Molecular Biology and Evolution, 2025, **42**, 1–15
https://doi.org/10.1093/molbev/msaf082
Advance access publication 7 April 2025
Discoveries

MBE OXFORD

Unequally Abundant Chromosomes and Unusual Collections of Transferred Sequences Characterize Mitochondrial Genomes of *Gastrodia* (Orchidaceae), One of the Largest Mycoheterotrophic Plant Genera

Hanchen Wang [1,2] Deyi Wang [3,4] Bingyi Shao,[1,2] Jingrui Li [1] Zhanghai Li [5] Mark W. Chase,[6,7] Jianwu Li,[8] Yanlei Feng [9] Yingying Wen [1] Shiyu Qin [1,10] Binghua Chen [11] Zhiqiang Wu [12,*] Xiaohua Jin [1,*]

[1]State Key Laboratory of Plant Diversity and Specialty Crops, Institute of Botany, Chinese Academy of Sciences, Beijing, China
[2]College of Life Sciences, University of Chinese Academy of Sciences, Beijing, China
[3]Chengdu Institute of Biology, Chinese Academy of Sciences, Chengdu, China
[4]Naturalis Biodiversity Center, Leiden, the Netherlands
[5]Key Laboratory of Chemistry in Ethnic Medicinal Resources, Ministry of Education, Yunnan Minzu University, Kunming, China
[6]Department of Environment and Agriculture, Curtin University, Bentley, Australia
[7]Royal Botanic Gardens, Kew, Richmond, UK
[8]Xishuangbanna Tropical Botanical Garden, Chinese Academy of Sciences, Mengla County, Yunnan, China
[9]ZJU-Hangzhou Global Scientific and Technological Innovation Center, Zhejiang University, Hangzhou, China
[10]School of Life Sciences, Nanchang University, Nanchang, China
[11]College of Life Sciences, Fujian Normal University, Fuzhou, China
[12]Shenzhen Branch, Guangdong Laboratory of Lingnan Modern Agriculture, Key Laboratory of Synthetic Biology, Ministry of Agriculture and Rural Affairs, Agricultural Genomics Institute at Shenzhen, Chinese Academy of Agricultural Sciences, Shenzhen, China

***Corresponding authors:** E-mails: wuzhiqiang@caas.cn; xiaohuajin@ibcas.ac.cn.
Associate editor: Jun Gojobori

Abstract

The mystery of genomic alternations in heterotrophic plants is among the most intriguing in evolutionary biology. Compared to plastid genomes (plastomes) with parallel size reduction and gene loss, mitochondrial genome (mitogenome) variation in heterotrophic plants remains underexplored in many aspects. To further unravel the evolutionary outcomes of heterotrophy, we present a comparative mitogenomic study with 13 de novo assemblies of *Gastrodia* (Orchidaceae), one of the largest fully mycoheterotrophic plant genera, and its relatives. Analyzed *Gastrodia* mitogenomes range from 0.56 to 2.1 Mb, each consisting of numerous, unequally abundant chromosomes or contigs. Size variation might have evolved through chromosome rearrangements followed by stochastic loss of "dispensable" chromosomes, with deletion-biased mutations. The discovery of a hyper-abundant (~15 times intragenomic average) chromosome in two assemblies represents the hitherto most extreme copy number variation in any mitogenomes, with similar architectures discovered in two metazoan lineages. Transferred sequence contents highlight asymmetric evolutionary consequences of heterotrophy: despite drastically reduced intracellular plastome transfers convergent across heterotrophic plants, their rarity of horizontally acquired sequences sharply contrasts parasitic plants, where massive transfers from their hosts prevail. Rates of sequence evolution are markedly elevated but not explained by copy number variation, extending prior findings of accelerated molecular evolution from parasitic to heterotrophic plants. Putative evolutionary scenarios for these mitogenomic convergence and divergence fit well with the common (e.g. plastome contraction) and specific (e.g. host identity) aspects of the two heterotrophic types. These idiosyncratic mycoheterotrophs expand known architectural variability of plant mitogenomes and provide mechanistic insights into their content and size variation.

