New Phytologist

Research

Unparalleled replacement of native mitochondrial genes by foreign homologs in a holoparasitic plant

M. Virginia Sanchez-Puerta[1,2], Laura E. García[1,2], Josefina Wohlfeiler[2] and Luis F. Ceriotti[2]

[1]IBAM, Facultad de Ciencias Agrarias, CONICET, Universidad Nacional de Cuyo, Almirante Brown 500, M5528AHB, Chacras de Coria, Argentina; [2]Facultad de Ciencias Exactas y Naturales, Universidad Nacional de Cuyo, 5500 Mendoza, Argentina

Author for correspondence:
M. Virginia Sánchez Puerta
Tel: +54 261 4135000, ext. 1307
Email: mvsanchezpuerta@fca.uncu.edu.ar

Received: *19 September 2016*
Accepted: *27 October 2016*

New Phytologist (2016)
doi: 10.1111/nph.14361

Key words: evolution, holoparasite, horizontal gene transfer (HGT), *Lophophytum*, mtDNA.

Summary

- Horizontal gene transfer (HGT) among flowering plant mitochondria occurs frequently and, in most cases, leads to nonfunctional transgenes in the recipient genome. Parasitic plants are particularly prone to this phenomenon, but their mitochondrial genomes (mtDNA) have been largely unexplored.
- We undertook a large-scale mitochondrial genomic study of the holoparasitic plant *Lophophytum mirabile* (Balanophoraceae). Comprehensive phylogenetic analyses were performed to address the frequency, origin, and impact of HGT.
- The sequencing of the complete mtDNA of *L. mirabile* revealed the unprecedented acquisition of host-derived mitochondrial genes, representing 80% of the protein-coding gene content. All but two of these foreign genes replaced the native homologs and are probably functional in energy metabolism. The genome consists of 54 circular-mapping chromosomes, 25 of which carry no intact genes.
- The likely functional replacement of up to 26 genes in *L. mirabile* represents a stunning example of the potential effect of rampant HGT on plant mitochondria. The use of host-derived genes may have a positive effect on the host–parasite relationship, but could also be the result of nonadaptive forces.

Introduction

Parasitic plants account for 1% of all flowering plants with *c.* 4200 species in 270–275 genera (Heide-Jorgensen, 2008; Westwood *et al.*, 2010). Hemiparasitic species are capable of photosynthesis but depend upon their host for water and nutrients, although some are facultative parasites. By contrast, holoparasitic plants lack Chl and are completely host-dependent (Heide-Jorgensen, 2008; Westwood *et al.*, 2010). Less than 10% of parasitic plants are strict holoparasites, including almost 400 species in eight angiosperm families (Heide-Jorgensen, 2008). Tremendous changes involving both the phenotype and genotype have been observed in the transition to a parasitic lifestyle (Heide-Jorgensen, 2008; Wicke *et al.*, 2013). Few genomic studies outside plastid genomes have focused on parasitic plants and thus little is known about the independent genomic transformations experienced by each of the *c.* 12–13 parasitic lineages that originated separately during the evolution of flowering plants (Barkman *et al.*, 2007; Westwood *et al.*, 2010). Parasitic relationships facilitate the exchange of genetic information between parasite and host, and make parasitic plants particularly prone to horizontal gene transfer (Davis & Xi, 2015), although the extent and repercussions of this phenomenon remain unclear.

Horizontal gene transfer (HGT) is the lateral transmission of genetic material between unrelated organisms, unlike the more familiar transfer through sexual or asexual reproduction. HGT is common in prokaryotes and has been increasingly reported in multicellular eukaryotes (Keeling & Palmer, 2008). Plants have been recognized as donors and targets of HGT involving the three domains of life (Keeling & Palmer, 2008). Among eukaryotes, the frequency of HGT between flowering plants is surprisingly high (Sanchez-Puerta *et al.*, 2008; Mower *et al.*, 2010; Rice *et al.*, 2013; Davis & Xi, 2015; Park *et al.*, 2015). Plant-to-plant HGT is far more common in the mitochondria than in the chloroplast or nuclear compartments (Keeling & Palmer, 2008; Sanchez-Puerta, 2014) and most often involves mitochondrial sequences from other plants (Bergthorsson *et al.*, 2004; Barkman *et al.*, 2007; Sanchez-Puerta *et al.*, 2008; Mower *et al.*, 2010; Rice *et al.*, 2013). Angiosperms particularly rich in mitochondrial HGT include two free-living species, *Amborella trichopoda* (Rice *et al.*, 2013) and *Geranium brycei* (Park *et al.*, 2015), and the holoparasitic plant *Sapria hymalayana* of the family Rafflesiaceae (Xi *et al.*, 2013; Molina *et al.*, 2014). In most cases, foreign copies are redundant with native genes, often becoming pseudogenes, while only a few mitochondrial foreign genes remain intact and are putatively functional (Xi *et al.*, 2013; Sanchez-Puerta, 2014). In addition, chimeric genes can be formed by gene conversion between the native and transiently present foreign homologs (Bergthorsson *et al.*, 2003; Barkman *et al.*, 2007; Hao *et al.*, 2010; Hepburn *et al.*, 2012; Park *et al.*, 2015). The highly

invasive *cox1* intron represents the most outstanding case of chimeric gene across angiosperms (Cho *et al.*, 1998; Sanchez-Puerta *et al.*, 2008).

Although increasing numbers of plant mitochondrial genome sequences (mtDNAs) become available every year, those of parasitic plants have been largely unexplored. Recently, the first complete mtDNAs from two hemiparasitic plants of the genus *Viscum* (family Viscaceae, order Santalales) uncovered the massive loss of genes encoding ribosomal and respiratory proteins, with no evidence of HGT, except for the *cox1* intron (Petersen *et al.*, 2015; Skippington *et al.*, 2015). Additional insight comes from a large-scale study of mitochondrial sequences from the endophytic holoparasitic family Rafflesiaceae (Xi *et al.*, 2013; Molina *et al.*, 2014). This study revealed that a sizeable fraction of the gene content appears to have been obtained from their host plants in both ancient and recent horizontal transfers, and that the extent of HGT was higher in the mtDNA than in the nuclear genome (Xi *et al.*, 2012, 2013). Even though most of the foreign mitochondrial sequences in Rafflesiaceae are full-length and transcribed, the majority are redundant with native genes, which are actively transcribed and most likely still functional (Xi *et al.*, 2013). Overall, despite the great diversity of parasitic plants, not much is known about the properties of their mtDNAs and the impact of horizontal acquisition of genes from their hosts.

The family Balanophoraceae *sensu stricto* encompasses 14 genera of obligate root parasites and is a member of the order Santalales (Su *et al.*, 2015). This family represents an independent origin of holoparasitism (Barkman *et al.*, 2007) and presents an opportunity to explore the potential complexities and properties particular of this lineage, and to provide an independent comparison of the mitochondrial features reported in the only two lineages of parasitic plants studied so far. The genus *Lophophytum* belongs to this family and includes five species, all of which are exclusively found in South America (Hansen, 1980). *Lophophytum mirabile* has a more or less spherical subterranean vegetative body or 'tuber' with only mature inflorescences emerging above soil level (Gonzalez & Mauseth, 2010). It is found mainly parasitizing roots of trees of *Anadenanthera colubrina* (Fabaceae) (Sato, 2014), although one study reported three other mimosoid legumes as hosts (Hansen, 1980).

To contribute to our understanding of the genomic evolution of parasitic plants, we sequenced and characterized the complete mitochondrial genome of the holoparasitic plant *L. mirabile* (Balanophoraceae). We performed comprehensive evolutionary analyses to address the frequency, origin, and impact of HGT on the mtDNA of *L. mirabile*, and uncovered an extraordinarily high degree of gene transfers from its host plant.

Materials and Methods

Plant material and DNA extraction

Plants of the parasitic angiosperm *Lophophytum mirabile* Schot. & Endl. ssp. *bolivianum* (Wedd.) B. Hansen (H. A. Sato 430) and its host *Anadenanthera colubrina* var. *cebil* (Griseb.) Altschul (H. A. Sato 429) were collected on 8 January 2013 from the Calilegua National Park, Jujuy, Argentina. Samples of the holoparasite *L. leandri* Eichler (H. A. Sato 438) and its host *Parapiptadenia rigida* (Benth.) Brenan (H. A. Sato 439) were collected on 21 July 2013 from the city of San Ignacio, Misiones, Argentina. These populations of *L. mirabile* and *L. leandri* have been thoroughly characterized (Gonzalez & Mauseth, 2010; Sato & Gonzalez, 2013, 2016; Sato, 2014). DNA was extracted from leaves of individual plants of *A. colubrina* and *P. rigida* using a cetyl-trimethyl-ammonium-bromide DNA-extraction protocol (Doyle & Doyle, 1987). Genomic DNA from the holoparasites was extracted from the inflorescence (*L. mirabile*) or from sections of the tuber sufficiently distant from host tissue (*L. leandri*) using the DNeasy Plant Mini kit (Qiagen). Sample preparation was done with extreme care to avoid contamination from host tissue (see later).

