrophic angiosperms. Mol. Biol. Evol. 31, 3095–3112. https://doi.org/10.1093/molbev/msu252

Bellot, S., Cusimano, N., Luo, S., Sun, G., Zarre, S., Gröger, A., Temsch, E., Renner, S.S., 2016. Assembled plastid and mitochondrial genomes, as well as nuclear genes, place the parasite family Cynomoriaceae in the Saxifragales. Genome Biol. Evol. 8, 2214–2230. https://doi.org/10.1093/gbe/evw147

Bergthorsson, U., Richardson, A.O., Young, G.J., Goertzen, L.R., Palmer, J.D., 2004. Massive horizontal transfer of mitochondrial genes from diverse land plant donors to the basal angiosperm Amborella. Proc. Natl. Acad. Sci. 101, 17747–17752. https://doi.org/10.1073/pnas.0408336102

Bjorkman, E., 1960. Monotropa hypopitys L. - an epiparasite on tree roots. Physiol. Plant. 13, 308–327. https://doi.org/10.1111/j.1399-3054.1960.tb08034.x

Boetzer, M., Pirovano, W., 2012. Toward almost closed genomes with GapFiller. Genome Biol. 13, R56. https://doi.org/10.1186/gb-2012-13-6-r56

bioRxiv preprint doi: https://doi.org/10.1101/681718; this version posted June 25, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Bromham, L., Cowman, P.F., Lanfear, R., 2013. Parasitic plants have increased rates of molecular evolution across all three genomes. BMC Evol. Biol. 13, 126.

Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K., Madden, T.L., 2009. BLAST+: architecture and applications. BMC Bioinformatics 10, 421. https://doi.org/10.1186/1471-2105-10-421

Chen, L., 2013. Characterization and comparison of human nuclear and cytosolic editomes. Proc. Natl. Acad. Sci. 110, E2741–E2747. https://doi.org/10.1073/pnas.1218884110

Cho, Y., Qiu, Y.L., Kuhlman, P., Palmer, J.D., 1998. Explosive invasion of plant mitochondria by a group I intron. Proc. Natl. Acad. Sci. U. S. A. 95, 14244–14249.

Cusimano, N., Renner, S.S., 2019. Sequential horizontal gene transfers from different hosts in a widespread Eurasian parasitic plant, Cynomorium coccineum. Am. J. Bot. 106, 679–689. https://doi.org/10.1002/ajb2.1286

Cusimano, N., Zhang, L.-B., Renner, S.S., 2008. Reevaluation of the cox1 group I intron in Araceae and angiosperms indicates a history dominated by loss rather than horizontal transfer. Mol. Biol. Evol. 25, 265–276. https://doi.org/10.1093/molbev/msm241

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, 336–343. https://doi.org/10.1016/j.gene.2013.11.104

Fan, W., Zhu, A., Kozaczek, M., Shah, N., Pabón-Mora, N., González, F., Mower, J.P., 2016. Limited mitogenomic degradation in response to a parasitic lifestyle in Orobanchaceae. Sci. Rep. 6, 36285. https://doi.org/10.1038/srep36285

Grewe, F., Edger, P.P., Keren, I., Sultan, L., Pires, J.C., Ostersetzer-Biran, O., Mower, J.P., 2014. Comparative analysis of 11 Brassicales mitochondrial genomes and the mitochondrial transcriptome of Brassica oleracea. Mitochondrion 19, 135–143. https://doi.org/10.1016/j.mito.2014.05.008

Grigoriev, I.V., Nikitin, R., Haridas, S., Kuo, A., Ohm, R., Otillar, R., Riley, R., Salamov, A., Zhao, X., Korzeniewski, F., Smirnova, T., Nordberg, H., Dubchak, I., Shabalov, I., 2014. MycoCosm portal: gearing up for 1000 fungal genomes. Nucleic Acids Res. 42, D699-704. https://doi.org/10.1093/nar/gkt1183

Gruzdev, E.V., Mardanov, A.V., Beletsky, A.V., Kochieva, E.Z., Ravin, N.V., Skryabin, K.G., 2016. The complete chloroplast genome of parasitic flowering plant Monotropa hypopitys: extensive gene losses and size reduction. Mitochondrial DNA in press.

