Plant Molecular Biology (2020) 103:623–638
https://doi.org/10.1007/s11103-020-01014-x

Multichromosomal structure and foreign tracts in the *Ombrophytum subterraneum* (Balanophoraceae) mitochondrial genome

M. Emilia Roulet[1] · Laura E. Garcia[1,2] · Carolina L. Gandini[1] · Hector Sato[3] · Gabriela Ponce[1,2] · M. Virginia Sanchez-Puerta[1,2]

Received: 20 December 2019 / Accepted: 5 May 2020 / Published online: 22 May 2020

Abstract

Horizontal gene transfer (HGT) is frequent in parasitic plant mitochondria as a result of vascular connections established in host-parasite relationships. Recent studies of the holoparasitic plant *Lophophytum mirabile* (Balanophoraceae) revealed the unprecedented acquisition of a large amount of mitochondrial sequences from its legume host. We focused on a close relative, the generalist holoparasite *Ombrophytum subterraneum*, to examine the incidence of HGT events in the mitochondrial genome (mtDNA). The mtDNA of *O. subterraneum* assembles into 54 circular chromosomes, only 34 of which contain the 51 full-length coding regions. Numerous foreign tracts (totaling almost 100 kb, ~ 14% of the mtDNA), including 12 intact genes, were acquired by HGT from the Asteraceae hosts. Nine chromosomes concentrate most of those regions and eight are almost entirely foreign. Native homologs of each foreign gene coexist in the mtDNA and are potentially functional. A large proportion of shorter regions were related to the Fabaceae (a total of ~ 110 kb, 15.4%), some of which were shared with *L. mirabile*. We also found evidence of foreign sequences donated by angiosperm lineages not reported as hosts (Apocynaceae, Euphorbiaceae, Lamiaceae, and Malvales). We propose an evolutionary hypothesis that involves ancient transfers from legume hosts in the common ancestor of *Ombrophytum* and *Lophophytum* followed by more recent transfer events in *L. mirabile*. Besides, the *O. subterraneum* mtDNA was also subjected to additional HGT events from diverse angiosperm lineages, including large and recent transfers from the Asteraceae, and also from Lamiaceae.

Key message

The *O. subterraneum* mtDNA harbors foreign mitochondrial regions from diverse angiosperms, including 14 and 15% acquired from host lineages Asteraceae and Fabaceae. Twelve intact foreign genes coexist with native homologs.

Keywords Horizontal gene transfer · Santalales · *Lophophytum* · Holoparasite · mtDNA · Evolution

Introduction

Most angiosperm lineages include species that evolved independently in a successful heterotroph life strategy (Barkman et al. 2007; Westwood et al. 2010). Parasitic plants are characterized by their ability to feed directly from other plants, invading roots or stems of their host through a specialized structure called the haustorium (Davis and Xi 2015; Westwood et al. 2010). This vascular connection between host and parasite allows the passage of water, nutrients, pathogens, and even nucleic acids (Barkman et al. 2007; Davis and Wurdack 2004; Kim et al. 2014; Mower et al. 2004; Nickrent et al. 2004; Roney et al. 2007; Sanchez-Puerta 2014; Westwood et al. 2009; Yang et al. 2016). Consequently, parasitic relationships facilitate the exchange of genetic information,

---

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s11103-020-01014-x) contains supplementary material, which is available to authorized users.

✉ M. Virginia Sanchez-Puerta
mvsanchezpuerta@fca.uncu.edu.ar

[1] IBAM, Universidad Nacional de Cuyo, CONICET, Facultad de Ciencias Agrarias, Almirante Brown 500, M5528AHB Chacras de Coria, Mendoza, Argentina

[2] Facultad de Ciencias Exactas y Naturales, Universidad Nacional de Cuyo, Padre Jorge Contreras 1300, M5502JMA Mendoza, Argentina

[3] Facultad de Ciencias Agrarias, Universidad Nacional de Jujuy, Cátedra de Botánica General–Herbario JUA, Alberdi 47, 4600 San Salvador de Jujuy, Jujuy, Argentina

a phenomenon known as horizontal gene transfer (Davis and Xi 2015; Kim et al. 2014; Mower et al. 2010; Sanchez-Puerta 2014; Westwood et al. 2010; Xi et al. 2013).

Horizontal gene transfer (HGT) refers to the movement of genetic material between non-mating organisms. Most cases of plant-to-plant HGT involve the mitochondria, others affect the nucleus, and only rarely the chloroplast (Gandini and Sanchez-Puerta 2017; Keeling and Palmer 2008; Sanchez-Puerta 2014). This frequency of HGT in angiosperm mitochondria could be explained by properties of plants mtDNAs, such as the presence of large intergenic regions, which enables the acquisition of foreign DNA without interrupting the function of native genes (Davis and Xi 2015; Keeling and Palmer 2008; Sanchez-Puerta 2014); successive fusion and fission events of seed plant mitochondria during the cell cycle (Arimura et al. 2004), and frequent homologous recombination of plant mitochondrial genomes (Manchekar et al. 2006; Sanchez-Puerta 2014; Shedge et al. 2010).

A "mitochondrial fusion compatibility model" has been proposed for plant mitochondrial HGT (Rice et al. 2013). This model states that HGT involves the capture of entire foreign mitochondria from the donor plant followed by fusion with the native organelles (Keeling and Palmer 2008; Rice et al. 2013; Sanchez-Puerta 2014). Then, both mitochondrial genomes (mtDNAs) recombine, and foreign mitochondrial fragments may be introduced in the mtDNA of the recipient plant (Sanchez-Puerta 2014). Besides, plant mtDNAs can also import sequences from other cellular compartments, such as nuclear and plastid fragments acquired by intracellular gene transfer (IGT) (Kubo and Mikami 2007). Under this model, plant mitochondrial HGT initiates with the transfer of mtDNA resulting in the initial duplication of the mitochondrial gene content (Sanchez-Puerta 2014). In many cases, foreign genes are non-functional in the recipient genome, they become pseudogenes over time and are likely lost (*silent* HGT) (Bergthorsson et al. 2004; Mower et al. 2010; Park et al. 2015; Rice et al. 2013; Sanchez-Puerta et al. 2017). In other cases, both copies in the recipient mtDNA are maintained as full-length genes (*duplicative* HGT) (Bergthorsson et al. 2004; Rice et al. 2013; Sanchez-Puerta et al. 2017; Xi et al. 2013); or native and foreign alleles recombine generating chimeric genes (*chimeric* HGT) (Barkman et al. 2007; Bergthorsson et al. 2003; Hao et al. 2010; Hepburn et al. 2012; Mower et al. 2010; Park et al. 2015; Sanchez-Puerta et al. 2019, 2017; Xi et al. 2013). Another possible result is the replacement of a native copy by a foreign one (*replacement HGT*) (Bellot et al. 2016; Sanchez-Puerta et al. 2017; Xi et al. 2013).

There are a few examples of angiosperms rich in mitochondrial HGT, including the parasitic plants *Lophophytum mirabile* of the family Balanophoraceae (Sanchez-Puerta et al. 2017), *Rafflesia cantleyi, R. tuan-mudae,* and *Sapria himalayana* of the family Rafflesiaceae (Molina et al. 2014; Xi et al. 2013); and two free-living species, *Amborella trichopoda* (Rice et al. 2013) and *Geranium brycei* (Park et al. 2015). A recent study of the holoparasitic plant *L. mirabile* revealed the unprecedented acquisition of about 60% of the mtDNA (intergenic and coding regions) from its legume host and the replacement of the majority of the native genes by foreign alleles (Sanchez-Puerta et al. 2017, 2019). These findings raise new questions regarding the consequences of HGT in the host-parasite relationship. Parasitic plants are particularly susceptible to this phenomenon and represent an excellent system to examine the role of HGT in the evolution of the mitochondrial genome.

We focus on the family Balanophoraceae sensu stricto (Santalales), which encompasses 14 genera of obligate root parasites distributed across tropical and subtropical regions (Su et al. 2015). This family is characterized by species with an irregular subterranean mass (vegetative body or 'tuber') that grows attached to the host roots (Hansen 1980). The genus *Ombrophytum* belongs to this family and is sister to *Lophophytum* (Su et al. 2015). It includes seven species: *O. subterraneum* (Asplund) B. Hansen, *O. violaceum* B. Hansen, *O. microlepis* B. Hansen, *O. peruvianum* Poepp. & Endl., *O. guayanensis* P.G. Delprete, *O. chilensis* Kuijt & Delprete, and *O. villamariensis* Guzmán-Guzmán (Delprete 2014; Guzmán-Guzmán 2019; Hansen 1980; Kuijt and Delprete 2019; Sato 2014; Zuloaga et al. 2008). *Ombrophytum subterraneum* is a holoparasitic plant, distributed in South America (Argentina, Bolivia, Chile, and Ecuador), growing in dry bushland and occasionally in tropical and subtropical areas (Delprete 2014; Hansen 1980; Mauseth et al. 1992; Sato 2014; Su et al. 2015). This species is a generalist parasite and it has been found associated with diverse hosts. Reported hosts belong mainly to the angiosperm family Asteraceae (*Baccharis grisebachii*, *B. petiolata*, *Eupatorium bupleurifolium*, *Heterothalamus spartioides*, *Lepidophyllum quadrangulare*, *Scalesia peduculata*, *Tessaria absinthioides*, *Viguieria mollis*), but also to the families Dioscoreaceae (*Dioscorea megalantha*), Fabaceae (*Medicago sativa*), and Solanaceae (*Nicotiana glauca*) (Hansen 1980; Mauseth et al. 1992). In Argentina, *Ombrophytum subterraneum* has been found in Catamarca, Jujuy, Salta, and Tucumán (Zuloaga et al. 2008), always parasitizing *Baccharis* sp. (Hector Sato unpublished results).

Fortunately, mitochondrial sequences of many of the reported host lineages are available. There are complete mtDNAs for seven genera of Asteraceae, more than twenty Fabaceae, several Solanaceae, and none of Dioscoreaceae. In this study, we performed comparative and evolutionary genomic analyses of the mitochondrial genome of the holoparasite *Ombrophytum subterraneum* to address the following questions: (i) is there evidence of HGT in the mtDNA of *O. subterraneum*?; (ii) which are the donor lineages of those foreign regions?; (iii) were foreign sequences transferred to the parasite via DNA or mRNA?; (iv) are there cases of silent, duplicative, chimeric, or replacement HGT?; (v) what fraction of the *O. subterraneum* mtDNA is derived from the Asteraceae host?; and (vi) is there evidence of HGT events from Fabaceae that occurred in the ancestor of *Lophophytum* and *Ombrophytum*?