Genome sequencing, assembly, and validation

DNA from *L. mirabile* was used to construct a 700 bp paired-end library and was sequenced at the Beijing Genomics Institute using the Illumina HiSeq 2000 sequencing system (San Diego, CA, USA). This generated 65 million clean 100 bp paired-end reads that were assembled on the Mason large-memory computer cluster at Indiana University-Bloomington (USA). For the reconstruction of the mtDNA, we followed the methodology described by Sanchez-Puerta *et al.* (2015). First, assembly parameters were optimized with the program Velvet Optimiser 2.2 (Simon Gladman, CSIRO and Monash University, Australia). Then, we performed *de novo* assembly of the mtDNA of *L. mirabile* using Velvet 1.2.03 (Zerbino & Birney, 2008) without scaffolding and with the optimized parameters: hash length of 81, expected coverage of 20, and coverage cutoff of 2.8. Velvet assembled 281 *de novo* contigs larger than 1 kb, with N50 of 11 234 bp and a maximum contig size of 27 149 bp. Taking advantage of typical differences in read depth between nuclear and organellar sequences (Straub *et al.*, 2012), 83 mitochondrial contigs (average read depth of 80) were identified. Manual editing, joining, and closing of the mitochondrial contigs were done based on consistent paired-end reads visualized in Consed v.29 (Gordon & Green, 2013). The resulting organellar genome assembly of 54 chromosomes was compared against all Illumina reads to identify errors until the read depth and paired-end read depth were as expected at each base and the high-quality mismatches were very low or zero. The assembly showed a fairly even read depth of 80× and contained no gaps or low-coverage regions. A circular organization of all mitochondrial chromosomes was supported by paired-end reads having one mate mapping to the start of a contig and the other to the end.

Raw sequence data are available from the NCBI Sequence Read Archive as accession number SRP071697. The annotated *L. mirabile* mitochondrial chromosomes were deposited in the GenBank data libraries under accession numbers KU992322–KU992380 and KX792461.

Characterization of gene and repetitive content in the mtDNA

The mtDNA was annotated using MITOFY (Alverson *et al.*, 2010), BLAST (Altschul *et al.*, 1990), and the tRNAscan-SE algorithm (Lowe & Eddy, 1997). Graphical genome maps were generated using OGDRAW software (Lohse *et al.*, 2007). Nonsynonymous sites of RNA editing were predicted with PREP-Mt (Mower, 2005) with a cutoff value of 0.5. Repeats larger than the DNA library size (*c.* 700 bp) were detected by mapping all Illumina reads to the mitochondrial assembly using the program CONSED v.29 (Gordon & Green, 2013) and comparing the read depth across the genome. We assembled each repeat pair into a single chromosome to minimize subcircular maps, even though this is only one of the possible alternative representations given the known high recombinational activity of large repeats in mtDNAs. Dispersed perfect and imperfect (> 80% identity) repeats shorter than 700 bp were identified by CONSED v.29. Recombination between these repeats was evaluated based on paired-end reads with the aid of CONSED v.29.

Sequencing of selected mitochondrial genes and phylogenetic analyses

Maximum likelihood (ML) and maximum parsimony (MP) phylogenetic analyses were performed for each of the mitochondrial protein and rRNA genes, as well as for the plastid- and nuclear-derived sequences. Each mitochondrial gene data set included *L. mirabile* and 38 free-living angiosperms (Supporting Information Table S1) and additional species of Fabales, Santalales, and Balanophoraceae when available. Sequences from the hemiparasitic genus *Viscum* (Viscaceae, Santalales) were excluded from the analyses because of the extremely fast substitution rate of the mitochondrial genes (Petersen *et al.*, 2015; Skippington *et al.*, 2015). Selected mitochondrial genes of the holoparasite *L. leandri* and the legume hosts were amplified by PCR with primers that had been developed earlier (Bergthorsson *et al.*, 2004). PCR products were sequenced using an ABI 3730 (Applied Biosystems, Waltham, MA, USA). Nucleotide sequences were aligned manually with MACCLADE 4.0 (Maddison & Maddison, 2000). RNA editing sites were removed from the alignments before the analyses because both genomic and cDNA sequences were included (Table S1). ML analyses were performed with GARLI 2.0 (Zwickl, 2006) under the General Time Reversible model with parameters for invariable sites and gamma-distributed rate heterogeneity (GTR + I + Γ4; four rate categories). A thousand rapid bootstrap replicates were done under the same model of evolution using RAxML (Stamatakis, 2014). MP analyses and a thousand bootstrap replicates were run in PAUP*4.0 (Swofford *et al.*, 2002). No strongly supported differences were observed in the topologies of the ML and MP trees; thus, MP bootstrap values are shown on the ML trees below each branch.

Alternative topology tests were performed using the approximately unbiased (AU) test as implemented in the CONSEL package (Shimodaira, 2002). We tested two different constrained topologies: monophyly of the Fabaceae excluding a clade with *L. mirabile* and all sequences from other available Santalales in each data set (constraint #1); and the monophyly of *L. mirabile*, other available Santalales, asterids, and Caryophyllales (constraint #2), as expected in vertically inherited genes (The Angiosperm Phylogeny Group, 2016). The most likely tree under each constraint was obtained by searching for the best tree compatible with that constraint using PAUP*4.0 (Swofford *et al.*, 2002). The site likelihoods for the constrained trees and for the best tree in the unconstrained analysis were exported from PAUP*4.0, and the AU *P*-values were calculated from these data.

Sequence analyses

We evaluated the possibility that some of the genes in *L. mirabile* were chimeric, as a result of recombination between host and native genes, using the program onepop of the OrgConv 1.1 package (Hao, 2010). Given the lack of data from close relatives to *L. mirabile*, we were only able to analyze those genes for which sequences from other Santalales were available (*atp1*, *atp4*, *atp6*, *atp8*, *cox1*, *matR*, *nad1x2x3*, *nad3*, *nad4L*, *nad5x1x2*, *nad9*, *rps3*, *rrnS*). The input files consisted in individual gene alignments of sequences of *L. mirabile,* Santalales and mimosoid legumes. The Bonferroni correction for multiple comparisons was considered (Hao, 2010).

To analyze the selective pressure on *L. mirabile* genes, we tested individual protein coding regions in a phylogenetic context with PAML v.4.8 (Yang, 2007). For each gene, a ML phylogenetic tree was constructed using GARLI 2.0 (Zwickl, 2006) based on a reduced data set, from which those taxa with accelerated substitution rates, such as *Silene*, *Ajuga* and *Geranium*, and sequences based on cDNAs (*Acacia*, *Prosopis*), were excluded. To test if there were differences in the ratio of synonymous (d*S*) and nonsynonymous (d*N*) substitution rates ($\omega = dN/dS$) on branches of *L. mirabile* compared with the rest of the angiosperms, we contrasted two likelihood models. In the first case, it was assumed that all branches have the same ω ratio (model = 0 in codeml), while in the second case two different ω ratios were allowed for *L. mirabile* and the other branches (model = 2 in codeml), respectively. Both models were compared using a likelihood ratio test (LRT) and *P*-values were calculated based on a χ^2 distribution with 1 df. In addition, we aimed to detect positive selection affecting only a few sites in *L. mirabile* (foreground branch) compared with the rest of the branches on the tree (background branches) using a branch-site model in codeml. The null model (no sites under positive selection) was tested against the alternative model (some sites under positive selection on the foreground branch), using a χ^2 distribution as described earlier, with 1 df. Positively selected codons were identified by the Bayes empirical Bayes inference (Yang *et al.*, 2005).