Gualberto, J.M., Mileshina, D., Wallet, C., Niazi, A.K., Weber-Lotfi, F., Dietrich, A., 2014. The plant mitochondrial genome: Dynamics and maintenance. Biochimie 100, 107–120. https://doi.org/10.1016/j.biochi.2013.09.016

H., H., K., I., H., S., 2002. Structural features and expression analysis of a linear mitochondrial plasmid in rapeseed (Brassica napus L.). Mol. Genet. Genomics 267, 797–805. https://doi.org/10.1007/s00438-002-0711-4

Handa, H., 2008. Linear plasmids in plant mitochondria: peaceful coexistences or malicious invasions? Mitochondrion 8, 15–25.

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 Res. 31, 5907–5916.

bioRxiv preprint doi: https://doi.org/10.1101/681718; this version posted June 25, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Janouškovec, J., Tikhonenkov, D.V., Mikhailov, K.V., Simdyanov, T.G., Aleoshin, V.V., Mylnikov, A.P., Keeling, P.J., 2013. Colponemids Represent Multiple Ancient Alveolate Lineages. Curr. Biol. 23, 2546–2552. https://doi.org/10.1016/j.cub.2013.10.062

Joyce, P.B., Gray, M.W., 1989. Chloroplast-like transfer RNA genes expressed in wheat mitochondria. Nucleic Acids Res. 17, 5461–5476.

Katoh, K., Rozewicki, J., Yamada, K.D., 2017. MAFFT online service: multiple sequence alignment, interactive sequence choice and visualization. Brief. Bioinform. https://doi.org/10.1093/bib/bbx108

Kim, G., LeBlanc, M.L., Wafula, E.K., dePamphilis, C.W., Westwood, J.H., 2014. Plant science. Genomic-scale exchange of mRNA between a parasitic plant and its hosts. Science 345, 808–811. https://doi.org/10.1126/science.1253122

Kozik, A., Rowan, B., Lavelle, D., Berke, L., Schranz, M.E., Michelmore, R.W., Christensen, A.C., 2019. The alternative reality of plant mitochondrial DNA. bioRxiv. https://doi.org/10.1101/564278

Krzywinski, M., Schein, J., Birol, I., Connors, J., Gascoyne, R., Horsman, D., Jones, S.J., Marra, M.A., 2009. Circos: An information aesthetic for comparative genomics. Genome Res. 19, 1639–1645. https://doi.org/10.1101/gr.092759.109

Kudla, J., Igloi, G.L., 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 J. 11, 1099–1103.

Logacheva, M.D., Schelkunov, M.I., Nuraliev, M.S., Samigullin, T.H., Penin, A.A., 2014. The plastid genome of mycoheterotrophic monocot Petrosavia stellaris exhibits both gene losses and multiple rearrangements. Genome Biol. Evol. 6, 238–246. https://doi.org/10.1093/gbe/evu001

Logacheva, M.D., Schelkunov, M.I., Shtratnikova, V.Y., Matveeva, M.V., Penin, A.A., 2016. Comparative analysis of plastid genomes of non-photosynthetic Ericaceae and their photosynthetic relatives. Sci. Rep. 6, 30042. https://doi.org/10.1038/srep30042

Mariotti, M., Lobanov, A.V., Guigo, R., Gladyshev, V.N., 2013. SECISearch3 and Seblastian: new tools for prediction of SECIS elements and selenoproteins. Nucleic Acids Res. 41, e149. https://doi.org/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, 1–13. https://doi.org/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. Mol. Gen. Genet. MGG 247, 61–72.

Olson, A.R., 1993. Patterns of embryo and endosperm formation in Monotropa hypopitys (Monotropaceae) from North America and Sweden. Am. J. Bot. 80, 839–846. https://doi.org/10.1002/j.1537-2197.1993.tb15299.x

Olson, A.R., 1990. Observation on the floral shoots of Monotropa hypopitys (Monotropaceae). Rhodora 92, 54–56.