Materials and methods

DNA extraction and sequencing

Plants of the holoparasitic angiosperm *O. subterraneum* (Balanophoraceae) and its host plant *Baccharis* sp. (Asteraceae) were collected in May 16th, 2015 from the banks of the lagoon of Mr. Mamaní's estate (S22°55′02.4" W65°35′32.8" 3689 mamsl), Tres Cruces, Humahuaca, Jujuy, Argentina. Total DNA was extracted from leaves of individual plants of *Baccharis* sp. using a cetyl-trimethylammonium-bromide DNA-extraction protocol (Doyle and Doyle 1987), and from a section of a tuber of *O. subterraneum* sufficiently distant from host tissue using the DNeasy Plant Mini kit (Qiagen).

DNA of *O. subterraneum* was used to construct two paired-end libraries that were sequenced at the Beijing Genomics Institute using the Illumina® HiSeq™ sequencing technology. This generated two sets of reads: (i) 44 million clean 150 bp paired-end reads with an insert size of 250 bp (dataset250) with Illumina® HiSeq™ 4000 and (ii) 64 million clean 100 bp paired-end reads with an insert size of 800 bp (dataset800) with Illumina® HiSeq™ 2500. Both sets of reads were assembled independently on the Mason large-memory computer cluster at Indiana University-Bloomington (USA).

To discard contamination from the host, we designed primers in Primer3 (Koressaar and Remm 2007; Untergasser et al. 2012) to amplify in the host plant *Baccharis* sp. those genes identified as foreign in the holoparasite. PCR products were sequenced using an ABI 3730 DNA Analyzer (Applied Biosystems, Waltham, MA, USA).

Assembly of the mitochondrial genome

Assembly parameters were optimized with the program Velvet Optimiser 2.2 (Simon Gladman, CSIRO and Monash University, Australia). Then, we used Velvet 1.2.03 (Zerbino and Birney 2008) without scaffolding and with a hash length of 105 and 73 for the dataset250 and dataset800, respectively, to perform de novo assembly of the mtDNA of *O. subterraneum*. Using the dataset250, Velvet assembled 328 contigs larger than 1 kb, with N50 10,446 bp and a maximum contig size of 24,613 bp. With the dataset800, Velvet assembled 375 contigs larger than 1 kb, with N50 5,687 bp and a maximum contig size of 23,159 bp. Raw sequence data are available from the NCBI BioProject PRJNA596804. The annotated *O. subterraneum* mitochondrial chromosomes were deposited in the GenBank data libraries under accession numbers MT076267-MT076320 and genes sequences of *Baccharis* sp. under accession numbers MT036989-MT036999.

For each data set, the Illumina paired-end reads were aligned on the Velvet contigs to calculate the read depth. Using the differences in read depth among cellular compartments (Straub et al. 2012) and BLASTn searches against mitochondrial databases, we identified putatively mitochondrial contigs. We generated a set of mitochondrial contigs combining both datasets and compared them with Local BLAST to discard duplicate contigs. Based on information of paired-end reads, contigs were joined, edited, and finished in Consed v.29 (Gordon and Green 2013). The structure of the mtDNA consisting of 54 circular chromosomes was supported by paired-end reads having one mate mapping to the start of a contig and the other to the end. The total read depth for each chromosome using both sets of reads was calculated using Bowtie2 (Langmead and Salzberg 2012) with the following presets: –end-to-end –very-fast, and plots were generated using R (Supporting Information Figure S1).

Mitochondrial genome annotation

The mtDNA of *O. subterraneum* was annotated using Mitofy (Alverson et al. 2010), BLAST (Camacho et al. 2009), and tRNAscan-SE algorithm (Lowe and Eddy 1997). To predict gene and exon boundaries for each gene, we used the information of those angiosperm mtDNAs that were annotated based on RNAseq data. Non-synonymous RNA editing sites were predicted with PREP-Mt (Mower 2005) with a cutoff value of 0.5. Plastid-derived mitochondrial sequences (MTPTs) larger than > 100 bp were identified by BLAST searches against a custom angiosperm chloroplast database. Pseudogenes larger than > 100 bp were annotated and further analyzed. For those noncoding chromosomes in *O. subterraneum,* unknown ORFs with a minimum length of 300 nucleotides were identified using ORFfinder.

Repeats were identified in Consed v.29 (Gordon and Green 2013) using crossmatch. Repeats larger than 500 bp with identity > 90% were evaluated for possible recombination across them based on information of paired-end reads on repeats boundaries. Repeats larger than the greater library size (~ 800 bp) and those with evidence of recombination were assembled into a single contig to minimize subcircular maps. Graphical genome maps were generated using OGDRAW software (Greiner et al. 2019).

Comparative and evolutionary analyses

To infer the origin of the intergenic regions, we performed BLAST searches of each of the *O. subterraneum* mitochondrial chromosomes against a local database of angiosperm mtDNA available at NCBI Organelle Genome Database using the BLASTn v.2.4.0 (Camacho et al. 2009). BLAST hits with an *e* value $< 2 \times 10^{-10}$ were plotted using the Sushi R package v.1.20.0 (Phanstiel 2019) (Supporting Information Figure S2). Given that most regions of similarity were shared with the families Asteraceae and Fabaceae, BLAST hits to the available Asteraceae (*Artemisia annua*, *Chrysanthemum boreale*, *C. indicum*, *Diplostephium hartwegii*, *Dendrosenecio brassiciformis*, *Helianthus annuus*, *Lactuca sativa*, *L. serriola*, and *L. saligna*) and Fabaceae (*Acacia ligulata*, *Ammopiptanthus mongolicus*, *Cercis canadensis*, *Glycine max*, *G. soja, Haematoxylum brasiletto*, *Leucaena trichandra, Libidibia coriaria*, *Lotus japonicus*, *Medicago truncatula*, *Millettia pinnata*, *Phaseolus vulgaris*, *Senna tora*, *S. occidentalis, Sophora flavescens*, *Styphnolobium japonicum*, *Tamarindus indica*, *Vicia faba*, *Vigna angularis*, and *V. radiata*) mtDNAs were grouped for better visualization (Supporting Information Figure S2).

We considered a foreign region, as a result of HGT from the host lineages (Asteraceae or Fabaceae) when any of the following criteria were met: (i) all the hits of a particular region belong exclusively to the host family; (ii) hits to the host lineage were longer and with higher identity than those from other angiosperms, or (iii) ML phylogenetic analyses showed a strong relationship (bootstrap support > 70%) between *Ombrophytum* and one of the host families (Supporting Information Figure S3).

Phylogenetic analyses

Phylogenetic analyses were performed under Maximum Likelihood (ML) and Maximum Parsimony (MP) for each mitochondrial protein or rRNA coding gene. The nucleotide alignments include 43 free-living angiosperms, *Baccharis* sp. and additional species of Asterales, Santalales, or Fabales, when available (Supporting Information Table S1). Nucleotide sequences were aligned manually with AliView v.1.26 (Larsson 2014). RNA editing sites predicted for *O. subterraneum* with PREP-Mt and those identified in other angiosperms (Edera et al. 2018) were removed from the alignments for phylogenetic inferences. ML analyses were performed with RAxML v.8.2.11 (Stamatakis 2014) under the GTR + gamma model. A thousand rapid bootstrap replicates were done under the same model of evolution using RAxML. MP analysis were run in PAUP v.4.0 with 100 bootstrap replicates (Swofford et al. 2002). The trees were visualized with FigTree v.1.4.4 (https://tree.bio.ed.ac.uk/software/figtree/).

The approximately unbiased (AU) test implemented in the CONSEL package (Shimodaira 2002) was used to evaluate alternative topologies for each foreign gene. For those genes that have two foreign copies, we evaluated them individually, excluding the other foreign copy. Different constrained topologies that excluded the foreign copy from the asterids or that forced the foreign copy to be within the Santalales (as expected under vertical inheritance) were compared to the unconstrained tree. The ML tree for each constraint was found by searching for the best tree compatible with that constraint using PAUP v.4.0 (Swofford et al. 2002). The site likelihoods for the unconstrained and constrained trees were obtained with PAUP v.4.0, and the AU *p*-values were calculated from these data with CONSEL.

We used the program Geneconv v.1.81a (Sawyer 1989) to evaluate the presence of chimeric genes in the mtDNA of *O. subterraneum* and to re-evaluate chimerism in the *L. mirabile* foreign mitochondrial genes.

Analysis of host contamination in the assembly of *Ombrophytum subterraneum*

We dismissed the possibility of contamination in the mtDNA of *O. subterraneum* for several reasons. First, sample preparation of the parasite was done with extreme care, using a section of the tuber distal to the area of attachment to the host. Second, earlier studies found that the tuber tissue of *O. subterraneum* consists exclusively of parasite cells (Mauseth et al. 1992). Third, initial exploratory PCR amplifications and sequencing of 12 genes (*atp1*, *atp4*, *atp8*, *ccmC*, *ccmFN*, *cob*, *cox1*, *matR*, *nad6*, *nad9*, *rpl5*, *rps4*) in different samples of *O. subterraneum* revealed the presence of double peaks in the chromatograms of seven of those genes (*atp4*, *ccmC*, *ccmFN*, *matR*, *nad6*, *nad9*, *rps4*), indicating the likely presence of multiple copies. Analyses of such peaks agree with the presence of the foreign and native alleles in the mtDNA of *O. subterraneum.* Fourth, PCR products of seven genes (*atp4*, *ccmC*, *ccmFN*, *matR*, *nad6*, *nad9*, and *rps4*), identified as horizontally transferred from Asteraceae in the mtDNA of *O. subterraneum,* were amplified using DNA extracted from the actual host individual of *Baccharis* sp. In all cases, *O. subterraneum* sequences presented differences with those of *Baccharis* sp. Fifth, phylogenetic gene trees found *O. subterraneum* as sister to the family Asteraceae, and not to *Baccharis* sp. suggesting that *Baccharis* sp. is not the donor of those sequences. Sixth, all mitochondrial chromosomes show a similar read depth, including those regions identified as foreign, with a single foreign chromosome, OsChr46 exhibiting a lower average read depth (Supporting Information Figure S1 and Table S2). Taking all these into account, we can rule out the possibility of contamination from the host. In addition, the relative even read depth within and among chromosomes indicate that foreign sequences reside in the same cellular compartment.