Assessment of host contamination in the genomic assembly

We ruled out the possibility of contamination from host tissue in the mitochondrial genome assembly of *L. mirabile* for several reasons. First, a section of the inflorescence of

L. mirabile was carefully dissected for DNA extraction. Given that the inflorescences of *Lophophytum* are not in direct contact with host tissue (Gonzalez & Mauseth, 2010), the risk of host contamination is negligible. Second, no complete plastid genes were found in the assembly. If there were any host contamination, plastid sequences from the host should be easily detected. Third, 13 chromosomes contain both foreign and native regions, indicating that these sequences are clearly integrated into the mitochondrial genome of *L. mirabile*. Fourth, all chromosomes show relatively even read-depth coverage, including regions identified as foreign. A similar read-depth between native and foreign sequences is not expected if transgenes were the result of host contamination or were located in the nuclear genome instead of in the mtDNA of the parasite. Fifth, PCR products of six mitochondrial genes were generated from independent DNA extractions of *L. mirabile* (inflorescence) and *L. leandri* (tuber samples from two individuals). In all cases, the sequences obtained from both specimens of *L. leandri* were identical to each other, and they shared the same evolutionary history with *L. mirabile* homologs. Lastly, those six genes showed sequence divergence when compared with homologs from their host plants. The mean p-distance between the DNA sequences of *L. mirabile* and *A. colubrina* was 0.019, similar to the divergence between *L. leandri* and its host *P. rigida* (mean *P*-distance = 0.024).

Results

The multichromosomal *Lophophytum mirabile* mitochondrial genome

The mtDNA of the holoparasite *L. mirabile* is 821 919 bp long and assembles into 54 circular-mapping chromosomes, the largest of which is 58 274 bp and the shortest of which is 7177 bp (Fig. 1). No other mtDNAs from this plant family are available for comparison. The GC content of the *L. mirabile* mtDNA is 44.5%, ranging between 42.2 and 46.1% among chromosomes, comparable to other flowering plants. Direct and inverted repeats represent 14% of the genome (Table 1). Only repeats shorter than the average library insert size (700 bp) could be evaluated for their recombinational activity based on paired-end reads. Out of 55 repeats of length 150–700 bp, only 11 showed evidence of the alternative configuration (Fig. 1, underlined). For those chromosomes with recombining repeats, the majority of sequencing paired-end reads support the genomic conformation shown in Fig. 1.

The final assembly of the *L. mirabile* mtDNA showed an average sequencing depth of *c.* 80× (Fig. S1). Mapped read depth in a genome assembly can be used to estimate the relative abundance of different sequences. The fairly even coverage across the 54 chromosomes of *L. mirabile* mtDNA indicates that they have a similar abundance (Fig. S1). The only exception is chromosome 39 that shows four times the read depth across half its sequence. That region is similar to a transposable element and thus the increased coverage is probably a result of mismapping of reads that were derived from the nucleus.

The *Lophophytum mirabile* mtDNA gene content

The mitochondrial genome of *L. mirabile* presents 35 protein, three rRNA, and six tRNA genes, but many are present in multiple copies, resulting in a total gene content of 56 (Table 1). Intron content includes 22 group II introns and one group I intron in the gene *cox1* (Table 1). RNA editing sites were predicted in all full-length protein-coding genes (Table 2).

Only 29 chromosomes of *L. mirabile* mtDNA contain all the full-length mitochondrial genes, whereas 25 are devoid of intact known genes (Fig. 1). Nevertheless, all chromosomes contain matches to other sequenced plant mitochondrial genomes. About 30% of the mtDNA of *L. mirabile* shares similarity with other mitochondrial genomes (including genic and nongenic sequences), while *c.* 67% lacks detectable similarity to sequences in GenBank (Table 1). The *L. mirabile* mtDNA also contained chloroplast- and nuclear-derived regions, representing 0.6% and 2.2% of the genome, respectively, although the nuclear contribution may be greater but unrecognizable at this stage because of lack of genomic data. The identified nuclear-derived sequences were truncated retroelements (1.6%) or DNA transposons (0.03%), and typically nuclear-encoded gene fragments (0.55%).

Rampant horizontal gene transfer from its host in *L. mirabile* mtDNA

About 80% of the 44 protein-coding gene copies found in the *L. mirabile* mtDNA showed a foreign origin (Table 2). All 35 foreign sequences were placed phylogenetically with species of the family Fabaceae, almost always as sister to or within the subfamily Mimosoideae, to which the *L. mirabile* host plant (*Anadenanthera colubrina*) belongs (Fig. S2). The genes *cox3*, *nad3*, *rpl5*, *rpl16*, *rps3*, *rps12*, *rps14* and *sdh4* form part of gene clusters that are conserved in most angiosperms and are separated by a few nucleotides (3–48 nt) or have overlapping open reading frames (ORFs). Because HGT typically does not correlate with gene boundaries, each gene cluster was analyzed as a whole (Fig. S2). One example of host-to-parasite HGT is the phylogeny of *atp1* in which *Lophophytum* spp. sequences formed a strongly supported clade with the subfamily Mimosoideae (Fig. 2a). The affiliation of *L. mirabile* with the Fabaceae in the phylogenies of the 35 foreign protein genes was supported by bootstrap support (BS) values ≥ 75% (Fig. S2). Furthermore, vertical inheritance of 20 foreign genes was rejected by the AU test (Table 2; Fig. S3). However, the topologies that were not rejected by the AU test are not necessarily supported (Shimodaira, 2002). In fact, only the unconstrained (best) trees were strongly supported by the AU test. Given the lack of sequence data of *L. mirabile* relatives in the individual gene data sets, constraints for the AU tests were difficult to apply, and, in some cases, impossible (*ccmC*, *nad2*), limiting the utility of the test considerably. The exclusion of *L. mirabile* from the clade of Fabaceae (constraint #1 in Fig. S3) was the most useful constraint, particularly when no other Santalales were available. The best tree under such a constraint always showed *L. mirabile* as sister to the Fabaceae. Thus, even though the AU test did not reject that hypothesis, both trees (the

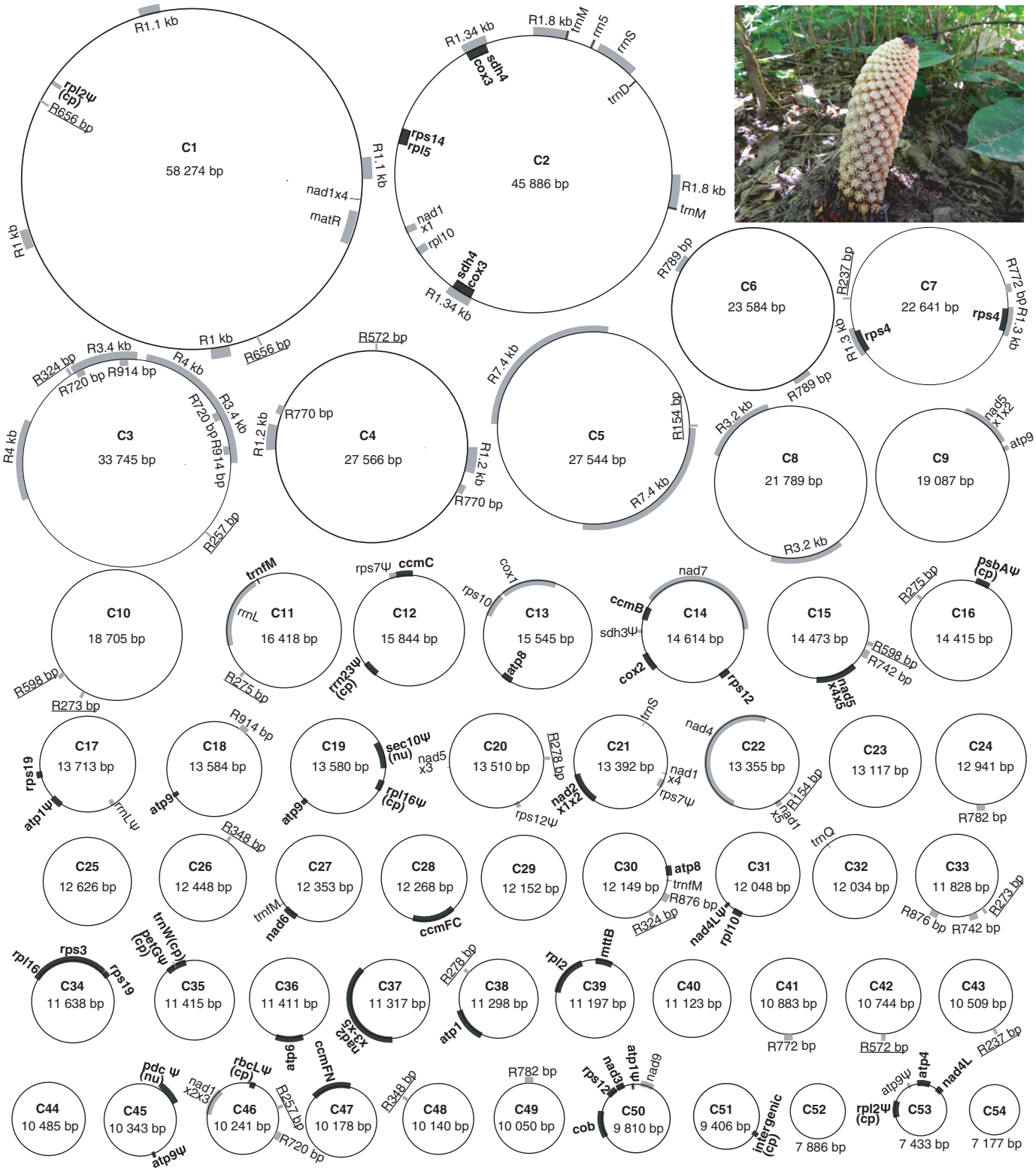

Fig. 1 Map of the mitochondrial genome (mtDNA) of *Lophophytum mirabile*. The mtDNA is 821 919 bp long and is subdivided in 54 circular chromosomes (C*n*) of diverse lengths. Genes drawn inside and outside each circle are transcribed clockwise and counterclockwise, respectively. Shown are full-length genes, pseudogenes > 100 bp (indicated by 'Ψ'), repeats > 700 bp (labeled 'R', followed by the repeat length), and sequences derived from the chloroplast (cp) or nucleus (nu) longer than 100 bp. Repeats for which there is evidence of recombination are underlined. Gene names in black or gray font are foreign or native, respectively. The photograph shows an inflorescence of *L. mirabile* (courtesy of H. A. Sato).