Petersen, G., Cuenca, A., Møller, I.M., Seberg, O., 2015. Massive gene loss in mistletoe (Viscum, Viscaceae) mitochondria. Sci. Rep. 5. https://doi.org/10.1038/srep17588

Rice, D.W., Alverson, A.J., Richardson, A.O., Young, G.J., Sanchez-Puerta, M.V., Munzinger, J., Barry, K., Boore, J.L., Zhang, Y., dePamphilis, C.W., Knox, E.B., Palmer, J.D., 2013.

bioRxiv preprint doi: https://doi.org/10.1101/681718; this version posted June 25, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Horizontal transfer of entire genomes via mitochondrial fusion in the angiosperm Amborella. Science 342, 1468–1473. https://doi.org/10.1126/science.1246275

Samigullin, T.H., Logacheva, M.D., Penin, A.A., Vallejo-Roman, C.M., 2016. Complete plastid genome of the recent holoparasite Lathraea squamaria reveals earliest stages of plastome reduction in Orobanchaceae. PloS One 11, e0150718. https://doi.org/10.1371/journal.pone.0150718

Sanchez-Puerta, M.V., Abbona, C.C., Zhuo, S., Tepe, E.J., Bohs, L., Olmstead, R.G., Palmer, J.D., 2011. Multiple recent horizontal transfers of the cox1 intron in Solanaceae and extended co-conversion of flanking exons. BMC Evol. Biol. 11. https://doi.org/10.1186/1471-2148-11-277

Sanchez-Puerta, M.V., García, L.E., Wohlfeiler, J., Ceriotti, L.F., 2017. Unparalleled replacement of native mitochondrial genes by foreign homologs in a holoparasitic plant. New Phytol. 214, 376–387. https://doi.org/10.1111/nph.14361

Santesmasses, D., Mariotti, M., Guigó, R., 2017. Computational identification of the selenocysteine tRNA (tRNASec) in genomes. PLoS Comput. Biol. 13, e1005383. https://doi.org/10.1371/journal.pcbi.1005383

Schelkunov, M.I., Penin, A.A., Logacheva, M.D., 2018. RNA-seq highlights parallel and contrasting patterns in the evolution of the nuclear genome of fully mycoheterotrophic plants. BMC Genomics 19. https://doi.org/10.1186/s12864-018-4968-3

Skippington, E., Barkman, T.J., Rice, D.W., Palmer, J.D., 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 Biol. 17. https://doi.org/10.1186/s12870-017-0992-8

Skippington, E., Barkman, T.J., Rice, D.W., Palmer, J.D., 2015. Miniaturized mitogenome of the parasitic plant Viscum scurruloideum is extremely divergent and dynamic and has lost all nad genes. Proc. Natl. Acad. Sci. 112, E3515–E3524. https://doi.org/10.1073/pnas.1504491112

Sloan, D.B., 2013. One ring to rule them all? Genome sequencing provides new insights into the 'master circle' model of plant mitochondrial DNA structure. New Phytol. 200, 978–985. https://doi.org/10.1111/nph.12395

Sloan, D.B., Alverson, A.J., Wu, M., Palmer, J.D., Taylor, D.R., 2012. Recent acceleration of plastid sequence and structural evolution coincides with extreme mitochondrial divergence in the angiosperm genus Silene. Genome Biol. Evol. 4, 294–306.

Xi, Z., Wang, Y., Bradley, R.K., Sugumaran, M., Marx, C.J., Rest, J.S., Davis, C.C., 2013. Massive mitochondrial gene transfer in a parasitic flowering plant clade. PLoS Genet. 9, e1003265. https://doi.org/10.1371/journal.pgen.1003265

Xu, G., Zhang, J., 2014. Human coding RNA editing is generally nonadaptive. Proc. Natl. Acad. Sci. 111, 3769–3774. https://doi.org/10.1073/pnas.1321745111