Results

General features of the *Ombrophytum subterraneum* mitochondrial genome

The mtDNA of *O. subterraneum* is composed of 54 circular chromosomes (4–27 kb in length) totaling 713,777 bp (Fig. 1). This multichromosomal architecture was recently described in a close relative, *Lophophytum mirabile,* with a mtDNA of 821,919 bp that also assembled into 54 circular-mapping chromosomes (Sanchez-Puerta et al. 2017). Despite the identical number of chromosomes, only 19.19% (136,966 bp) of the *O. subterraneum* mtDNA shows similarity to the *L. mirabile* mtDNA, with an average identity of 85.88%. About half of such regions are coding sequences. The homologous sequences are highly fragmented and sporadically distributed across 52 chromosomes of *O. subterraneum* mtDNA (Supporting Information Figure S2). The largest homologous region is 4,556 bp long (this region includes the genes *cox1* and *rps10*, which are native in both species).

The GC content of the mtDNA of *O. subterraneum* is 44.22% similar to other flowering plants. Direct and inverted repeats represent 2% of the genome (11,956 bp). Two repeats larger than the greater library size (~ 800 bp) were identified and, assuming they actively recombine, they were assembled into a single contig (Fig. 1, underlined). The rest of the repeats were smaller than the library size and we did not find evidence of recombination based on paired-end read information. The final assembly of *O. subterraneum* has an average read depth of ~ 71×, ranging between 40 × and 113 × across the 54 mitochondrial chromosomes (Supporting Information Table S2). The genome shows a continuous and relatively even read depth with no gaps or low-coverage regions (Supporting Information Figure S1).

Only 34 chromosomes contain full-length mitochondrial genes, while the remaining 20 are devoid of any identifiable coding region. Nevertheless, a total of 26 unknown ORFs were identified in 15 of the 20 noncoding chromosomes; six of those ORFs are also found in other angiosperm species (Supporting Information Table S3). The coding fraction of the mtDNA represents 10.16% including 3 ribosomal RNA, 13 transfer RNA, and 36 protein coding genes. The majority of these genes are present in multiple copies, resulting in a total of 71 genes (Table 1). Three of the tRNAs (*trnD, trnS, trnW*) are of plastid origin (Supporting Information Table S4). The *O. subterraneum* genome contains 18 cis-splicing (*ccmFC, cox2, nad1, nad2, nad4, nad5, nad7, rpl2, rps3, rps10*) and six trans-splicing (*nad1, nad2, nad5*) group II introns and only one group I intron in the gene *cox1*. We identified nine mitochondrial pseudogenes that were truncated or frameshifted and 17 chloroplast-derived sequences representing 0.95% of the genome (Table 1 & Fig. 1). Non-synonymous RNA editing sites were predicted for all full-length protein-coding genes (Table 2).

Large horizontal gene transfers from Asteraceae hosts to parasite

To unveil the impact of horizontal gene transfer from the diverse hosts (or other donors) on the mtDNA of *O. subterraneum*, we performed similarity searches and phylogenetic analyses of genic and intergenic regions. BLASTn searches of the *O. subterraneum* mitochondrial chromosomes against a local database of angiosperm mtDNAs found that ~ 73% of the mtDNA has similarity to other angiosperm mitochondrial sequences. BLAST hits were distributed sporadically along each chromosome (Supporting Information Figure S2). Then, we inferred the origin of the mitochondrial regions of *O. subterraneum* based on the length and sequence identity of BLAST hits and phylogenetic affiliations of those regions with multiple hits of similar length (Supporting Information Figure S2 and S3).

About 14% (98,069 bp) of the *O. subterraneum* mtDNA is related to the family Asteraceae, with an average identity of 96.6%. Less than 11% of those foreign tracts involve a genic region and the largest intergenic foreign tract is 6.7 kb long in the chromosome OsChr9 (Supporting Information Figure S2). A detailed analysis of each chromosome of *O. subterraneum* demonstrates that the amount of sequences with affinity to Asteraceae is unevenly distributed (Fig. 2). Nine chromosomes (OsChr9, OsChr12, OsChr16, OsChr25, OsChr36, OsChr38, OsChr39, OsChr46, OsChr49) concentrate most of those regions and are almost entirely foreign (Supporting Information Figure S2). The limited availability of complete mtDNAs from diverse Asteraceae (only nine species of the > 25,000 described in the family) probably precludes the identification of additional foreign tracts.

A sizeable proportion of foreign sequences were acquired from legume hosts

By analyzing the BLAST search results, we found that, besides Asteraceae, there is a great amount of hits with high identity to mtDNAs of the family Fabaceae. A detailed analysis of such regions revealed that 15.4% (109,821 bp)

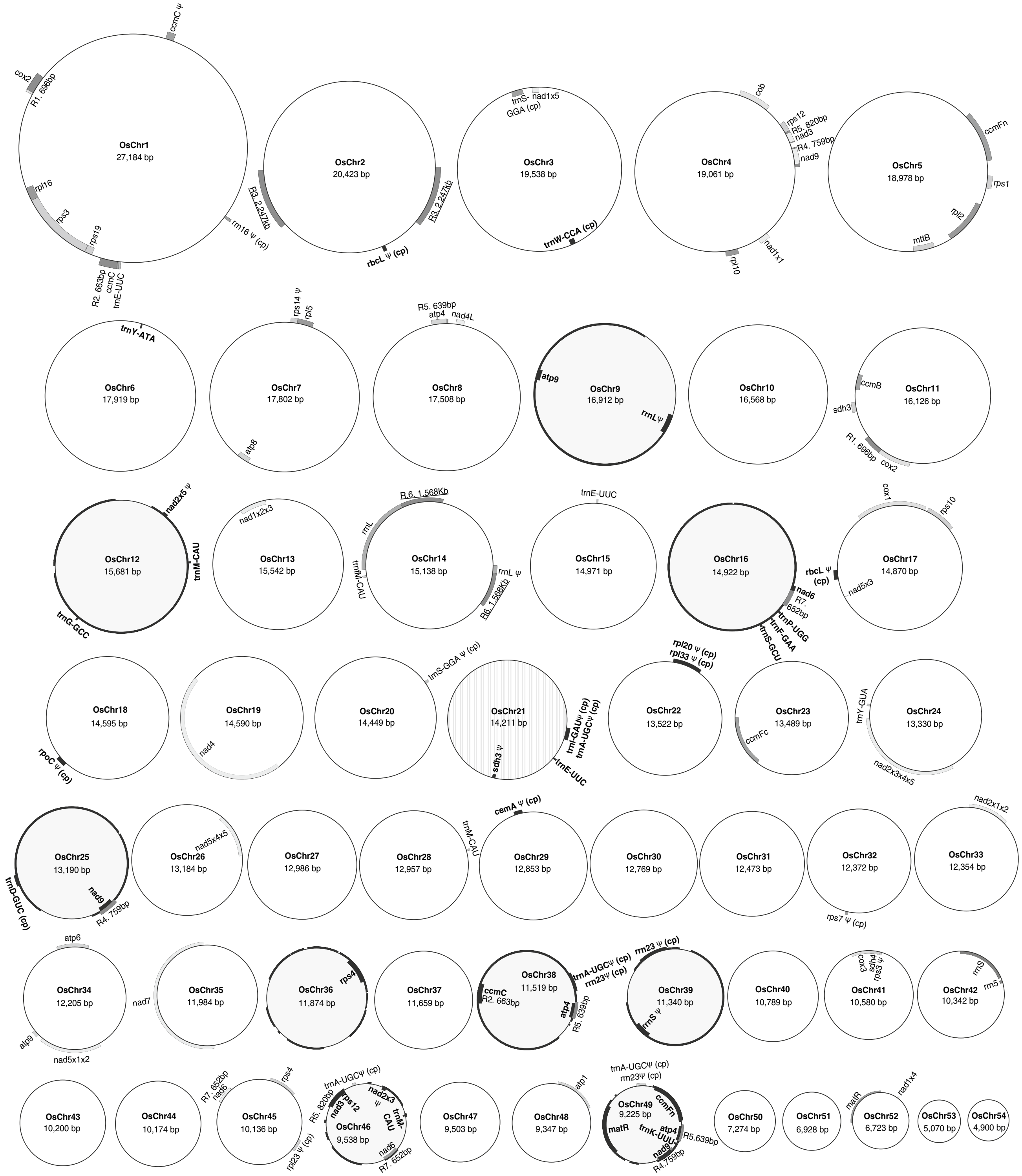

Fig. 1 Map of the mitochondrial genome (mtDNA) of *Ombrophytum subterraneum*. The mtDNA is 713,777 bp long and is subdivided in 54 circular chromosomes (OsChr) of diverse lengths. Genes drawn inside and outside each circle are transcribed clockwise and counterclockwise, respectively. Shown are full-length genes, pseudogenes > 100 bp (indicated by ‘Ψ’), repeats > 500 bp (labeled ‘R’, followed by the repeat length), and sequences derived from the chloroplast (cp) longer than 100 bp. Repeats assumed to recombine are underlined. Gene names in black or gray font are foreign or native, respectively. Thicker lines depict foreign intergenic regions related to Asteraceae. Solid and stripe circles represent chromosomes with a high proportion of sequences related to Asteraceae or Lamiaceae, respectively

Table 1 General features of the mitochondrial genome of *Ombrophytum subterraneum*

Genome length (bp)	713,777
GC content	44.22%
Protein-coding genes (including repeats)	36 (51)
rRNA-coding genes	3
tRNA-coding genes (including repeats)	12 (17)
Pseudogenes (including repeats)	8 (9)
Group II introns	
Cis-splicing	18
Trans-splicing	6
Group I introns	1
Repeats (bp, % of genome)	11,956 bp; 2%
Large repeats (> 1 kb)	3815 bp
Chloroplast derived sequences (# MTPTs)	17
Chloroplast derived sequences (bp, % of genome)	6786 bp; 0.95%
Mitochondrial genes (including cis-spliced introns)	10.16%
Mitochondrial-like sequences (including genes)	72%

of the *O. subterraneum* mtDNA has affinity with this family (Fig. 2), even though a single report of a legume host is known for *O. subterraneum.* Such regions are distributed in 49 chromosomes, have an average identity of 90.4%, and the largest foreign tract is 3920 bp long in the chromosome OsChr24 (Supporting Information Figure S2). Interestingly, all legume-derived regions are noncoding sequences.