Table 1 Features of the mitochondrial genome of *Lophophytum mirabile*

Genome length (bp)	821 919
Protein-coding genes[a]	35 (44)
rRNA genes[a]	3 (3)
tRNA genes[a]	6 (9)
Group II introns	
cis-splicing	16
trans-splicing	6
Group I introns	1
Repeats (kb; % of genome)[b]	113.3 (14.0%)
Large repeats (> 1 kb)[b]	51.81 (6.4%)
Chloroplast-derived sequences	0.6%
Nuclear-derived sequences	2.2%
Mitochondrial genes (exons and cis-spliced introns)	7.68%
Mitochondrial-like (including genes)	30.05%
Uncharacterized	67.17%

[a]First value excludes duplicates; value in parentheses includes them.
[b]Total length of repeats.

unconstrained and the constrained one) showed a legume origin for *L. mirabile* genes.

The remaining protein gene copies present in the mtDNA showed *L. mirabile* in an unsupported position probably because of the lack of data from close relatives, and were classified as putatively native (Table 2). One example of native genes is the phylogeny of the gene *matR*, in which *L. mirabile* sequences are related to the Santalales (Fig. 2b). In addition, two of the six tRNAs in *L. mirabile* (*trnfM*, *trnW-cp*) are foreign and were probably acquired from the legumes (Figs S2, S4). Two rRNA genes (*rrn5*, *rrnS*) showed no evidence of a foreign origin, while the *rrnL* sequence was placed with the legumes, albeit with limited statistical support (Fig. S2).

Foreign genes are known occasionally to recombine with native homologs upon arrival into the recipient mitochondria. We evaluated evidence for gene conversion events in *L. mirabile* genes, for which comparative data from Santalales were available (*atp1*, *atp4*, *atp6*, *atp8*, *cox1*, *matR*, *nad1x2x3*, *nad3*, *nad4L*, *nad5x1x2*, *nad9*, *rps3*, *rrnS*). The program onepop from the OrgConv package only detected significant evidence for recombination in one (*atp6*) of the 13 genes analyzed (excluding the known chimeric origin of the intron-containing *cox1*). The gene *atp6* shows a 150 nt recombinant segment with L/N and L-N *P*-values of $2.51e^{-4}$ and $9.7e^{-04}$, respectively.

Besides the 56 full-length genes described earlier, the *L. mirabile* mtDNA has 10 mitochondrial pseudogenes > 100 bp, four of which were related to the Fabaceae (Fig. S5). In total, a minimum of 55 kb (6.78%) of the mtDNA of *L. mirabile* were probably transferred from the legumes, including *c.* 30 kb of foreign genes and *c.* 25 kb of nongenic sequences only similar to Fabaceae mtDNAs. A greater fraction of legume-derived sequences in *L. mirabile* mtDNA may be recognized when mimosoid mitochondrial genome sequences become available.

Test of selection of *Lophophytum mirabile* protein genes

For each protein-coding gene in the *L. mirabile* mtDNA, we tested whether different selective pressures can be detected on the holoparasite. The ω ratio ($dN:dS$) measures the direction and magnitude of selection on proteins, with $\omega > 1$ indicating positive selection, $\omega < 1$ purifying selection, and $\omega = 1$ neutral selection (Kimura, 1983; Messier & Stewart, 1997). The most basic model, which assumes a single ω ratio across all branches could

Table 2 Features of protein-coding genes in *Lophophytum mirabile* mtDNA

Gene name[a]	No. of intact genes	No. of predicted editing sites	Origin[b]	AU test[c]
cox1	1	12	Native	na
matR	1	15	Native	na
nad1	1	38	Native	na
nad4	1	40	Native	na
nad5x1x2	1	14	Native	na
nad5x3	1	0	Too short	na
nad7	1	25	Native	na
nad9	1	8	Native	na
rps10	1	2	Native	na
atp9	3	1;2;5	Foreign (2); native	ns (both)
rpl10	2	4;5	Foreign; native	$P < 0.05$
atp1	1	3	Foreign	$P < 0.05$
atp4	1	11	Foreign	$P < 0.05$
atp6	1	18	Foreign	$P < 0.05$
atp8	2	3; 3	Foreign (2)	$P < 0.05$ (one)
ccmB	1	26	Foreign	$P < 0.05$
ccmC	1	30	Foreign	nd
ccmFC	1	17	Foreign	ns
ccmFN	1	31	Foreign	ns
cob	1	14	Foreign	ns
cox2	1	10	Foreign	$P < 0.05$
cox3	2	13;9	Foreign (2)	$P < 0.05$ (both)
mttB	1	26	Foreign	$P < 0.05$
nad2x1x2	1	16	Foreign	nd
nad2x3x5	1	13	Foreign	nd
nad3	1	11	Foreign	$P < 0.05$
nad4L	1	9	Foreign	$P < 0.05$
nad5x4x5	1	8	Foreign	$P < 0.05$
nad6	1	11	Foreign	ns
rpl2	1	6	Foreign	$P < 0.05$
rpl5	1	9	Foreign	ns
rpl16	1	6	Foreign	ns
rps3	1	7	Foreign	ns
rps4	2	18;19	Foreign (2)	$P < 0.05$ (one)
rps12	2	5; 5	Foreign (2)	$P < 0.05$ (both)
rps14	1	3	Foreign	ns
rps19	2	3; 3	Foreign (2)	$P < 0.05$ (one)
sdh4	2	3; 3	Foreign (2)	$P < 0.05$ (both)

[a]Intron-containing genes are shown in bold face.
[b]Origin based on phylogenetic analyses. Numbers in parenthesis indicate the no. of intact genes.
[c]Details on the approximately unbiased (AU) test can be found in Supporting Information Fig. S3; na, not applicable; nd, not determined; ns, not significant at $P = 0.05$.