Yang, Z., Zhang, Y., Wafula, E.K., Honaas, L.A., Ralph, P.E., Jones, S., Clarke, C.R., Liu, S., Su, C., Zhang, H., Altman, N.S., Schuster, S.C., Timko, M.P., Yoder, J.I., Westwood, J.H., dePamphilis, C.W., 2016. Horizontal gene transfer is more frequent with increased heterotrophy and contributes to parasite adaptation. Proc. Natl. Acad. Sci. U. S. A. https://doi.org/10.1073/pnas.1608765113

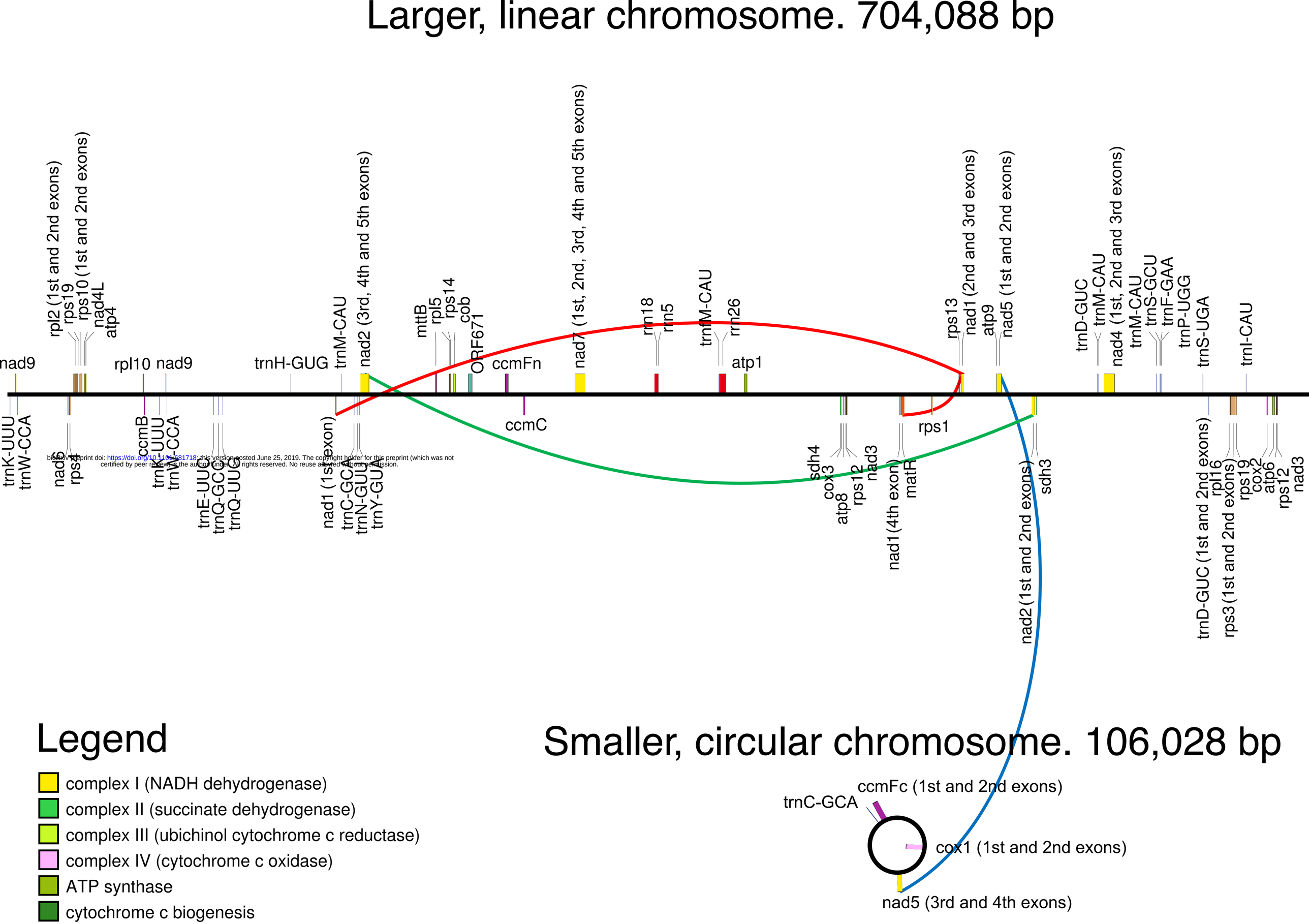
Yuan, Y., Jin, X., Liu, J., Zhao, X., Zhou, J., Wang, X., Wang, D., Lai, C., Xu, W., Huang, J., Zha, L., Liu, D., Ma, X., Wang, L., Zhou, M., Jiang, Z., Meng, H., Peng, H., Liang, Y., Li, R.,