Recently, our group found evidence for a genome-scale transfer of mitochondrial sequences (genic and intergenic) from legume hosts to the holoparasitic plant *L. mirabile* (Sanchez-Puerta et al. 2017, 2019). Given that *Lophophytum* and *Ombrophytum* are sister genera (Su et al. 2015), we wondered whether the Fabaceae regions in *O. subterraneum* mtDNA represented ancestral events, in which mitochondrial sequences were acquired in the common ancestor before species divergence. We found 39 legume-derived regions of *O. subterraneum* that were also found in *L. mirabile* (totaling 19.2 kb). We performed phylogenetic analyses of nine of those regions where there was sufficient taxon sampling for meaningful phylogenetic inference (Supporting Information Figure S4). In these trees *O. subterraneum* and *L. mirabile* are not sister taxa but both are associated to the Fabaceae, although with low bootstrap support to distinguish between ancestral or independent horizontal transfers. We identified a single case where *O. subterraneum* and *L. mirabile* are sister taxa within the Fabaceae (Supporting Information Figure S4-OsChr33). Interestingly, the majority of the sequences of the *O. subterraneum* mtDNA derived from Fabaceae (90.6 kb out of 109.8 kb) were not found in *L. mirabile* (Supporting Information Figure S2).

Origin of the *Ombrophytum subterraneum* mitochondrial coding sequences

Phylogenetic analyses of each gene revealed that of the 51 protein-coding gene copies in the *O. subterraneum* mtDNA, 12 show a foreign origin (Table 2). Some genes (*cox3-sdh4*, *nad3-rps12*, and *rpl5-rps14*) were concatenated in the phylogenetic trees because they overlap in *O. subterraneum* and/or are conserved gene clusters in most angiosperms (Richardson et al. 2013). Foreign gene sequences are always placed phylogenetically as sister to or within the angiosperm family Asteraceae with bootstrap support values ≥ 70% (Supporting Information Figure S5). In those trees, sequences from the host plant (*Baccharis* sp., Asteraceae) are not sister to *O. subterraneum* even though they are grouped in the same clade (Supporting Information Figure S5). We classified as native those genes associated with Santalales and those with an unsupported position, probably due to limited information from close relatives. Examples of trees with native genes (*ccmFN*-chr5 and *cox1*) that were vertically inherited and foreign genes (*ccmFN-chr49*) that were transferred from an Asteraceae host to *O. subterraneum* mtDNA are shown in Fig. 3. Noticeably, the tree of the gene *atp4* shows an unexpected and strongly supported clade of diverse angiosperms (Solanaceae, Fabaceae, *L. mirabile*-foreign-, and Brassicales) that share a ~ 60 bp stretch of high identity probably as a result of HGT (unpublished results).

The BLAST analysis of the complete chromosomes shows that all foreign genes are embedded in longer foreign regions that take up approximately the full length of those chromosomes (Supporting Information Figure S2). Furthermore, vertical inheritance of 11 of the 12 foreign gene copies was rejected by the AU test (Table 2). In all cases, the unconstrained trees, which place the foreign copy of the gene in the family Asteraceae, show a high *p* value in the AU test (Supporting Information Figure S6). The scarce availability of mitochondrial genes of other Santalales limited the design of constrained topologies to force a native origin. The only available mtDNAs of Santalales included *Viscum* spp., which presents an extremely fast rate of evolution (Petersen et al. 2015; Skippington et al. 2015) and *L. mirabile,* in which a large fraction of the native genes were replaced by foreign homologs (Sanchez-Puerta et al. 2017).

In contrast to *L. mirabile* mtDNA (Sanchez-Puerta et al. 2017), all identified foreign genes coexist with native homologs in the mtDNA of *O. subterraneum.* Surprisingly, both native and foreign homologs are full length and potentially functional. Sequence comparisons of homologs of different origin identified several synonymous and nonsynonymous differences, indicating that the protein products would be quite different. Analyses of recombination between native and foreign alleles using Geneconv did not find any significant result. Also, there are two native copies of the

Table 2 Features of full-length protein-coding genes in *Ombrophytum subterraneum* mtDNA

Gene name	Chromosome	CDS length (gene length) bp	No. of predicted editing sites	Origin	AU test
atp1	OsChr48 (MT076314)	1530	4	Native	na
atp4	OsChr49 (MT076315)	582	11	Foreign	$p < 0.05$
atp4	OsChr38 (MT076304)	576	11	Foreign	$p < 0.05$
atp4	OsChr8 (MT076274)	564	10	Native	na
atp6	OsChr34 (MT076300)	1182	14	Native	na
atp8	OsChr7 (MT076273)	477	2	Native	na
atp9	OsChr9 (MT076275)	429	9	Foreign	$p < 0.05$
atp9	OsChr34 (MT076300)	225	1	Native	na
ccmB	OsChr11 (MT076277)	618	31	Native	na
ccmC	OsChr38 (MT076304)	744	30	Foreign	$p < 0.05$
ccmC	OsChr1 (MT076267)	756	22	Native	na
ccmFC	OsChr23 (MT076289)	1269 (1986)	16	Native	na
ccmFN	OsChr49 (MT076315)	1743	35	Foreign	$p < 0.05$
ccmFN	OsChr5 (MT076271)	1887	34	Native	na
cob	OsChr4 (MT076270)	1173	13	Native	na
cox1	OsChr17 (MT076283)	1617 (2578)	14	Native	na
cox2	OsChr1 (MT076267)	774	10	Native	na
cox2	OsChr11 (MT076277)	726 (1856)	10	Native	na
cox3	OsChr41 (MT076307)	798	11	Native	na
matR	OsChr49 (MT076315)	1920	12	Foreign	$p < 0.05$
matR	OsChr52 (MT076318)	1761	15	Native	na
mttB	OsChr5 (MT076271)	822	25	Native	na
nad1	OsChr4 (MT076270)-OsChr13 (MT076279)-OsChr52 (MT076318)-OsChr3 (MT076269)	978 (1745)	20	Native	na
nad2	OsChr33 (MT076299)-OsChr24 (MT076290)	1467 (5751)	27	Native	na
nad3	OsChr46 (MT076312)	357	15	Foreign	$p < 0.05$
nad3	OsChr4 (MT076270)	357	12	Native	na
nad4	OsChr19 (MT076285)	1488 (6781)	39	Native	na
nad4L	OsChr8 (MT076274)	303	10	Native	na
nad5	OsChr34 (MT076300)-OsChr17 (MT076283)-OsChr26 (MT076292)	2016 (4036)	24	Native	na
nad6	OsChr16 (MT076282)	792	11	Foreign	$p < 0.05$
nad6	OsChr46 (MT076312)	603	9	Native	na
nad6	OsChr45 (MT076311)	615	11	Native	na
nad7	OsChr35 (MT076301)	1176 (6094)	25	Native	na
nad9	OsChr49 (MT076315)	582	10	Foreign	ns
nad9	OsChr25 (MT076291)	573	11	Foreign	$p < 0.05$
nad9	OsChr4 (MT076270)	573	6	Native	na
rpl2	OsChr5 (MT076271)	996 (1740)	5	Native	na
rpl5	OsChr7 (MT076273)	591	10	Native	na
rpl10	OsChr4 (MT076270)	477	4	Native	na
rpl16	OsChr1 (MT076267)	525	5	Native	na
rps1	OsChr5 (MT076271)	483	3	Native	na
rps3	OsChr1 (MT076267)	1686 (2984)	9	Native	na
rps4	OsChr36 (MT076302)	996	13	Foreign	$p < 0.05$
rps4	OsChr45 (MT076311)	1053	16	Native	na
rps10	OsChr17 (MT076283)	372 (1116)	3	Native	na
rps12	OsChr46 (MT076312)	378	8	Foreign	$p < 0.05$
rps12	OsChr4 (MT076270)	378	3	Native	na
rps19	OsChr1 (MT076267)	291	2	Native	na
sdh3	OsChr11 (MT076277)	381	3	Native	na
sdh4	OsChr41 (MT076307)	462	4	Native	na