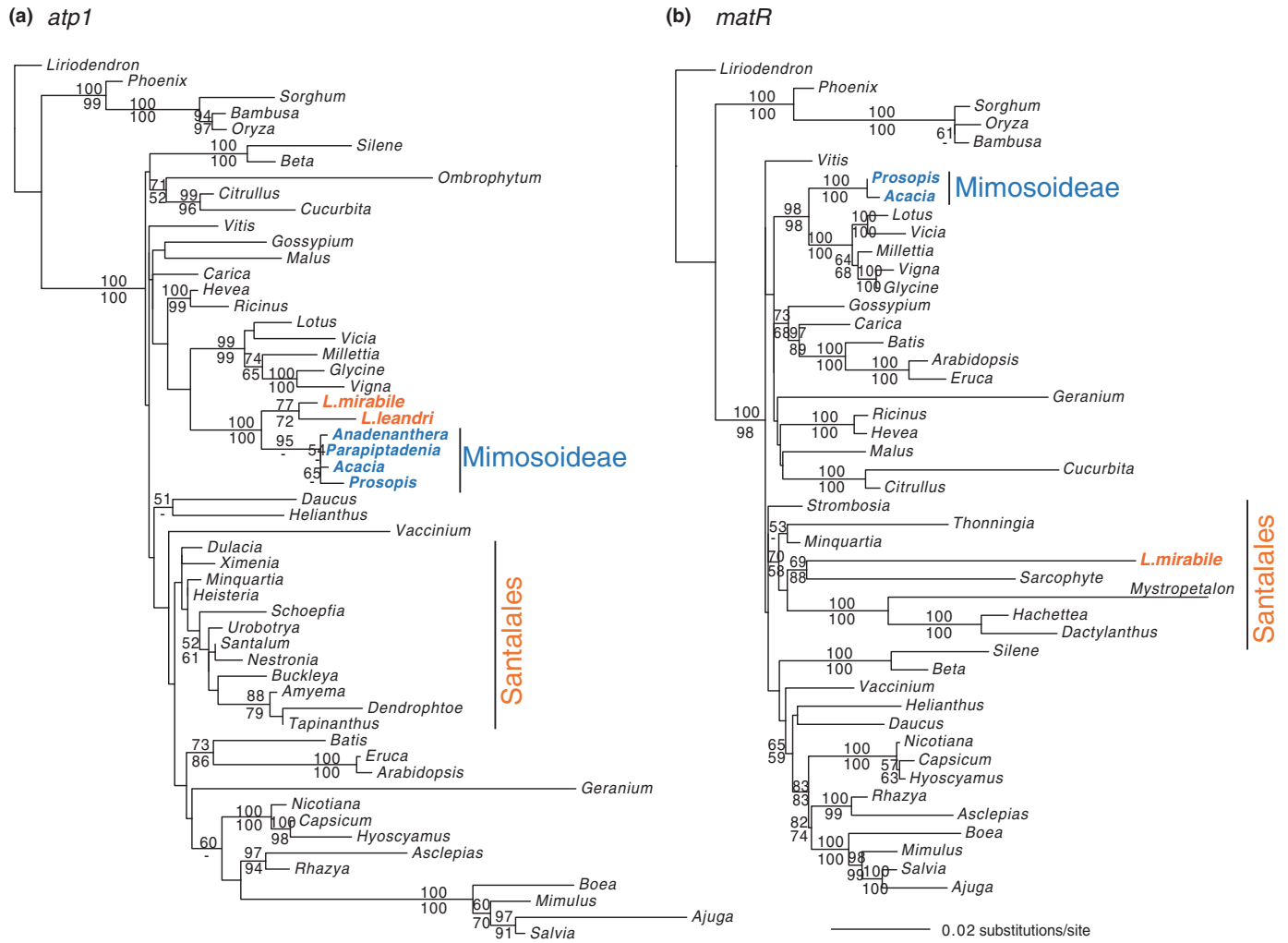

Fig. 2 Maximum likelihood phylogenetic analyses of mitochondrial genes of *Lophophytum mirabile*. (a) Phylogeny of the gene *atp1*. This tree provides an example of gene transfer from host (subfamily Mimosoideae, Fabaceae; in blue) to parasite (*Lophophytum* spp.; in orange). (b) Phylogeny of the gene *matR*. This tree provides an example of vertically inherited sequences in *Lophophytum*. ML and MP bootstrap support values ≥ 50% are shown above and below each branch, respectively. Scale bar corresponds to substitutions per site.

not be rejected, except for two genes (*atp9*_C9 and *nad9*) that showed significant differences in ω values in *L. mirabile* compared with the rest of the angiosperms (Table S2). In general, the branch model estimated $\omega < 1$ for all but five protein-coding genes in *L. mirabile* (*nad1*x5, *nad9*, *atp6*, *atp8*_C11, *ccmB*), suggesting they are evolving mostly under purifying selection (Table S2). In the case of the genes with $\omega > 1$, four of them had very small dS values (d$S < 0.02$), which might lead to an overestimation of the ω ratio. A comparison of the synonymous rates and ω values among foreign and native *L. mirabile* genes showed great variability but not marked differences (Table S2).

Because in the basic and branch models the ω ratio is averaged over all sites, and it is rarely > 1 (Yang, 2007), we used a branch-site model to detect positive selection in individual codons of *L. mirabile* genes. For most protein-coding regions, we did not find significant differences in selection patterns (Table S2). For seven genes the branch-site model fitted better than the null model, which indicates that a few sites evolved under positive selection on the *L. mirabile* branch (Table S2). Positively selected codons with posterior probability > 0.95 were identified for the foreground branch: *ccmFN* (147T, 346K, 576H), *cox2*, *matR* (181A, 191H, 207P, 208F, 243T, 245K,247E, 248F, 249Y, 257K), *nad9* (4Q, 16K, 19V, 20K, 53R, 110G, 134S, 154V, 178Y, 182A), *rpl2* (125G), and *rps3* (273V).

Some of the horizontal transfers predate speciation events in the parasitic lineage

To estimate the timing of some of the horizontal gene transfers, we gathered comparative data from another holoparasitic species of the genus *Lophophytum*, *L. leandri*, and also from the current legume hosts of *L. leandri* and *L. mirabile*: *P. rigida* and *A. colubrina*, respectively. In three cases (*atp1*, *atp8* and *cob*) the two species of *Lophophytum* grouped together with high bootstrap support and were sister to the subfamily Mimosoideae, but were not particularly related to their hosts, suggesting an older horizontal transfer that predates the speciation events in *Lophophytum* (Figs 2a, S2). In *rpl5* (Fig. S2), *L. leandri* was sister to

Anadenanthera + *Parapiptadenia* with high bootstrap support (BS = 86%) and *L. mirabile* was sister to *Acacia* with lower support (BS = 67%); this would weakly suggest two independent acquisitions of the gene *rpl5* from the mimosoid lineage. In *ccmC* there was no strong support for relationships within the *Lophophytum* + Fabaceae clade to reach a conclusion (Fig. S2). In addition, the *cox1* intron (Fig. S2) has been horizontally transferred from an as yet undetermined donor before the divergence of *Lophophytum* spp. and *Ombrophytum subterraneum* (both Balanophoraceae).

All chloroplast- and nuclear-derived sequences in the *L. mirabile* mtDNA were foreign

Eight chloroplast-derived sequences of length > 100 bp were identified in the mtDNA of *L. mirabile* (Fig. 1). These regions contained mostly gene fragments, except for short genes such as *petG* and *trnW*. Six of the plastid sequences were related to legumes, in particular to the subfamily Mimosoideae (Fig. S4). The remaining two chloroplast-derived regions (*rpl2* and *psbA*) showed weakly supported phylogenetic relationships with the lineages Malpighiales and magnoliids, respectively (Fig. S4). All chloroplast-derived sequences integrated in the mtDNA of *L. mirabile* were foreign, indicating that no native chloroplast sequences were transferred to the mitochondrion by intracellular gene transfer as reported in all other plant mitochondrial genomes (Kubo & Mikami, 2007).

Five nuclear gene fragments reside in the *L. mirabile* mtDNA. Only two were long enough for meaningful phylogenetic analyses, encoding a pyruvate decarboxylase (*pdc*) and an exocyst complex component (*sec10*). Both of them grouped with sequences of the tribe Mimosoideae with strong bootstrap support (Fig. S6). The lack of introns in the foreign nuclear genes indicates that the transfer was RNA-mediated and then integrated into the mitochondrial genome via reverse transcription.

Discussion

Highly multichromosomal mtDNA

The mitochondrial genome of *L. mirabile* assembles into 54 largely autonomous, circular chromosomes (7–58 kb in length). A multipartite chromosomal architecture was first described among plants in the cucumber mitochondria with three circular chromosomes of lengths 45, 84 and 1556 kb, respectively (Alverson *et al.*, 2011). Later, four to 128 chromosomes (5–394 kb in length) were described in *Silene* spp. mtDNAs (Sloan *et al.*, 2012) and five circular-mapping mitochondrial chromosomes (119–3179 kb) in *A. trichopoda* (Rice *et al.*, 2013). In contrast to the typical angiosperm mtDNA, the aforementioned multichromosomal genomes cannot be mapped to a single circular chromosome. Multicircular mitochondrial genomes have also been reported in other eukaryotic lineages, such as metazoa, fungi, and kinetoplastids (Burger *et al.*, 2003).

It is puzzling that almost half (25 out of 54) of the chromosomes of *L. mirabile* mtDNA are devoid of intact known genes, even though they have regions with similarity to other sequenced plant mitochondrial genomes. Likewise, 20 of the 59 chromosomes in *Silene noctiflora*, 86 of the 128 chromosomes in *S. conica* (Sloan *et al.*, 2012), and the two smaller chromosomes of cucumber (Alverson *et al.*, 2011) bear no intact genes. The existence of possibly noncoding chromosomes raises questions about their origin and perpetuation, for which both neutral and adaptive explanations have been advanced, but the questions remain unanswered (Alverson *et al.*, 2011; Sloan *et al.*, 2012; Rice *et al.*, 2013; Wu *et al.*, 2015). Adaptive hypotheses for the existence of noncoding chromosomes suggest the presence of unidentified genes or regulatory regions that are functionally relevant and are maintained by natural selection. Alternatively, gene-lacking chromosomes may replicate and be perpetuated by genetic drift (or even selfishly) as a result of an insignificant fitness cost on the organism (Sloan *et al.*, 2012; Wu *et al.*, 2015). A population study of *S. noctiflora* revealed some differences in the presence/absence of chromosomes that lack any identifiable genes (Wu *et al.*, 2015). These findings suggest that at least some chromosomes are evolving under genetic drift (suffering stochastic loss), while others (those carrying essential genes) may be subjected to natural selection (Wu *et al.*, 2015). Analyses of other individuals of *L. mirabile* and/or other species of the genus are necessary to better understand the evolution of the mtDNA in this lineage.