bioRxiv preprint doi: https://doi.org/10.1101/681718; this version posted June 25, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

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. Nat. Commun. 9. https://doi.org/10.1038/s41467-018-03423-5

Zervas, A., Petersen, G., Seberg, O., 2019. Mitochondrial genome evolution in parasitic plants. BMC Evol. Biol. 19. https://doi.org/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, 1047–1057. https://doi.org/10.1534/genetics.114.161414



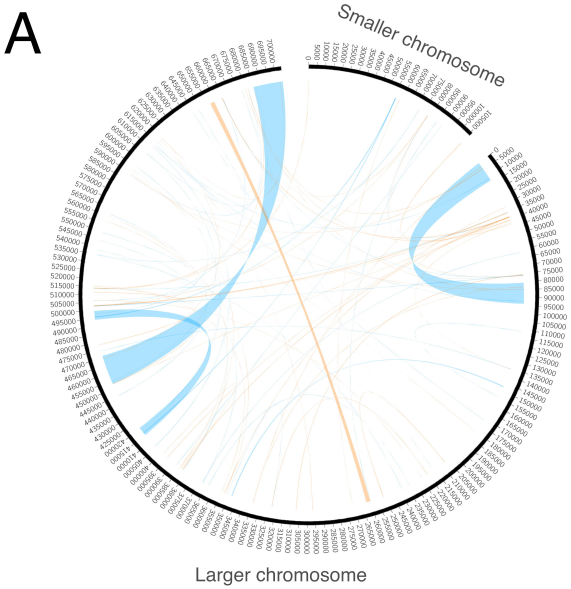
bioRxiv preprint doi: https://doi.org/10.1101/681718; this version posted June 25, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