na not applicable, *ns* not significant

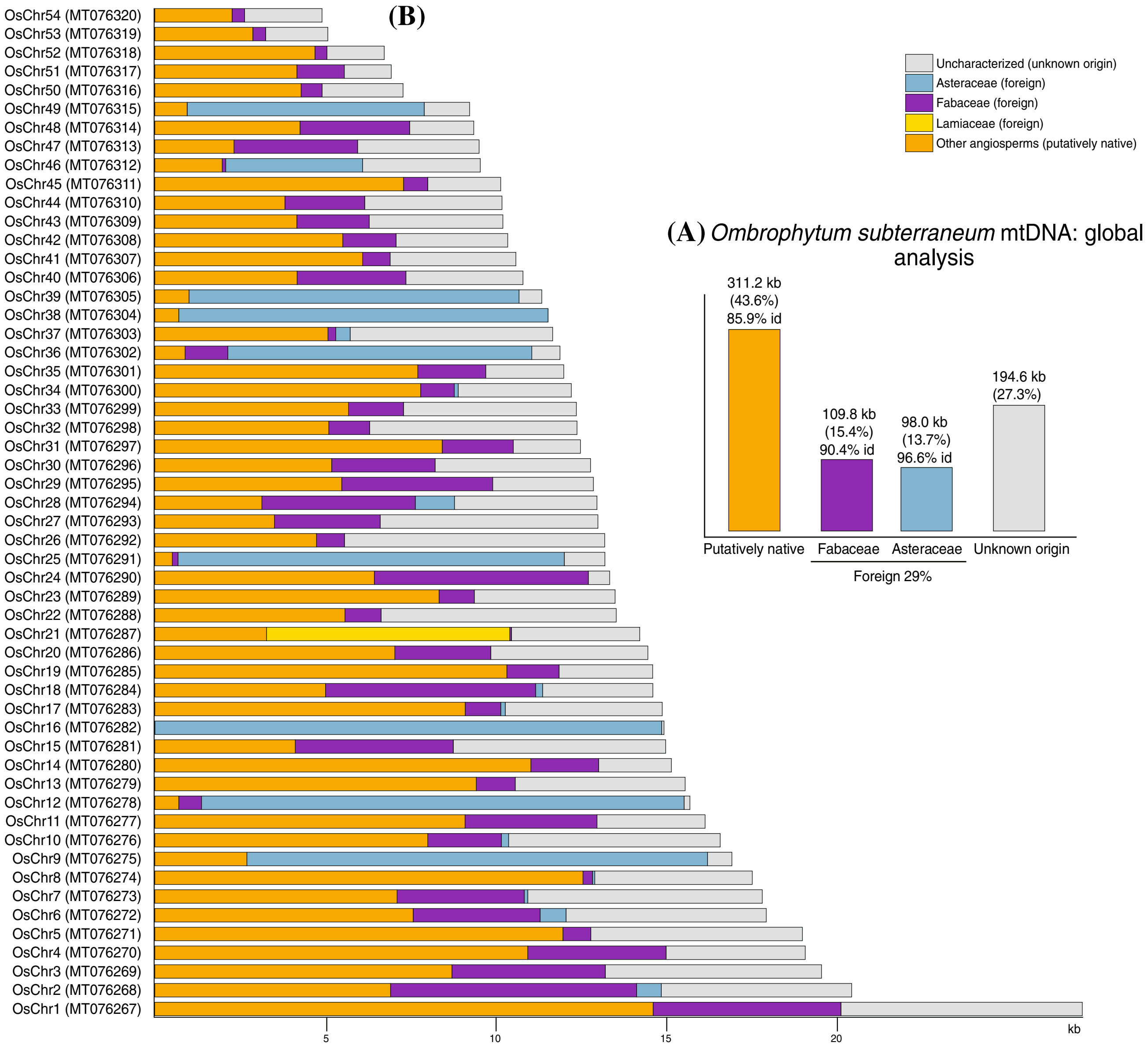

Fig. 2 Evolutionary origin of the mtDNA of *Ombrophytum subterraneum*. Shown is the amount of sequences with affinity to the family Asteraceae (blue), Fabaceae (purple), Lamiaceae (yellow), or other angiosperm mitochondrial genomes (putatively native; in orange), and those regions that found no significant BLAST hits against the angiosperm mtDNAs (light grey), considering the whole genome **(a)** or each mitochondrial chromosome individually **(b)**. The total length, percent of the genome, and average sequence identity (id) of native and foreign regions are indicated for the global analysis **(a)**

gene *cox2*, which differ in the 3′ sequence, predicted editing sites, and intron content.

The three ribosomal RNA coding genes (*rrn5*, *rrnS* and *rrnL*) in *O. subterraneum* are identified as native in the evolutionary trees (Supporting Information Figure S5), while transfer RNAs are too short for phylogenetic reconstruction. Thus, we inferred the origin of tRNAs through BLASTn similarity searches of the surrounding sequences (Supporting Information Figure S2). A total of 12 of the 17 tRNA genes are foreign. Within foreign tracts of Asteraceae origin, there are eight tRNAs (*trnD-GUC, trnF-GAA, trnG-GCC, trnK-UUU, trnM-CAU*- 2 copies, *trnP-UGG, trnS-GCU*). Three tRNAs are associated with the family Fabaceae (*trnM- CAU*, *trnW-CCA*, *trnY-ATA*) and one with the family Lamiaceae (*trnE-UUC*; see below). No native homologs of eight foreign transfer RNAs were identified in the *O. subterraneum* mtDNA (Supporting Information Table S4).

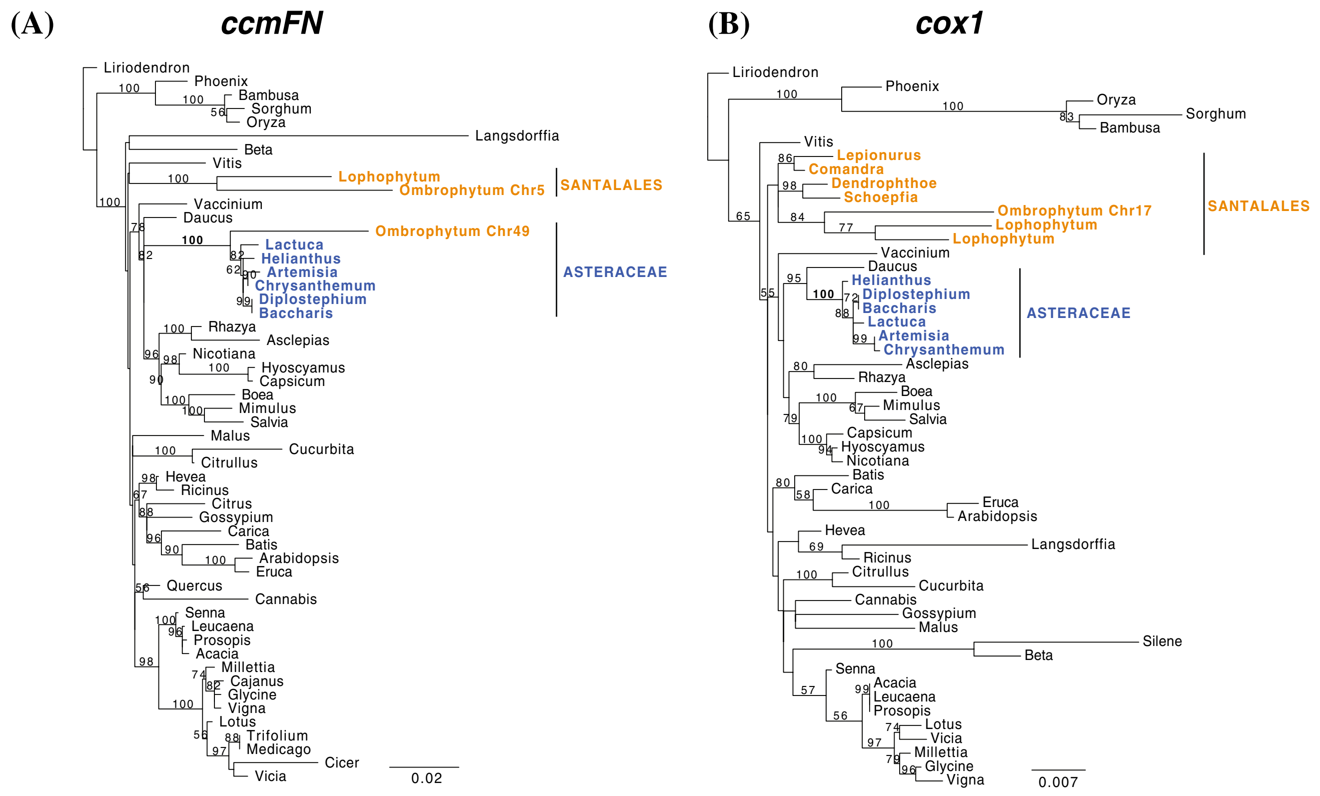

Fig. 3 Maximum likelihood phylogenetic analyses of native and foreign mitochondrial gene copies of *Ombrophytum subterraneum*. **a** Phylogeny of the gene *ccmFN* of *O. subterraneum* (in orange) that shows a native and a foreign copy acquired by horizontal gene transfer from the host (family Asteraceae; in blue). **b** Phylogeny of the gene *cox1* that is vertically inherited in *O. subterraneum*

New data recognizes chimeric genes in *Lophophytum mirabile* mitochondrial genes

The availability of the *O. subterraneum* mtDNA allowed us to re-examine the origin of the *L. mirabile* mitochondrial genes. In particular, we were able to re-assess the chimeric origin of those genes identified previously as foreign, which was not possible without any close relative available for comparison (Sanchez-Puerta et al. 2017, 2019). We evaluated the presence of recombination in all foreign protein-coding genes previously identified in *L. mirabile* (Sanchez-Puerta et al. 2017). Two chimeric genes (*rpl2* and *rps3*) were identified by Geneconv (p value < 0.005). Additional evidence of the chimeric status comes from contrasting affiliations in the trees when these gene sequences were divided into 5′ and 3′ subregions; and from the BLAST similarity searches of flanking sequences (Supporting Information Figure S8). The third chimeric gene in *L. mirabile* mtDNA is *nad5* because it is composed of foreign and native trans-spliced exons (Sanchez-Puerta et al. 2017, 2019).

Furthermore, the inclusion of *O. subterraneum* sequences in the phylogenetic analyses indicated that the genes *cob* and *ccmFN* of *L. mirabile* are likely native. Both genes from *L. mirabile* are found as sister to *O. subterraneum* with high bootstrap support (Supporting Information Figure S5) and no significant BLAST hits are found in their flanking regions in the *L. mirabile* mtDNA (Sanchez-Puerta et al. 2019). Taking all these results into account, the *L. mirabile* mtDNA has 44 protein-coding genes, of which 29 are foreign, 12 are native, and three are chimeric (Supporting Information Table S5).

Chloroplast derived sequences in *Ombrophytum subterraneum* are foreign

Seventeen regions of chloroplast derived sequences (MTPT) were identified in the mtDNA of *O. subterraneum* (Fig. 1). These chloroplast regions include only fragments of chloroplast genes, except for short genes such as *trnS-GGA*, *trnD-UGC*, and *trnW-CCA*. Multiple origins were identified for these MTPTs (Supporting Information Figure S8). Two of them are native (*rrn16* and *rps7*) as they are placed phylogenetically with species of the family Santalales and are probably the result of IGT. This is consistent with the plastid genome (cpDNA) of species belonging to the family Balanophoraceae: *Rhopalocnemis phalloides* (Schelkunov et al. 2019), *Balanophora laxiflora* and *B. reflexa* (Su et al. 2019) that maintain these genes, despite their drastic genome reduction. Interestingly, there is a high difference in GC content between the native MTPTs found in the mtDNA of *O. subterraneum* (GC% = 45%) and the homologous sequences in the cpDNA of the family Balanophoraceae (GC% = 15%). Plastid sequences from Balanophoraceae were not included in the phylogenetic analyses due to the extraordinary AT content and high substitution rate (Schelkunov et al. 2019; Su et al. 2019).

In addition, plastid-derived regions including two copies of *trnS-GGA* (OsChr3 and OsChr20), two of *trnA-UGC* (OsChr46 and OsChr49), and part of the gene *rpl23* were classified as native because they show a weak support in the phylogenetic analyses (Supporting Information Figure S8). Besides, the flanking sequences did not show evidence of a foreign origin in BLAST searches (Supporting Information Figure S2). These tRNAs and the gene *rpl23* are absent in the cpDNAs of the family Balanophoraceae (Schelkunov et al. 2019; Su et al. 2019) including the cpDNA of *O. subterraneum* (unpublished results), raising the possibility of a foreign origin from an unsampled lineage.