Rampant mitochondrial HGT from host to parasite

Lophophytum mirabile harbors 35 intact protein transgenes that were acquired from its legume hosts in one or multiple HGT events. Foreign genes in *L. mirabile* were transferred as DNA, most likely as large fragments (or even as whole chromosomes or the entire genome), given the presence of foreign gene clusters, introns, and cytosine to uracil RNA editing sites. Parasitic plants form vascular connections with the host plant through a haustorium that enables regular transfer of water, nutrients, proteins, mRNAs, and pathogens (Westwood *et al.*, 2010; Kim *et al.*, 2014). The intimate association between plant holoparasites and their hosts could facilitate the occasional transfer of large fragments of DNA or whole mitochondria (Rice *et al.*, 2013). The existence of chromosomes in *L. mirabile* mtDNA carrying exclusively mitochondrial genes derived from the Fabaceae (e.g. C18, C28, C31, C34, C36, C37, C38, C39 and C47) opens the possibility that complete subgenomic circular molecules were acquired from their host by HGT. Hundreds of small circles have been theoretically predicted for *Glycine max* mtDNA through homologous recombination between large repeats (Chang *et al.*, 2013). Also, minicircles were observed in electron micrographs of soybean and pea mitochondrial DNA isolates (Synenki *et al.*, 1978). It is possible that circular molecules of legume mtDNA were horizontally transferred, giving rise to foreign autonomous chromosomes in *L. mirabile* mitochondria.

In addition to the mitochondrial transgenes, chloroplast and nuclear sequences were also transferred from the legume host into the *L. mirabile* mtDNA. The route by which these plastid and nuclear sequences entered the *L. mirabile* mitochondria remains unknown. However, it is more likely that plastid genes were first

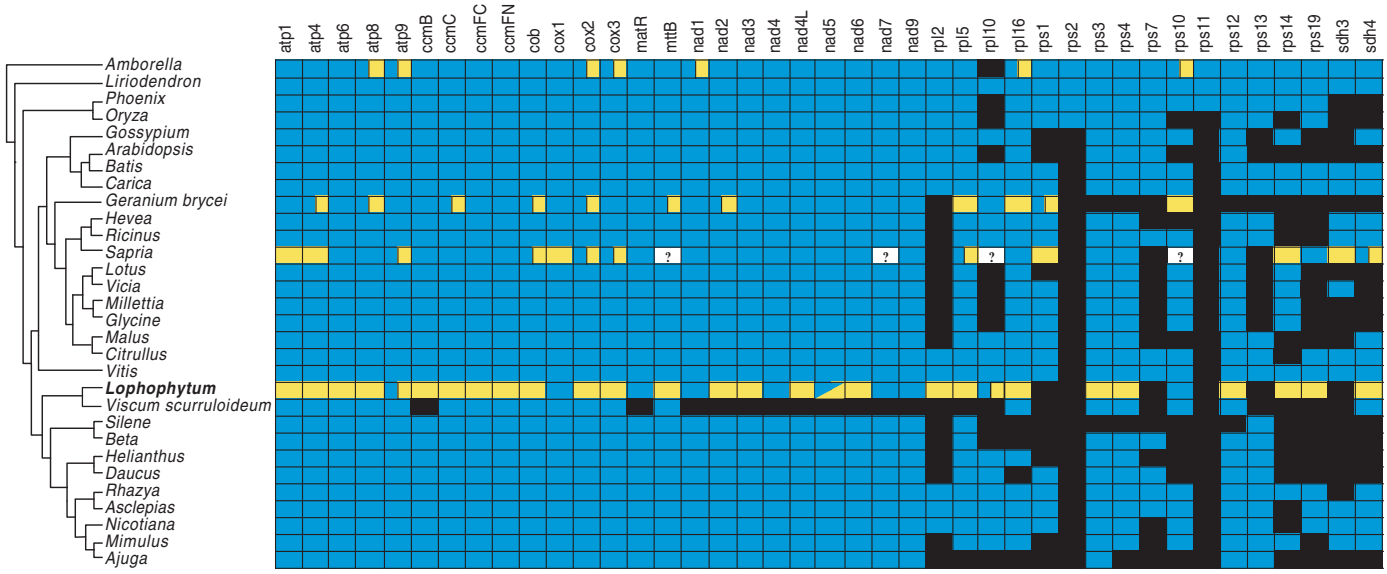

Fig. 3 Comparison of the mitochondrial protein-coding gene content in *Lophophytum mirabile* and selected angiosperms. Colors indicate native (blue), foreign (yellow), and absent (black) genes; a question mark is shown for undetermined status. The presence of both native and foreign copies of a gene is depicted by a subdivided rectangle. The gene *nad5* of *L. mirabile* (two triangles) is trans-spliced and consists of native and foreign individual gene regions. The tree shown is the best estimate of relationships among these plants. Only intact genes are shown; pseudogenes are not included.

transferred to the mtDNA by intracellular gene transfer (IGT) within the legume donor and then horizontally transferred to *L. mirabile* via mitochondrion-to-mitochondrion (mt-to-mt) HGT for the following reasons. First, mt-to-mt HGT is quite frequent among plant mitochondria (Rice *et al.*, 2013; Sanchez-Puerta, 2014). Second, plastid-to-mitochondrial IGT is fairly common in angiosperms (Kubo & Mikami, 2007). Third, plastid-to-plastid HGT is unknown in angiosperms and no clear cases of plastid-to-mitochondrial HGT exist (Sanchez-Puerta, 2014). Under the mt-to-mt HGT hypothesis, foreign plastid genes in *L. mirabile* mtDNA should be embedded within foreign mitochondrial tracts, which are, as of yet, unrecognizable, probably owing to the lack of sequencing information of mimosoid mitochondrial genomes.

Fate of foreign genes in *Lophophytum mirabile* mtDNA

Horizontal acquisition of mitochondrial genes results initially in gene duplication in the recipient mtDNA. Foreign genes commonly coexist with native homologs (duplicative HGT), as observed in the mtDNA of *Amborella* (Rice *et al.*, 2013) and *Geranium brycei* (Park *et al.*, 2015) (subdivided rectangles in Fig. 3). In those cases, transferred genes usually become pseudogenes while the native alleles remain functional (Mower *et al.*, 2010; Rice *et al.*, 2013). Occasionally, the foreign and native copies may undergo continuous or discontinuous gene conversion, generating chimeric gene copies (chimeric HGT), as described in earlier studies (Bergthorsson *et al.*, 2003; Barkman *et al.*, 2007; Hao *et al.*, 2010; Hepburn *et al.*, 2012; Park *et al.*, 2015). In other cases, the native copy is lost and only the foreign homolog remains (replacement HGT), as reported for a few transgenes in *S. hymalayana* (Xi *et al.*, 2013; Molina *et al.*, 2014) (yellow rectangles in Fig. 3). Another possibility is that native genes that were lost from the mtDNA (i.e. transferred to the nucleus) could be reacquired by HGT (recapture HGT) as reported for *Geranium brycei* (Park *et al.*, 2015), or the Rafflesiaceae (Xi *et al.*, 2013), and may or may not be functional in the recipient mitochondria.

All four scenarios are conceivable during the vast acquisition of foreign mitochondrial genes by the *L. mirabile* mtDNA. It is particularly striking that, with two exceptions (*atp9* and *rpl10*), all native homologs of the foreign genes are missing from the *L. mirabile* mtDNA. That is, there are two gene overlaps between the native and the foreign mitochondrial gene content (subdivided rectangles in Fig. 3). The distinct copies of the genes *atp9* and *rpl10*, at least one putatively native and another foreign, represent the only two cases of duplicative HGT in *L. mirabile*. However, *atp9* is the shortest protein-coding gene in the genome and therefore provides very little signal. As a result, there is no robust evidence that the putatively native copy of *atp9* is from a nonlegume lineage (Fig. S2).