Keywords: organelle genome, convergent evolution, plant-microbe interaction, horizontal gene transfer, mutation rate

Introduction

Plant cells carry two organelle genomes that both originated from ancient endosymbionts (Gray et al. 1999; Burger and Lang 2003) but have undergone distinct evolutionary trajectories (Giannakis et al. 2024; Wang et al. 2024). Compared to plastomes, mitogenomes of angiosperms are characterized by generally conserved coding contents but staggering diversity in size, architecture, and intergenic sequence components (Palmer and Herbon 1988; Bergthorsson et al. 2003; Mower et al. 2012b; Mower 2020). Angiosperm mitogenomes usually contain a subset of around 41(mostly 20 to 40) protein-encoding genes putatively present in their common ancestor (Richardson et al. 2013), plus three ribosomal RNA (rRNA) and variable numbers of transfer RNA (tRNA) genes (Michaud et al. 2011; Warren et al. 2021). Protein-encoding genes typically evolve at slower paces than those of plastid and nuclear genomes (Palmer and Herbon 1988; Bromham et al. 2013; Mower 2020). Yet, rampant

Received: October 12, 2024. **Revised:** March 12, 2025. **Accepted:** April 1, 2025

sequence turnover in intergenic regions often leads to drastic variation in size and architecture between mitogenomes of closely related plant taxa (Ward et al. 1981; Alverson et al. 2010; Sloan et al. 2012; Lin et al. 2022b; Yu et al. 2023) and in some cases, within species (Allen et al. 2007; Wu et al. 2015). Angiosperm mitogenomes assembled to date vary in size by approximately 180 folds, spanning from 66 kb to over 11 Mb (Sloan et al. 2012; Wu et al. 2022). Though often visualized as "master circles" by convention, recent studies have suggested the potential prevalence of multichromosomal architectures in plant mitogenomes, with some complicated ones comprising dozens of, sometimes unequally abundant, chromosomes (Alverson et al. 2011; Sloan et al. 2012, 2018; Wu et al. 2015, 2022). Such size and structural diversity is related to the transfer of foreign DNAs and homologous recombination mediated by repeated sequences (Backert et al. 1997; Gandini and Sanchez-Puerta 2017; Kozik et al. 2019). Sequences of various origins, including conspecific plastid and nuclear genomes, as well as distant plant mitogenomes (Mower et al. 2012a, 2012b) prevail in angiosperm mitogenomes, but their amount can vary greatly (Sanchez-Puerta et al. 2017; Wang et al. 2024). Many of these insights derive from comparative mitogenomic studies that sample closely related taxa (e.g. Sloan et al. 2012; Kozik et al. 2019; Liu et al. 2020; Fan et al. 2022). However, the limited taxonomic coverage of such studies, in particular for species with unusual mitogenomic features, has hampered a deeper understanding of how diversity and alternations of plant mitogenomes emerged.

Among known variations of plant genomes, many extremes have been observed in heterotrophic plants—those obtaining nutrients from other plants (parasites) or symbiotic fungi (mycoheterotrophs) (Merckx and Freudenstein 2010; Wicke and Naumann 2018; Yuan et al. 2018). Plastomes of both heterotrophic types, for example, have undergone convergent reductions in size and gene content (Krause 2008; Braukmann et al. 2017; Graham et al. 2017), due mainly to relaxed selection on photosynthetic capability. Similar reductions have been discovered in the mitogenomes of two parasitic plants. Being only 66 and 131 kb, respectively, mitogenomes of the aerobic parasitic *Viscum scurruloideum* (Viscaceae) and the Balanophoraceae member *Rhopalocnemis phalloides* are the smallest among assembled vascular plants (Skippington et al. 2015; Yu et al. 2022). The *V. scurruloideum* mitogenome is also distinctive for the loss of nearly half of the "core genes", coupled with extreme sequence divergence in the remaining ones (Skippington et al. 2015). Another mitochondrial feature common to parasitic plants is the incorporation of foreign genes from their hosts, which has been documented in four of six holoparasitic lineages with published mitogenomes (Barkman et al. 2007; Yu et al. 2023). This is best exemplified by some members of Balanophoraceae and Rafflesiaceae, where about half of the mitochondrial genes are horizontally acquired (Xi et al. 2013; Sanchez-Puerta et al. 2017). These modifications are rooted in several aspects of the parasitic lifestyle, such as relaxed selection of genes due to nutritional dependence on the host (Skippington et al. 2015) and intimate physical associations that facilitate transmission of DNA (Mower et al. 2012a).