The other ten MTPTs resulted of foreign origin with high support in the phylogenetic trees of seven of them (Supporting Information Figure S8). Surprisingly, the MTPTs show diverse origins, including the families Asteraceae (*rrn23*, *trnA-UGC*, *trnD-GUC*), Fabaceae (*cemA*, *trnW-CCA*), Lamiaceae (*trnI-GAU-trnA-UGC*), Apocynaceae (*rbcL*, *rpl20-rpl33*), Euphorbiaceae (*rbcL*), and the order Malvales (*rpoC*). BLAST hits of flanking mitochondrial sequences to each MTPT were grouped according to the origin of the MTPT. The majority of these MTPTs were embedded in larger tracts of foreign mitochondrial sequences of the same lineage as the MTPT (Supporting Information Figure S8). These results indicate that chloroplast sequences were transferred by mitochondrion-to-mitochondrion HGT from a donor lineage that had previously acquired the plastid sequences in its mtDNA by IGT. Of particular relevance is one mitochondrial chromosome of *O. subterraneum*, OsChr21 of 14,211 bp in length, that contains an outstanding amount of sequences (~ 7 kb) with high identity (95.43%) to the mtDNA of *Salvia miltiorrhiza* (Lamiaceae), including the MTPT mentioned above (Supporting Information Figure S8).

Given that ten chromosomes are almost entirely foreign, we compared the read depth and GC content in native and foreign chromosomes and did not find a strong bias (Supporting Information Table S2). Six of the ten foreign chromosomes and 24 of the 44 native chromosomes have an average read depth below the overall mean for the whole mtDNA (71×). The chromosome with the smallest average read depth (40×) is the mostly foreign OsChr46 that encodes the intact genes *nad3*, *nad6*, *rps12*, and *trnM*. The chromosome with the highest read depth (113×) is the native OsChr42 that encodes the genes *rrnS* and *rrn5*. No differences in GC content were found among chromosomes, except in the native chromosome OsChr45 that has a lower GC content (41%).

Discussion

Coding and noncoding chromosomes in *Ombrophytum subterraneum*

The mtDNA of the holoparasite *O. subterraneum* (Balanophoraceae) assembles into 54 circular chromosomes. The largest chromosome is 27,184 bp and the smallest, 4900 bp. This representation is one of the possible alternatives considering the two pairs of direct repeats (in OsChr2 and OsChr14) that could recombine generating two additional smaller chromosomes. In contrast, the other 52 chromosomes do not share repeats with evidence of recombination activity and are likely structurally autonomous. Multichromosomal mitochondrial genomes were described in a few other angiosperms: *Amborella trichopoda* (5 chromosomes), *Cucumis sativus* (2 chromosomes), *L. mirabile* (54 chromosomes), *Saccharum officinarum* (2 chromosomes), *Silene noctiflora* (59 chromosomes), *S. conica* (128 chromosomes), and *S. vulgaris* (2–5 chromosomes) (Alverson et al. 2011; Rice et al. 2013; Sanchez-Puerta et al. 2017; Shearman et al. 2016; Sloan et al. 2012a, 2012b; Wu et al. 2015). Plant mtDNA replication has not been fully elucidated and these autonomous chromosomes may rely on a recombination-dependent replication mechanism (Maréchal and Brisson 2010). Other plant mtDNAs were described as multipartite and even assembled in more than one molecule, such as *Allium cepa* (Tsujimura et al. 2019), *Cynanchum wilfordii* (Kang et al. 2017), *Cynomorium coccineum* (Bellot et al. 2016), *Oryza sativa* (Kazama and Toriyama 2016), and *Solanum tuberosum* (Varré et al. 2019), but those subgenomic molecules could recombine with each other and may not represent autonomous chromosomes.

Like some of the multichromosomal mtDNAs mentioned above, 20 chromosomes in *O. subterraneum* are devoid of intact known genes, although unknown ORFs were identified in 15 of those. In *Silene* spp. the presence of chromosomes lacking any identifiable gene is highly variable among populations within a single species (Wu et al. 2015; Wu and Sloan 2019). The functional significance, maintenance, and the evolutionary role of these putative noncoding chromosomes remain unclear (Alverson et al. 2011; Rice et al. 2013; Sanchez-Puerta et al. 2017; Sloan et al. 2012a; Wu et al. 2015). One hypothesis involves the presence of unidentified coding regions or noncoding elements that play a role in cell function (Grimes et al. 2014; Sanchez-Puerta et al. 2017; Sloan 2013; Sloan et al. 2012a; Wu et al. 2015). Alternatively, noncoding chromosomes may perpetuate selfishly or evolve by genetic drift as they may represent an insignificant cost to the organism (Alverson et al. 2011; Sanchez-Puerta et al. 2017; Sloan 2013; Sloan et al. 2012a; Wu et al. 2015).

Frequent mitochondrion-to-mitochondrion horizontal transfers from diverse donors in *Ombrophytum subterraneum*

The physical association established between host and parasite allows the passage of nutrients, water, mRNA, and DNA (Barkman et al. 2007; Davis and Wurdack 2004; Mower et al. 2004; Nickrent et al. 2004; Roney et al. 2007; Westwood et al. 2009; Yang et al. 2016). Thus, the high rate of horizontal gene transfer could be promoted by this vascular connection (Davis and Xi 2015; Kim et al. 2014; Mower et al. 2010; Rice et al. 2013; Sanchez-Puerta 2014; Westwood et al. 2010; Xi et al. 2013). Plant-to-plant HGT is particularly frequent in mitochondria, which has been explained by active fusion and fission events of this organelle, repeated genomic recombination, and the presence of extremely large intergenic regions (Arimura et al. 2004; Davis and Xi 2015; Keeling and Palmer 2008; Manchekar et al. 2006; Sanchez-Puerta 2014; Shedge et al. 2010).

Numerous foreign tracts were identified in the mitochondrial genome of the holoparasite *O. subterraneum*. The identification of foreign intergenic regions, as well as introns and RNA editing sites in foreign genes, provide evidence for DNA-based transfers. In some cases, large fragments of DNA or even entire chromosomes may have been acquired by *O. subterraneum*. For example, foreign regions related to the family Asteraceae take up almost the full length of nine chromosomes of *O. subterraneum*. It is possible that host subgenomic circular molecules acquired regulatory regions to replicate as independent chromosomes in the recipient mitochondria (Sanchez-Puerta et al. 2019). Alternatively, foreign mitochondrial subgenomic molecules could have recombined with native mitochondrial DNA in *O. subterraneum* giving rise to chimeric chromosomes.

As expected for being the main host lineage, the Asteraceae are the major donors of long continuous mitochondrial sequences in *O. subterraneum*, totaling 98 kb (~ 14%) with high average identity (96.6%) and concentrated in nine chromosomes. Furthermore, foreign protein-coding genes are related only to the family Asteraceae. In addition to those foreign sequences, we also found evidence of regions subjected to mitochondrion-to-mitochondrion transfers from other angiosperm lineages. A large proportion of shorter regions distributed across 49 *O. subterraneum* chromosomes were acquired from legume hosts. These foreign regions encompass 15.4% of the *O. subterraneum* mtDNA. Noticeably, one chromosome gathers a total of ~ 7 kb (51% of the OsChr21) worth of sequences related to *Salvia miltiorrhiza* (Lamiaceae) mtDNA with high identity (95.43%). Species of Lamiaceae are not among the known hosts of *O. subterraneum*. The finding of these clustered foreign regions with such high sequence identity from a single donor lineage suggests that it is a recent HGT event.

Furthermore, analyses of MTPTs, which are particularly useful to discern phylogenetic affiliations, uncover the presence of foreign sequences from additional angiosperm lineages: Euphorbiaceae, Apocynaceae, and Malvales. Interestingly, these three angiosperm lineages have not been described as hosts of *O. subterraneum* either, suggesting that a different transfer mechanism could have taken place. Hypothetical transfer mechanisms include illegitimate pollination, natural grafting, and vector intermediates (Barkman et al. 2007; Bergthorsson et al. 2004; Davis et al. 2005; Davis and Wurdack 2004; Keeling and Palmer 2008; Mower et al. 2010, 2004; Nickrent et al. 2004; Sanchez-Puerta et al. 2017; Stegemann and Bock 2009). It also opens the possibility that *O. subterraneum* could parasitize or has parasitized a broader range of angiosperms that encompass a wide diversity of hosts. Alternatively, foreign sequences from non-host plants could have been transferred first to a host lineage, such as the Asteraceae, and then acquired by *O. subterraneum* through the haustorium from an Asteraceae host (plant-to-plant-to-plant HGT). For example, the MTPT containing *rpoC* from Malvales is flanked by sequences that have affinity to Fabaceae mtDNA. Thus, a legume mtDNA may have received the foreign plastid sequence from the Malvales and then passed it on to *O. subterraneum* by mitochondrion-to-mitochondrion HGT through the haustorium. A similar scenario has been shown for *L. mirabile* where foreign sequences from the order Piperales and the family Salicaceae were identified in both *L. mirabile* and it mimosoid host mtDNAs (Gandini and Sanchez-Puerta 2017; Sanchez-Puerta et al. 2019).

In the case of foreign MTPTs, the most likely route is the intracellular transfer from the chloroplast to the mitochondrion by IGT followed by mitochondrion-to-mitochondrion HGT between unrelated species (Gandini and Sanchez-Puerta 2017; Sanchez-Puerta et al. 2017). Under this hypothesis, foreign chloroplast sequences should be embedded within foreign mitochondrial tracts from the same donor lineage. We found evidence for this only in some MTPTs in *O. subterraneum* (OsChr29 from Fabaceae, OsChr2 from Euphorbiaceae, OsChr25, OsChr38, and OsChr39 from Asteraceae, OsChr21 from Lamiaceae). Another possible route is the horizontal transfer of a foreign plastid region into the recipient cell and its subsequent import into the mitochondria (Gandini and Sanchez-Puerta 2017). Even though less frequent than mitochondrion-to-mitochondrion HGT, plastid-to-mitochondrion HGT may be responsible for some of the foreign MTPTs (Gandini and Sanchez-Puerta 2017).

Other cases of angiosperm mtDNAs with foreign sequences from different donors have been reported. For example, *Cynomorium coccineum* acquired mitochondrial sequences from Caryophyllales and Sapindales (Bellot et al. 2016; Cusimano and Renner 2019), *Geranium* spp. from Gentianales, Lamiales, Malpighiales, and Solanales (Park et al. 2015); and *A. trichopoda* received almost full-length mtDNAs from mosses, green algae, and mitochondrial sequences from diverse angiosperms (Fagales, Oxalidales, Santalales, and Magnoliidae) (Rice et al. 2013). In those cases, diverse transfer mechanisms were proposed, including direct contact (e.g. host-parasite relationship, natural grafting, or illegitimate pollination) or via vector-mediated transfers (e.g. insects, fungi, virus) (Park et al. 2015; Rice et al. 2013).