The lack of comparative data from other members of the family Balanophoraceae makes it difficult to distinguish among the other three scenarios. Only two genera of the order Santalales have been examined so far. A Southern blot hybridization survey showed that *Lepionurus* (Opiliaceae, Santalales) conserves all mitochondrial genes present in the ancestral angiosperm, except for *rps2* and *rps11*, which were lost deeply in eudicot evolution (Adams *et al.*, 2002). By contrast, *Viscum* spp. (Viscaceae, Santalales) maintain only 14–25 genes after suffering massive gene loss (Petersen *et al.*, 2015; Skippington *et al.*, 2015).

Given that native homologs of genes otherwise universally present in angiosperms are absent from the *L. mirabile* mtDNA and have been replaced by foreign homologs (Fig. 3), we reason that the most widespread event in *L. mirabile* mtDNA was replacement HGT, with sporadic incidence of recapture HGT.

Estimating the frequency of chimeric HGT is particularly difficult based on the limited or lack of sequence data from close relatives to *L. mirabile*. A test to detect gene conversion in 13 *L. mirabile* mitochondrial genes (for which data on Santalales were available) was significant in only one case (*atp6*), suggesting a low frequency of chimerism. However, the Santalales available are only distant relatives to *L. mirabile* and could prevent the detection of additional events of recombination. Examining other species of *Lophophytum* or the sister genus *Ombrophytum* will allow to test these inferences.

Are foreign genes functional in *Lophophytum mirabile*?

One of the biggest questions in the field of plant HGT is whether foreign genes are functional in the recipient genome and what is the evolutionary impact of this phenomenon. Gene expression has been reported for some foreign genes in plant mitochondria (Rice *et al.*, 2013; Xi *et al.*, 2013). However, confirmation of gene function requires additional evidence beyond expression data, given that pseudogenes and noncoding regions can be transcribed in angiosperm mitochondria (Brandt *et al.*, 1993; Holec *et al.*, 2006). However, additional proof is more difficult to gather and is yet to be provided.

Here, we argue that foreign mitochondrial genes in *L. mirabile* mtDNA are most likely functional for several reasons: first, they have an intact ORF unlike foreign plastid genes in *L. mirabile* or the majority of the foreign genes found, for example, in *Amborella* (Rice *et al.*, 2013); second, one, and only one, full-length allele (except for *atp9* and *rpl10*), either native or foreign, is maintained in the *L. mirabile* mtDNA (Fig. 3), indicating that adaptive forces discriminated against mitochondria that carried two different alleles because their coexistence might have interfered with the overall gene expression or genetic homeostasis (Sanchez-Puerta *et al.*, 2015); native copies of genes that are universally present in angiosperm mtDNA (e.g. *atp1*, *atp6*, *ccmFN*, *cob*) are missing in *L. mirabile* and are probably essential for a functional mitochondria; thus, the foreign homologs could be functionally replacing them; a few transgenes persisted as intact genes through the speciation event that led to *L. mirabile* and *L. leandri*; and most foreign genes are evolving under purifying selection.

Although no direct evidence of functionality is provided for the foreign mitochondrial genes in *L. mirabile*, the claims listed earlier are compelling. Of all, the strongest proof is the presence of a full-length foreign copy of each of the missing native genes (but none of the rest, except for *atp9* and *rpl10*), suggesting that natural selection may be maintaining a complete set of functional genes, either native or foreign, in the *L. mirabile* mtDNA. The expected effect of genetic drift on nonfunctional gene sequences would be a random perpetuation of genes or gene fragments, leading to the arbitrary occurrence of foreign alleles with variable degrees of decay and redundancy with native genes. If the foreign homologs are, indeed, active in *L. mirabile*, the stunning functional replacement of up to 26 genes in *L. mirabile* (Fig. 3) represents an extraordinary example of the actual effect that rampant HGT can exert on plant mitochondria.

To explore why foreign gene copies are replacing the native homologs, we looked for evidence of relaxation of selection on native vs foreign homologs that would give foreign genes a competitive advantage. According to our results, native genes did not show evidence of diversifying selection, or evolutionary patterns or rate of synonymous substitutions different from foreign homologs, or from those in free-living angiosperms. Thus, the reasons for the functional replacement of native genes by host homologs remain elusive.

Conclusions

The mtDNA of *L. mirabile* revealed several remarkable features, the most extraordinary of which is the presence of 80% of the protein genes acquired from its host, while the native homologs were lost. The unprecedented replacement of native genes by likely functional foreign homologs raises new questions regarding the consequences of HGT on the host–parasite relationship and on the coevolution of organellar and nuclear genomes. Considering the extensive functional replacement of foreign mitochondrial genes in *L. mirabile* mtDNA, the magnitude and effect of HGT in the nuclear genome of this holoparasite are intriguing. Further analysis of parasitic species should contribute to uncover evolutionary aspects of acquiring foreign genes and address the effect of a hemiparasitic or a holoparasitic lifestyle on the proclivity for HGT.

Acknowledgements

We thank H. A. Sato for collecting samples of parasites and hosts, T. Valdovinos for conducting preliminary analyses and C. F. Delwiche, M. Rivarola, and A. Edera for critical reading of the manuscript. This work was supported by ANPCyT (PICT1762) and UNCUYO (06/M075) through funding to M.V.S-P.

Author contributions

M.V.S-P. designed the study. J.W. and L.F.C. obtained the DNA sequences generated for the study and analyzed them. M.V.S-P. and L.E.G. performed the evolutionary analyses and wrote the manuscript.

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, USA* **99**: 9905–9912.

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. 1990. Basic local alignment search tool. *Journal of Molecular Biology* **215**: 403–410.

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**: 2499–2513.

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**: 1436–1448.

Barkman TJ, McNeal JR, Lim SH, 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**: 248.

Bergthorsson U, Adams KL, Thomason B, Palmer JD. 2003. Widespread horizontal transfer of mitochondrial genes in flowering plants. *Nature* **424**: 197–201.

Bergthorsson U, Richardson A, Young G, Goertzen L, 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, USA* **101**: 17747–17752.

Brandt P, Unseld M, Eckertossenkopp U, Brennicke A. 1993. An *rps14* pseudogene is transcribed and edited in *Arabidopsis* mitochondria. *Current Genetics* **24**: 330–336.

Burger G, Gray MW, Lang BF. 2003. Mitochondrial genomes: anything goes. *Trends in Genetics* **19**: 709–716.

Chang S, Wang Y, Lu J, Gai J, Li J, Chu P, Guan R, Zhao T. 2013. The mitochondrial genome of soybean reveals complex genome structures and gene evolution at intercellular and phylogenetic levels. *PLoS ONE* **8**: e56502.

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, USA* **95**: 14244–14249.

Davis CC, Xi Z. 2015. Horizontal gene transfer in parasitic plants. *Current Opinion in Plant Biology* **26**: 14–19.

Doyle JJ, Doyle JL. 1987. A rapid isolation procedure for small quantities of fresh leaf tissues. *Phytochemical Bulletin* **19**: 11–15.

Gonzalez A, Mauseth J. 2010. Morphogenesis is highly aberrant in the vegetative body of the holoparasite *Lophophytum leandrii* (Balanophoraceae): all typical vegetative organs are absent and many tissues are highly modified. *International Journal of Plant Sciences* **171**: 499–508.

Gordon D, Green P. 2013. Consed: a graphical editor for next-generation sequencing. *Bioinformatics* **29**: 2936–2937.

Hansen B. 1980. Balanophoraceae. *Flora Neotrópica* **23**: 1–80.

Hao W. 2010. OrgConv: detection of gene conversion using consensus sequences and its application in plant mitochondrial and chloroplast homologs. *BMC Bioinformatics* **11**: 114.

Hao W, Richardson AO, Zheng Y, Palmer JD. 2010. Gorgeous mosaic of mitochondrial genes created by horizontal transfer and gene conversion. *Proceedings of the National Academy of Sciences, USA* **14**: 21576–21581.

Heide-Jorgensen HS. 2008. *Parasitic flowering plants*. Leiden, the Netherlands: Koninklijke Brill.

Hepburn N, Schmidt D, Mower JP. 2012. Loss of two introns from the *Magnolia tripetala* mitochondrial *cox2* gene implicates horizontal gene transfer and gene conversion as a novel mechanism of intron loss. *Molecular Biology and Evolution* **29**: 311–320.

Holec S, Lange H, Kuhn K, Alioua M, Borner T, Gagliardi D. 2006. Relaxed transcription in *Arabidopsis* mitochondria is counterbalanced by RNA stability control mediated by polyadenylation and polynucleotide phosphorylase. *Molecular and Cellular Biology* **26**: 2869–2876.

Keeling PJ, Palmer JD. 2008. Horizontal gene transfer in eukaryotic evolution. *Nature Reviews Genetics* **9**: 1–14.