Given their similarities with parasitic plants—obtaining nutrients from external hosts and comparable reduction in plastomes—one might expect that mycoheterotrophs exhibit mitogenomic features analogous to those identified in parasitic plants. Yet, this speculation remains largely unverified, as only six assemblies are currently available for full mycoheterotrophs: that of *Danxiaorchis yangii*, *Epipogium roseum* (Orchidaceae), *Epirixanthes elongata* (Polygalaceae), *Hypopitys monotropa* (Ericaceae), *Gastrodia elata*, and *G. pubilabiata* (Orchidaceae) (Yuan et al. 2018; Petersen et al. 2019; Shtratnikoya et al. 2020; Kim et al. 2023; Zhao et al. 2024; Liu et al. 2025). Aspects of mitogenomes examined in these studies—such as size, gene content, and horizontally transferred components—showed features overall typical to angiosperms. More extensive mitogenomic analyses with denser taxon sampling and a phylogenetic framework therefore appear to be critical for further uncovering potential mitogenomic peculiarities in mycoheterotrophs.

Here, we present a comparative mitogenomic study with a focus on *Gastrodia* (Orchidaceae). With ca. 100 described species (POWO 2023), *Gastrodia* is among the most species-rich genera that are fully mycoheterotrophic throughout their life cycle. Most members of this genus are found in Asia, and the others occur in Africa and Australasia (POWO 2023). Plastomes of *Gastrodia*, like those of many other heterotrophic plants, are drastically reduced in gene content and size (Wen et al. 2022). Prior analyses have reported mitogenome expansion for two *Gastrodia* species and provided insights into the transfer of plastid sequences (Yuan et al. 2018; Kim et al. 2023). In this study, we aim to further illuminate the mitogenomic variation and their evolution in mycoheterotrophic plants. Mitogenomes of 11 *Gastrodia* species and two of their relatives (including one with putatively independently evolved mycoheterotrophy) were assembled. Under a phylogenetic framework, extensive analyses were performed, with emphasis on their architectures, transferred sequence components, and rates of molecular evolution. We propose evolutionary scenarios for observed mitogenomic peculiarities by comparing with studies on parasitic plants and by placing the observations in the context of angiosperm organelle genomes. Our findings highlight the remarkable and diverse impacts of heterotrophy at the mitogenome level and promote a mechanistic understanding of mitogenomic alternations.

Results

Mitogenome Assembly and Annotation

We sampled 11 *Gastrodia* species plus two in related genera, *Epipogium aphyllum* (mycoheterotrophic) and *Nervilia mackinnonii* (autotrophic). All mitogenomes are assembled as multiple circular-mapping chromosomes ranging from 23 to 71 each, except for *G. peichatieniana* where 5 linear contigs are also present (Table 1). Given the consistently high sequencing depths (47.8 to 563.0 ×) within each assembly (supplementary fig. S1, Supplementary Material online) and the absence of potentially missing mitochondrial contigs, we assess the assembly for *G. peichatieniana* as "near complete", and all the rest as "complete". Analyzed *Gastrodia* species are split into three highly supported (all bootstrap supports [BS] > 0.95) clades according to mitochondrial molecular phylogenetics (Fig. 1a), with Clade 1 (sister to Clade 2 + Clade 3) consisting of *G. javanica*, Clade 2 of *G. elata* and *G. angusta*, and Clade 3 of all the rest. Sizes of assembled mitogenomes vary by 5 folds, spanning from a modest 417 kb in *E. aphyllum* to over 2 Mb in *G. javanica* (Table 1). Within *Gastrodia*, the size range is 3.8 folds (smallest: *G. peichatieniana*, 563 kb), with the maximum seen in the earliest-diverging Clade 1 and