Remarkable frequency of duplicative HGT in *Ombrophytum subterraneum* mtDNA

Even though most cases of mitochondrial HGT result in foreign pseudogenes, exceptionally, full length foreign and native copies are maintained in the recipient mtDNA, such as a few generally short genes in *A. trichopoda* (Bergthorsson et al. 2004; Rice et al. 2013), *Sapria himalayana* (Xi et al. 2013), and *L. mirabile* (Sanchez-Puerta et al. 2017). In *O. subterraneum* mtDNA, we identified a remarkable amount of duplicative HGT, where 12 foreign genes coexist with native homologs. Three lines of evidence support these horizontal transfers: phylogenetic analyses, in which foreign copies are always placed as sister to or within the family Asteraceae with strong bootstrap support; AU tests that refuted the vertical inheritance of 11 genes; and similarity searches of flanking regions that revealed that these genes are embedded in longer foreign tracts.

The coexistence of foreign and native homologs may generate genetic variability through recombination events. No clear evidence for chimeric genes in *O. subterraneum* genes were found, although cases in which gene conversion involves short gene regions or is discontinuous can be hard to detect (Hao et al. 2010; Palmer et al. 2003). Examining other species from the family Balanophoraceae will probably enable the identification of chimeric genes in *O. subterraneum*, as the availability of the mtDNA of *O. subterraneum* revealed the presence of two new chimeric genes in the *L. mirabile* mtDNA.

The simultaneous expression of two different alleles may be deleterious for the functionality of the cell (Sanchez-Puerta et al. 2015). When foreign and native copies coexist, it is likely that only one of them is actively transcribed and functional. Expression of co-resident foreign and native alleles has been examined in several genes in *A. trichopoda* (Rice et al. 2013) and *L. mirabile* (unpublished results) revealing that the native copies are the only functional alleles. Analyses of gene expression are needed to evaluate whether the native or foreign alleles are functional in the mtDNA of *O. subterraneum,* but we predict that only the native copies will show significant transcription levels.

Ancestral and more recent horizontal transfers in the evolution of *Ombrophytum* and *Lophophytum*

In contrast to the genus *Ombrophytum*, which has low host specificity, *Lophophytum spp.* are only known to parasitize mimosoid legumes (Hansen 1980; Sato 2014). Despite having both a mitochondrial genome of 54 chromosomes, *O. subterraneum* and *L. mirabile* only share less than 20% of their mtDNAs (137 kb) dispersed in short fragments along the chromosomes. The limited synteny among chromosomes of these two closely related parasites indicate that the presence of 54 chromosomes with dissimilar genetic content is the result of independent evolutionary events that involve multiple horizontal acquisitions from different donors in each parasite. However, the presence of a multichromosomal mitochondrial genome is likely an ancestral character.

Of the homologous regions in the mtDNAs of *O. subterraneum* and *L. mirabile*, only ~ 60 kb were recognized as native sequences vertically inherited by both holoparasites from the common ancestor, given that *L. mirabile* replaced most of its mitochondrial genome with mitochondrial sequences from the mimosoid legume hosts (Choi et al. 2019; Kovar et al. 2018; Sanchez-Puerta et al. 2019). Those transfers involved long continuous regions with extremely high identity (~ 96%) to the available mtDNAs from Fabaceae (Sanchez-Puerta et al. 2019). In the mtDNA of *O. subterraneum*, the foreign regions derived from Fabaceae (totaling ~ 110 kb) are much shorter and with lower sequence identity to the currently available legume mtDNAs (90.4% on average). About 20 kb of those legume-derived regions are also found in *L. mirabile* and may represent ancestral HGT events. In contrast, most legume-derived regions in the *O. subterraneum* mtDNA are not shared with *L. mirabile* mtDNA and could be the result of more recent HGT events from Fabaceae that took place after the divergence of the two genera. Alternatively, the short length and moderate identity of those sequences in *O. subterraneum* and Fabaceae may suggest older transfers from legumes to the common ancestor of *Ombrophytum* and *Lophophytum* followed by the loss (or replacement) of these legume-derived sequences from the *L. mirabile* mtDNA.

We propose an evolutionary hypothesis that involves ancestral and recent HGT events during the evolution of these Balanophoraceae (Fig. 4). Under this hypothesis, the common ancestor of *Ombrophytum* and *Lophophytum* acquired mitochondrial sequences by HGT from a legume

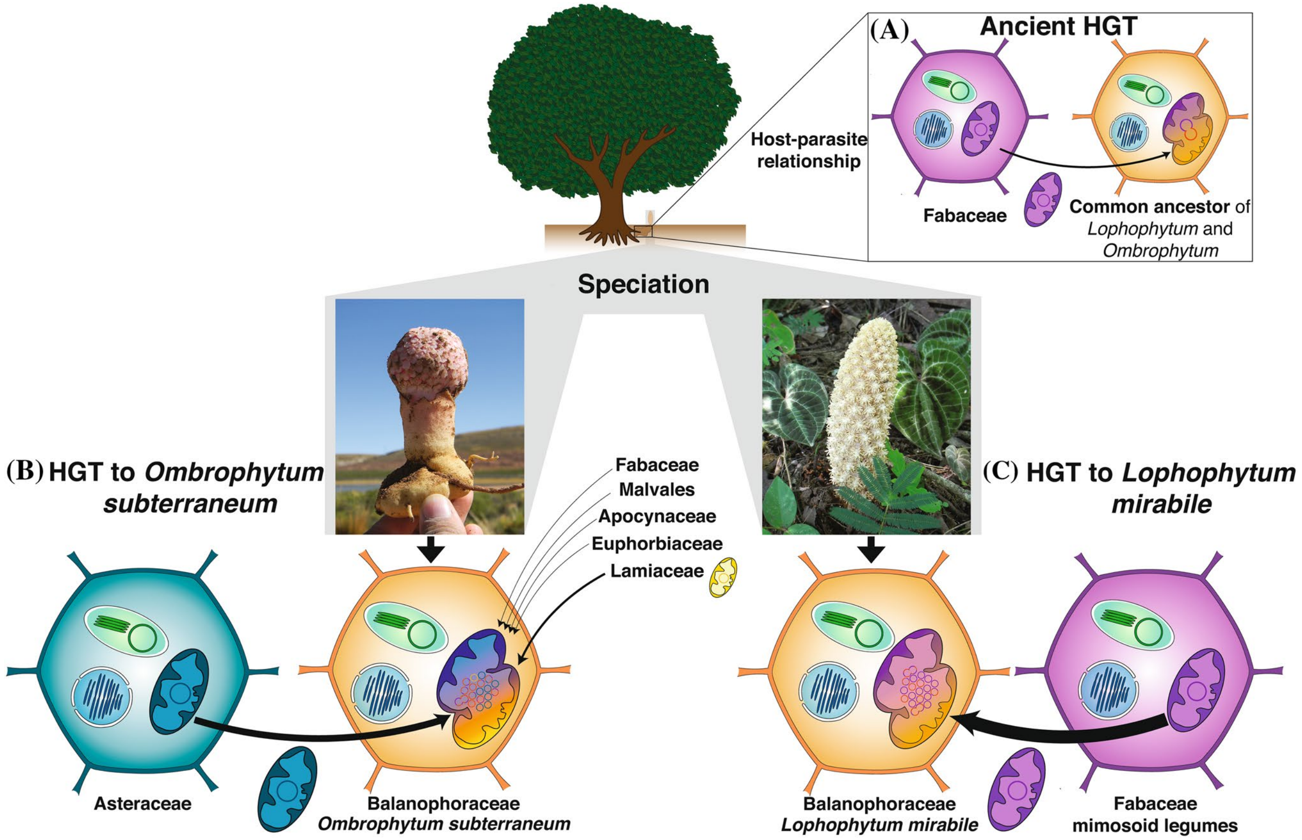

Fig. 4 Hypothesis of horizontal gene transfer in the evolution of two species of the family Balanophoraceae. **a** Ancient mitochondrion-to-mitochondrion horizontal gene transfer from a legume host (purple) to the common ancestor of *Ombrophytum subterraneum* and *Lophophytum mirabile* (orange). **b** Horizontal gene transfer events in the mitochondria of *O. subterraneum* from species of Asteraceae, Lamiaceae, Apocynaceae, Euphorbiaceae, Fabaceae, and Malvales. **c** Massive horizontal transfer of mitochondrial sequences of mimosoid legume hosts to the *L. mirabile* mtDNA

host. After species divergence, some of those ancestral foreign sequences were conserved in the descendants, but *L. mirabile* was subjected to massive and more recent HGT events from mimosoid legume hosts (explaining the high identity between most foreign sequences and mimosoid mtDNAs). Additional transfers of mitochondrial sequences from legume hosts to *O. subterraneum* may have occurred after species divergence, but there is not sufficient evidence to support this. In addition, the *O. subterraneum* mtDNA received mitochondrial and chloroplast sequences from four diverse angiosperm lineages not reported as hosts through unknown HGT mechanisms. Finally, the most recent incorporation of mitochondrial foreign sequences into *O. subterraneum* mtDNA came from Asteraceae and Lamiaceae, with the Asteraceae transfer(s) being outstanding with respect to the amount of DNA transferred. Evidence for recency of the latter HGT events include the concentration of foreign sequences in a few chromosomes, the high sequence identity with the available Asteraceae mtDNAs, and the presence of full-length foreign genes coexisting with native homologs.

Acknowledgements We thank A.A. Edera for help with scripts and graphs and J.D. Palmer for critical reading of the manuscript. This work was supported by Agencia Nacional de Promoción Científica y Tecnológica (Grant No PICT-0691) and Universidad Nacional de Cuyo (Grant No 06/A724) to M.V.S.P, and by NSF (Grant No 1062432) to Indiana University, which supports the computer cluster.

Author contributions MER and MVSP designed the study, analyzed the data and wrote the manuscript. H.S. was involved in sample collection and taxonomic identification. LEG and GP performed host and parasite DNA extraction and purification and participated in the analysis. CLG assembled the mtDNA and helped with scripts for data analyses.