Kim G, LeBlanc ML, Wafula E, dePamphilis CW, Westwood JH. 2014. Genomic-scale exchange of mRNA between a parasitic plant and its hosts. *Science* **345**: 808–811.

Kimura M. 1983. *The neutral theory of molecular evolution*. Cambridge, UK: Cambridge University Press.

Kubo T, Mikami T. 2007. Organization and variation of angiosperm mitochondrial genome. *Physiologia Plantarum* **129**: 6–13.

Lohse M, Drechsel O, Bock R. 2007. OrganellarGenomeDRAW (OGDRAW): a tool for the easy generation of high-quality custom graphical maps of plastid and mitochondrial genomes. *Current Genetics* **52**: 267–274.

Lowe TM, Eddy SR. 1997. tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. *Nucleic Acids Research* **25**: 955–964.

Maddison W, Maddison P. 2000. *MacClade version 4: analysis of phylogeny and character evolution*. Sunderland, MA, USA: Sinauer Associates.

Messier W, Stewart C-B. 1997. Episodic adaptive evolution of primate lysozymes. *Nature* **385**: 151–154.

Molina J, Hazzouri KM, Nickrent DL, Geisler M, Meyer RS, Pentony MM, Flowers JM, Pelser P, Barcelona J, Inovejas SA *et al.* 2014. Possible loss of the chloroplast genome in the parasitic flowering plant *Rafflesia lagascae* (Rafflesiaceae). *Molecular Biology and Evolution* **31**: 793–803.

Mower JP. 2005. PREP-Mt: predictive RNA editor for plant mitochondrial genes. *BMC Bioinformatics* **6**: 96.

Mower JP, Stefanovic S, Hao W, Gummow JS, Jain K, Ahmed D, Palmer JD. 2010. Horizontal acquisition of multiple mitochondrial genes from a parasitic plant followed by gene conversion with host mitochondrial genes. *BMC Biology* **8**: 150.

Park S, Grewe F, Zhu A, Ruhlman T, Sabir J, Mower JP, Jansen RK. 2015. Dynamic evolution of *Geranium* mitochondrial genomes through multiple horizontal and intracellular gene transfers. *New Phytologist* **208**: 570–583.

Petersen G, Cuenca A, Moller I, Seberg O. 2015. Massive gene loss in mistletoe (*Viscum*, Vicaceae) mitochondria. *Scientific Reports* **5**: 17588.

Rice DW, Alverson AJ, Richardson AO, Young GJ, Sanchez-Puerta MV, Munzinger J, Barry K, Boore JL, Zhang Y, dePamphilis CW *et al.* 2013. Horizontal transfer of entire genomes via mitochondrial fusion in the angiosperm *Amborella*. *Science* **342**: 1468–1473.

Sanchez-Puerta MV. 2014. Involvement of plastid, mitochondrial and nuclear genomes in plant-to-plant horizontal gene transfer. *Acta Societatis Botanicorum Poloniae* **83**: 317–323.

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**: 1762–1777.

Sanchez-Puerta MV, Zubko MK, Palmer JD. 2015. Homologous recombination and retention of a single form of most genes shape the highly chimeric mitochondrial genome of a cybrid plant. *New Phytologist* **206**: 381–396.

Sato H. 2014. *Anatomía reproductiva de las especies de Lophophytum Schott & Endl. (Balanophoraceae) de la Argentina y revisión taxonómica del género en América*. PhD thesis, Facultad de Ciencias Agrarias, Universidad Nacional del Nordeste, Corrientes, Argentina.

Sato H, Gonzalez A. 2013. Anatomía y desarrollo de la flor estaminada, microsporogénesis y microgametogénesis en especies de *Lophophytum* (Balanophoraceae) en la Argentina. *Boletín de la Sociedad Argentina de Botánica* **48**: 59–72.

Sato HA, Gonzalez AM. 2016. Floral development and anatomy of pistillate flowers of *Lophophytum* (Balanophoraceae), with special reference to the embryo sac inversion. *Flora* **219**: 35–47.

Shimodaira H. 2002. An approximately unbiased test of phylogenetic tree selection. *Systematic Biology* **51**: 492–508.

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, USA* **112**: E3515–E3524.

Sloan D, Alverson AJ, Chuckalovcak J, Wu M, McCauley D, Palmer JD, Taylor DR. 2012. Rapid evolution of enormous, multichromosomal genomes in flowering plant mitochondria with exceptionally high mutation rates. *PLoS Biology* **10**: e1001241.

Stamatakis A. 2014. RAxML Version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* **30**: 1312–1313.

Straub SCK, Parks M, Weitemier K, Fishbein M, Cronn R, Liston A. 2012. Navigating the tip of the genomic iceberg: next-generation sequencing for plant systematics. *American Journal of Botany* **99**: 349–364.

Su H-J, Hu J-M, Anderson FE, Der JP, Nickrent DL. 2015. Phylogenetic relationships of Santalales with insights into the origins of holoparasitic Balanophoraceae. *Taxon* **64**: 491–506.

Swofford D, Olsen G, Waddell P, Hillis D. 2002. *PAUP*. Phylogenetic Analysis Using Parsimony (*and other methods). Version 4*. Sunderland, MA, USA: Sinauer Associates.

Synenki R, Levings C, Shah D. 1978. Physicochemical characterization of mitochondrial DNA from soybean. *Plant Physiology* **61**: 460–464.

The Angiosperm Phylogeny Group. 2016. An update of the Angiosperm Phylogeny Group classification for the orders and families of flowering plants: APG IV. *Botanical Journal of the Linnean Society* **181**: 1–20.

Westwood JH, Yoder JA, Timko M, dePamphilis CW. 2010. The evolution of parasitism in plants. *Trends in Plant Science* **15**: 227–235.

Wicke S, Muller KF, de Pamphilis CW, Quandt D, Wickett N, Zhang Y, Renner SS, Schneeweiss GM. 2013. Mechanisms of functional and physical genome reduction in photosynthetic and nonphotosynthetic parasitic plants of the broomrape family. *Plant Cell* **25**: 3711–3725.

Wu Z, Cuthbert JM, Taylor DR, Sloan DB. 2015. The massive mitochondrial genome of the angiosperm *Silene noctiflora* is evolving by gain or loss of entire chromosomes. *Proceedings of the National Academy of Sciences, USA* **112**: 10185–10191.

Xi Z, Bradley R, Wurdack K, Wong K, Sugumaran M, Bomblies K, Rest J, Davis CC. 2012. Horizontal transfer of expressed genes in a parasitic flowering plant. *BMC Genomics* **13**: 227.

Xi Z, Wang Y, Bradley R, Sugumaran M, Marx C, Rest J, Davis CC. 2013. Massive mitochondrial gene transfer in a parasitic flowering plant clade. *PLoS Genetics* **9**: e1003265.

Yang Z. 2007. PAML 4: Phylogenetic Analysis by Maximum Likelihood. *Molecular Biology and Evolution* **24**: 1586–1591.

Yang Z, Wong WSW, Nielsen R. 2005. Bayes empirical Bayes inference of amino acid sites under positive selection. *Molecular Biology and Evolution* **22**: 1107–1118.

Zerbino D, Birney E. 2008. Velvet: algorithms for *de novo* short read assembly using de Bruijn graphs. *Genome Research* **18**: 821–829.

Zwickl DJ. 2006. *Genetic algorithm approaches for the phylogenetic analysis of large biological sequence datasets under the maximum likelihood criterion.* PhD thesis, School of Biological Sciences, University of Texas, Austin, TX, USA.

Supporting Information

Additional Supporting Information may be found online in the Supporting Information tab for this article:

Fig. S1 Total read depth of the mitochondrial genome of *Lophophytum mirabile*.

Fig. S2 Phylogenetic analyses of mitochondrial genes in the mtDNA of *Lophophytum mirabile*.

Fig. S3 Approximately unbiased test results.

Fig. S4 Phylogenetic analyses of chloroplast-derived sequences in the mtDNA of *Lophophytum mirabile*.

Fig. S5 Phylogenetic analyses of mitochondrial pseudogenes in the mtDNA of *Lophophytum mirabile*.

Fig. S6 Phylogenetic analyses of nuclear sequences found in the mtDNA of *Lophophytum mirabile*.

Table S1 Accession numbers and references of the taxa included in the phylogenetic analyses

Table S2 Estimation of synonymous and nonsynonyomous substitution rates and detection of positive selection in *Lophophytum mirabile*

Please note: Wiley Blackwell are not responsible for the content or functionality of any Supporting Information supplied by the authors. Any queries (other than missing material) should be directed to the *New Phytologist* Central Office.