Table 1 Summary of assembled orchid mitogenomes

Feature	*Epipogium aphyllum*	*Gastrodia angusta*	*G. crispa*	*G. elata*	*G. flexistyla*	*G. javanica*	*G. longistyla*	*G. menghaiensis*	*G. peichatieniana*	*G. shimizuana*	*G.* sp.	*G. uraiensis*	*Nervilia mackinnonii*
Genome size (bp)	416,775	1,151,371	974,735	1,508,320	1,102,691	2,131,312	629,292	579,585	562,999	955,081	959,500	1,011,371	1,302,313
Assembled structure	23 circular	53 circular	63 circular	65 circular	71 circular	66 circular	40 circular	37 circular	33 circular + 5 linear	61 circular	60 circular	70 circular	58 circular
Median chromosome length (bp)	16,440	21,054	14,898	22,161	14,782	27,852	14,479	15,183	14,608	15,290	15,870	13,853	21,055
GC Content (%)	45.2	43.1	43.6	44.5	43.7	44.9	43.7	44.2	44.3	43.8	43.7	43.7	43.8
Assembly version	Complete	Complete	Complete	Complete	Complete	Complete	Complete	Complete	Near complete	Complete	Complete	Complete	Complete
Proportion of protein-coding sequences (%)	8.03	2.86	3.37	2.19	2.99	1.54	5.21	5.65	5.82	3.43	3.41	3.26	2.55

decreasing toward more recently diverged species. The trend of variation in median chromosome length (2 folds) largely follows that of mitogenome size, although those in Clade 3 are generally stable compared to differences among clades (Table 1 and supplementary fig. S2, Supplementary Material online). Despite size variation, all assembled mitogenomes show stable guanine-cytosine (GC) contents of around 44% (Table 1) and harbor similar magnitudes of repetitive contents (9.3% to 16.1%). Among these, transposable elements (TEs) contribute 4.1% to 6.4% of the size of each assembly, of which long terminal repeat (LTR)-retrotransposons appear to be the most abundant type (2.2% to 3.1%). In some published orchid mitogenomes, such as that of *Paphiopedilum micranthum*, the proportional contribution of TEs reaches nearly 8% (supplementary table S1, Supplementary Material online).

All assemblies have an identical set of protein-encoding genes, including all 24 "core genes" and 14/17 "variable genes" (Mower 2020), resulting in a total of ca. 33-kb protein-encoding sequences (supplementary fig. S3, Supplementary Material online). Two nonidentical copies of *ccmC* and *nad6* genes are present in *N. mackinnonii*, and a duplicate gene fragment exists in *E. aphyllum* (*atp9*), *G. angusta* (*atp4*), *G. crispa* (*atp9*), and *G. flexistyla* (*atp9*), respectively. In addition to a common set of 25 Group II introns (Mower et al. 2012b), one Group I intron (cox1i729) is annotated in all *Gastrodia* mitogenomes, but not in that of *E. aphyllum* or *N. mackinnonii*. Notably, a unique insertion of 219 to 321 bp was identified in the *rps3* gene of all *Gastrodia* mitogenomes (supplementary fig. S4, Supplementary Material online). In addition, all assembled mitogenomes contain three rRNA genes and a variable number of tRNA genes ranging from seven in *G. peichatieniana* and *G. menghaiensis* to 26 in *N. mackinnonii* (including duplicates; supplementary fig. S5, Supplementary Material online). The number of mitochondrial tRNA genes in the mycoheterotroph *Gastrodia* and *Epipogium* falls behind their autotrophic relatives and many other angiosperms (supplementary fig. S5, Supplementary Material online).

Synteny among these assemblies is limited in general. Pairwise sequence similarity (proportion of mitochondrial sequences covered by syntenic blocks in pairwise comparisons) spans a broad range among *Gastrodia* congeners, with the highest detected between closely related species pairs, e.g. *G. menghaiensis*/*G. peichatieniana* (93.8%) and *G. shimizuana*/*G.* sp. (97.8%), and the lowest between more divergent pairs, e.g. *G. javanica*/*G. uraiensis* (10.7%) and *G. javanica*/*G. longistyla* (10.9%), while those between genera are mostly below 10%, of which nearly half are contributed by protein-encoding genes (Fig. 1 and supplementary fig. S2, Supplementary Material online). This is also reflected at the chromosome scale, as hardly any chromosome is conserved among clades, while within Clade 3 many have homologues in other species (supplementary table S2, Supplementary Material online).