References

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). Mol Biol Evol 27:1436–1448

Arimura S, Yamamoto J, Aida GP, Nakazono M, Tsutsumi N (2004) Frequent fusion and fission of plant mitochondria with unequal nucleoid distribution. Proc Nat Acad Sci USA 101:7805–7808

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 Evol Biol 7:248. https://doi.org/10.1186/1471-2148-7-248

Bellot S et al (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

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*. Proc Natl Acad Sci USA 101:17747–17752

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

Choi I-S et al (2019) Fluctuations in Fabaceae mitochondrial genome size and content are both ancient and recent. BMC Plant Biol 19:1–15

Cusimano N, Renner SS (2019) Sequential horizontal gene transfers from different hosts in a widespread Eurasian parasitic plant. Cynomorium coccineum Am J Bot 106:679–689

Davis C, Andersen W, Wurdack K (2005) Gene transfer from a parasitic flowering plant to a fern. Proc R Soc Lond B Biol Sci 272:2237–2242

Davis C, Wurdack K (2004) Host-to-parasite gene transfer in flowering plants: phylogenetic evidence from Malpighiales. Science 305:676–678

Davis CC, Xi Z (2015) Horizontal gene transfer in parasitic plants. Curr Opin Plant Biol 26:14–19

Delprete PG (2014) *Ombrophytum guayanensis*, the first record of subfamily Lophophytoideae (Balanophoraceae) in the Guayana Shield. Phytotaxa 175:263–269

Doyle JJ, Doyle JL (1987) A rapid isolation procedure for small quantities of fresh leaf tissues. Phytochem Bull 19:11–15

Edera AA, Gandini CL, Sanchez-Puerta MV (2018) Towards a comprehensive picture of C-to-U RNA editing sites in angiosperm mitochondria. Plant Mol Biol 97:215–231

Gandini CL, Sanchez-Puerta MV (2017) Foreign plastid sequences in plant mitochondria are frequently acquired via mitochondrion-to-mitochondrion horizontal transfer. Sci Rep 7:43402–43402. https://doi.org/10.1038/srep43402

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

Greiner S, Lehwark P, Bock R (2019) OrganellarGenomeDRAW (OGDRAW) version 1.3.1: expanded toolkit for the graphical visualization of organellar genomes. Nucleic Acids Res 47:W59–W64. https://doi.org/10.1093/nar/gkz238

Grimes BT, Sisay AK, Carroll HD, Cahoon AB (2014) Deep sequencing of the tobacco mitochondrial transcriptome reveals expressed ORFs and numerous editing sites outside coding regions. BMC Genomics 15:31. https://doi.org/10.1186/1471-2164-15-31

Guzmán-Guzmán S (2019) A new species of *Ombrophytum* (Balanophoraceae), a genus not previously recorded for Colombia. Phytotaxa 424:061–066

Hansen B (1980) Balanophoraceae. Flora Neotróp 23:1–80

Hao W, Richardson AO, Zheng Y, Palmer JD (2010) Gorgeous mosaic of mitochondrial genes created by horizontal transfer and gene conversion. Proc Natl Acad Sci USA 14:21576–21581

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. Mol Biol Evol 29:311–320

Kang S-H, Oh J-H, Kim H-J, Kim C-K (2017) The multipartite mitochondrial genome of *Cynanchum wilfordii* (Gentianales: Apocynaceae). Mitochondrial DNA Part B 2:720–721. https://doi.org/10.1080/23802359.2017.1390405

Kazama T, Toriyama. (2016). Whole mitochondrial genome sequencing and re-examination of a cytoplasmic male sterility-associated gene in Boro-Taichung- type cytoplasmic male sterile rice. PloS One

Keeling PJ, Palmer JD (2008) Horizontal gene transfer in eukaryotic evolution. Nat Rev Genet 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

Koressaar T, Remm M (2007) Enhancements and modifications of primer design program Primer3. Bioinformatics 23:1289–1291

Kovar L et al (2018) PacBio-based mitochondrial genome assembly of *Leucaena trichandra* (Leguminosae) and an intrageneric assessment of mitochondrial RNA editing. Genome Biol Evol 10:2501–2517. https://doi.org/10.1093/gbe/evy179

Kubo T, Mikami T (2007) Organization and variation of angiosperm mitochondrial genome. Physiol Plant 129:6–13

Kuijt J, Delprete PG (2019) A new species of *Ombrophytum* (Balanophoraceae) from Chile, with notes on subterranean organs and vegetative reproduction in the family. Phytotaxa 420:264–272

Langmead B, Salzberg SL (2012) Fast gapped-read alignment with Bowtie 2. Nat Methods 9:357–359. https://doi.org/10.1038/nmeth.1923

Larsson A (2014) AliView: a fast and lightweight alignment viewer and editor for large datasets. Bioinformatics 30:3276–3278. https://doi.org/10.1093/bioinformatics/btu531

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

Manchekar M, Scissum-Gunn K, Song D, Khazi F, McLean S, Nielsen BL (2006) DNA recombination activity in soybean mitochondria. J Mol Biol 356:288–299

Maréchal A, Brisson N (2010) Recombination and the maintenance of plant organelle genome stability. New Phytol 186:299–317

Mauseth JD, Hsiao S-C, Montenegro G (1992) Vegetative body of the parasitic angiosperm *Ombrophytum subterraneum* (Balanophoraceae). Bull Torrey Bot Club 119:407–417

Molina J et al (2014) Possible loss of the chloroplast genome in the parasitic flowering plant *Rafflesia lagascae* (Rafflesiaceae). Mol Biol Evol 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 Biol 8:150

Mower JP, Stefanovic S, Young GJ, Palmer JD (2004) Plant genetics: gene transfer from parasitic to host plants. Nature 432:165–166

Nickrent DL, Blarer A, Qiu YL, Vidal-Russell R, Anderson FE (2004) Phylogenetic inference in rafflesiales: the influence of rate heterogeneity and horizontal gene transfer. BMC Evol Biol 4:40

Palmer S, Schidkraut E, Lazarin R, Nguyen J, Nickoloff J (2003) Gene conversion tracts in *Saccharomyces cerevisiae* can be extremely short and highly directional. Nucleic Acids Res 31:1164–1173

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 Phytol 208:570–583

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

Phanstiel DH. (2019). Sushi: tools for visualizing genomics data. R package version 1.24.0.

Rice DW et al (2013) Horizontal transfer of entire genomes via mitochondrial fusion in the angiosperm *Amborella*. Science 342:1468–1473

Richardson AO, Rice DW, Young GJ, Alverson AJ, Palmer JD (2013) The "fossilized" mitochondrial genome of *Liriodendron tulipifera*: an ancestral gene content and order, ancestral editing sites, and extraordinarily low mutation rate. BMC Biol 11:29

Roney J, Khatibi P, Westwood J (2007) Cross-species translocation of mRNA from host plants into the parasitic plant doddler. Plant Physiol 143:1037–1043

Sanchez-Puerta MV (2014) Involvement of plastid, mitochondrial and nuclear genomes in plant-to-plant horizontal gene transfer. Acta Soc Bot Pol 83:317–323

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

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

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 Phytol 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, Universidad Nacional del Nordeste.

Sawyer S (1989) Statistical tests for detecting gene conversion. Mol Biol Evol 6:526–538

Schelkunov MI, Nuraliev MS, Logacheva MD (2019) *Rhopalocnemis phalloides* has one of the most reduced and mutated plastid genomes known. PeerJ 7:e7500–e7500. https://doi.org/10.7717/peerj.7500

Shearman JR et al (2016) The two chromosomes of the mitochondrial genome of a sugarcane cultivar: assembly and recombination analysis using long PacBio reads. Sci Rep. https://doi.org/10.1038/srep31533

Shedge V, Davila J, Arrieta-Montiel M, Mohammed S, Mackenzie S (2010) Extensive rearrangement of the *Arabidopsis* mitochondrial genome elicits cellular conditions for thermotolerance. Plant Physiol 152:1960–1970

Shimodaira H (2002) An approximately unbiased test of phylogenetic tree selection. Syst Biol 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. Proc Nat Acad Sci USA 112:E3515–E3524

Sloan D (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

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

Sloan D, Müller KM, McCauley D, Taylor DR, Štorchová H (2012b) Intraspecific variation in mitochondrial genome sequence, structure, and gene content in *Silene vulgaris*, an angiosperm with pervasive cytoplasmic male steritility. New Phytol 196:1228–1239

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

Stegemann S, Bock R (2009) Exchange of genetic material between cells in plant tissue grafts. Science 324:649–651. https://doi.org/10.1126/science.1170397

Straub SCK, Parks M, Weitemier K, Fishbein M, Cronn RC, Liston A (2012) Navigating the tip of the genomic iceberg: next-generation sequencing for plant systematics. Am J Bot 99:349–364. https://doi.org/10.3732/ajb.1100335

Su H-J et al (2019) Novel genetic code and record-setting AT-richness in the highly reduced plastid genome of the holoparasitic plant *Balanophora*. Proc Natl Acad Sci USA 116:934–943. https://doi.org/10.1073/pnas.1816822116

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. Sinauer Associates, Sunderland.

Tsujimura M, Kaneko T, Sakamoto T, Kimura S, Shigyo M, Yamagishi H, Terachi T (2019) Multichromosomal structure of the onion mitochondrial genome and a transcript analysis. Mitochondrion 46:179–186

Untergasser A, Cutcutache I, Koressaar T, Ye J, Faircloth BC, Remm M, Rozen SG (2012) Primer3—new capabilities and interfaces. Nucleic Acids Res 40:e115

Varré J-S et al (2019) Complete sequence, multichromosomal architecture and transcriptome analysis of the *Solanum tuberosum* mitochondrial genome. Int J Mol Sci 20:4788

Westwood J, Yoder JA, Timko M, dePamphilis CW (2010) The evolution of parasitism in plants. Trends Plant Sci 15:227–235

Westwood JH, Roney JK, Khatibi PA, Stromberg VK (2009) RNA translocation between parasitic plants and their hosts. Pest Manag Sci 65:533–539

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. Proc Nat Acad Sci USA 112:10185–10191

Wu Z, Sloan DB (2019) Recombination and intraspecific polymorphism for the presence and absence of entire chromosomes in mitochondrial genomes. Heredity 122:647–659

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 Genet 9:e1003265

Yang Z et al (2016) Horizontal gene transfer is more frequent with increased heterotrophy and contributes to parasite adaptation. Proc Natl Acad Sci U S A 113:E7010. https://doi.org/10.1073/pnas.1608765113

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

Zuloaga FO, Morrone ON, Belgrano MJ, Marticorena C, Marchesi E. (2008). Catálogo de las plantas vasculares del Cono Sur. Monogr Syst Bot Missouri Bot Gard, 107.