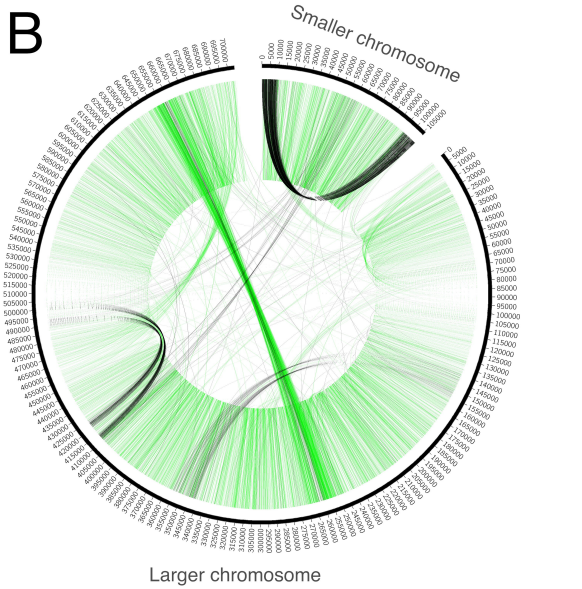
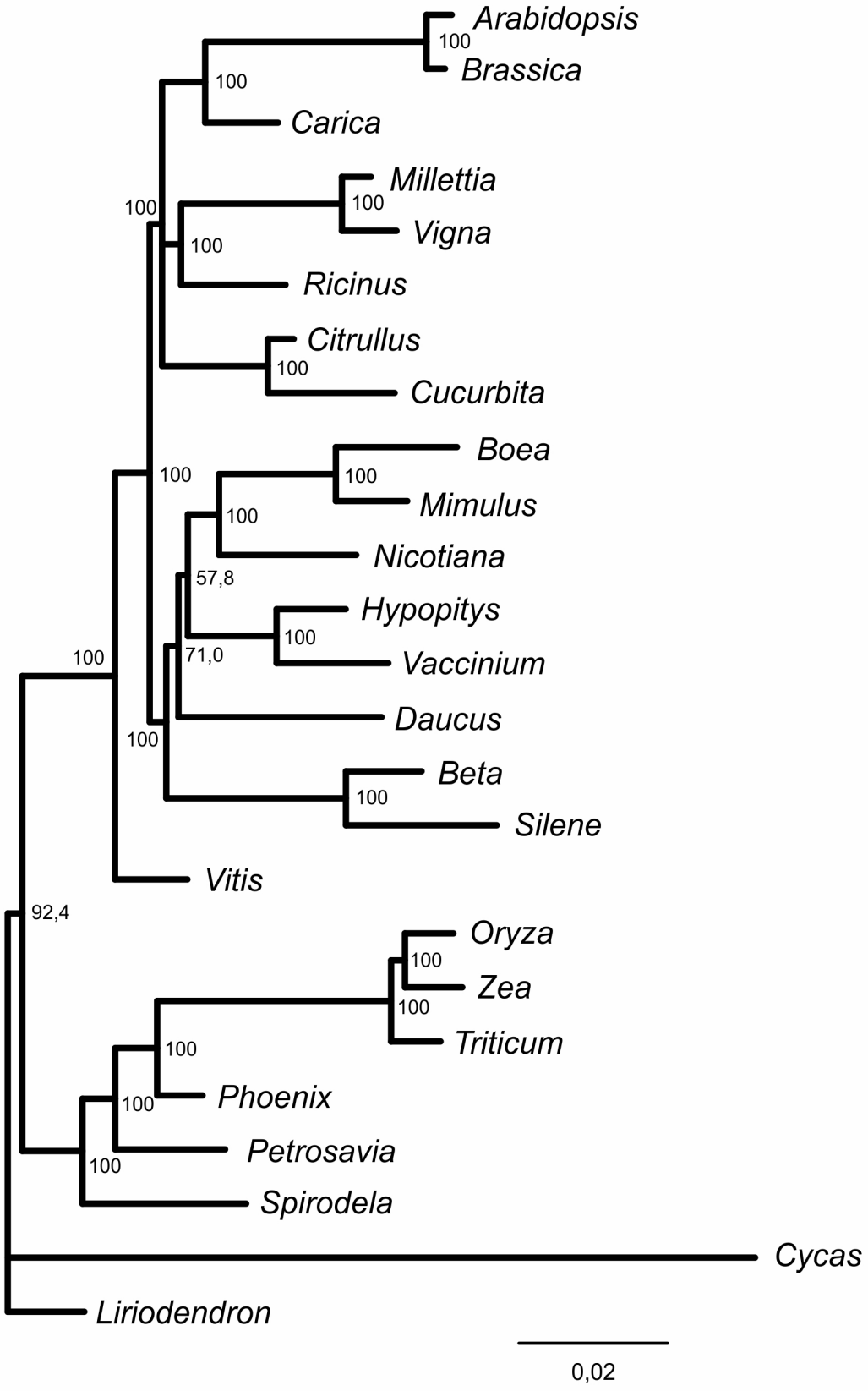
Table 1:
Summary data on the structure and annotation of *H. monotropa* mitogenome.

Chromosome	Accession number (NCBI)	Length	Genes	Pseudo genes	Genes with introns	Genes with cis-splicing	Genes with trans-splicing	Genes with intron I	tRNA	tRNA with introns	rRNA
Larger (linear)	MK990822	704088	35,5	2	7,5*	5	3	0	17-18	0	0
Smaller (circular)	MK990823	106028	2,5	0	2,5*	2	1	1	1	1	2
Total			38	1	10	7	3	1	18-19	1	2

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

bioRxiv preprint doi: https://doi.org/10.1101/681718; this version posted June 25, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.


A
Smaller chromosome
Larger chromosome


B
Smaller chromosome
Larger chromosome

bioRxiv preprint doi: https://doi.org/10.1101/681718; this version posted June 25, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.