Copy Number of Mitochondrial Chromosomes and Correlation With Length

The relative copy number of mitochondrial chromosomes (including linear contigs in *G. peichatieniana*) and the uncertainty of estimates were calculated based on sequencing depth (Wu et al. 2015; Sanchez-Puerta et al. 2017; Zwonitzer et al. 2024) (supplementary fig. S1,

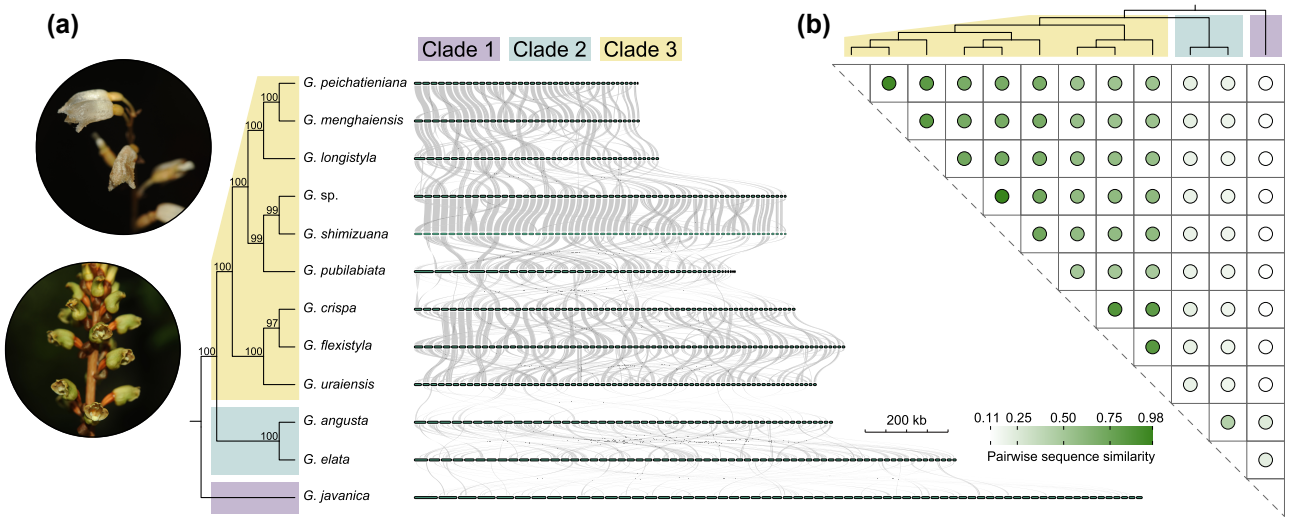

Fig. 1. The multichromosomal mitogenomes of *Gastrodia* with massive sequence turnover. a) Assembled structure and syntenic blocks among *Gastrodia* spp. mitogenomes, with their reconstructed phylogeny and photos of flowers of *G. elata* (lower) and *G. menghaiensis* (upper); dark green-colored blocks denote chromosomes or contigs and gray bands between species denote homologous sequences >1,000 bp long; chromosomes or contigs are ordered by descending length and drawn to the same scale (see scale at bottom-right); the phylogenetic tree is a subset of a larger one (same for Fig. 4a) shown in supplementary fig. S12, Supplementary Material online, with percentage bootstrap support labeled for each node. b) Heat map of pairwise sequence similarity (proportion of mitogenome covered by homologous sequences) among *Gastrodia* mitogenomes (also see supplementary fig. S2, Supplementary Material online).

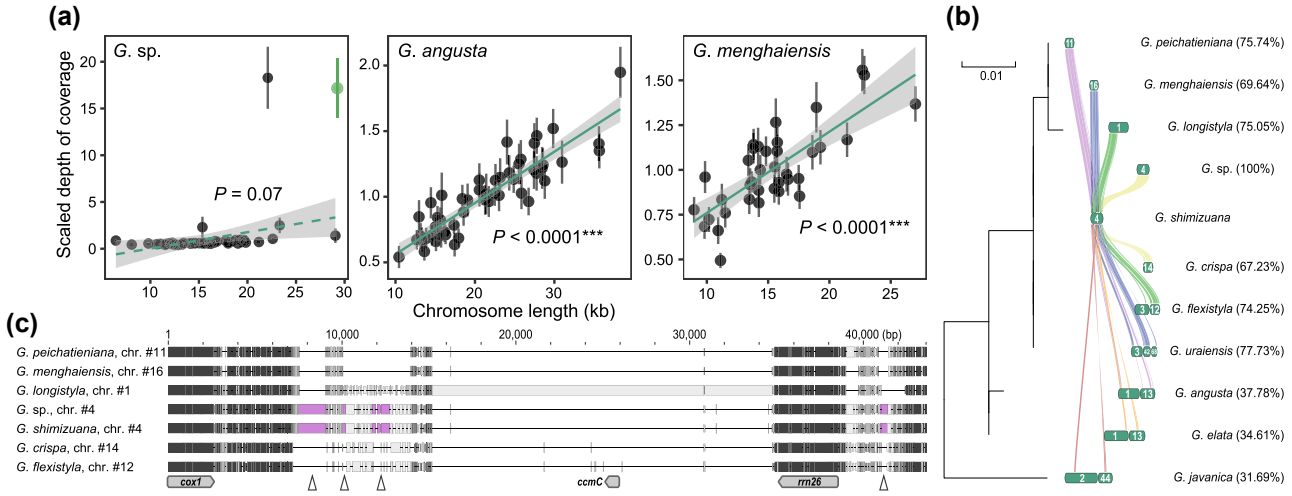

Fig. 2. Chromosomal copy number variation in *Gastrodia* mitogenomes. a) Chromosomes (or contigs) in assembled mitogenomes show two distinct patterns of copy number—length relationships: for *G.* sp. and *G. shimizuana* (exemplified by the former), a chromosome of ca. 15× average abundance is discovered, exceeding the copy number of conspecific plastomes (green dot in the leftmost panel); for most other assemblies (exemplified by *G. angusta* and *G. menghaiensis*), there is a significant positive correlation (all $P < 0.0001$) between chromosomal copy number and length. Scaled depths are shown so that the copy number of all chromosomes in a mitogenome averages to 1; lines with shaded areas denote fitted relationships with 95% confidence intervals. See supplementary fig. S6, Supplementary Material online for more details. b) Synteny between the hyper-abundant mitochondrial chromosome (chr. #4 of *G.* sp. and *G. shimizuana*) and other *Gastrodia* mitogenomes. Only chromosomes with >1,000-bp homologous sequence with #4 of *G. shimizuana* are shown. Bands denote homologous sequences >500 bp in length. The proportional length of *G. shimizuana* chromosome #4 covered by homologous sequences in each assembly is provided in brackets following taxa names. Branch length of *cox1* (the only protein-encoding gene in the hyper-abundant chromosomes) synonymous substitution rate is provided on the left. c) Alignment of the hyper-abundant chromosome with its homologues, with purple blocks highlighting regions unique to *G.* sp. and *G. shimizuana*; locations of protein-encoding genes, rRNA genes, and the four unique regions are indicated at the bottom row; the darkness of alignment positions is assigned based on nucleotide similarity across taxa (same for Fig. 4b).

Supplementary Material online). Significant differences in chromosome-specific copy numbers are detected for all assemblies (ANOVA tests; all $P < 0.001$). In general, the range of intragenomic chromosomal copy number variation is around three-fold, and larger chromosomes tend to have higher copy numbers (Fig. 2a and supplementary fig. S6, Supplementary Material online). Mitochondrial genes appear to distribute stochastically on chromosomes of varying abundance, although several tend to occur in clusters, such as *atp1-ccmFn*, *matR-nad1* exon3, and *atp8-nad4L-atp4* (supplementary table S2, Supplementary Material online). In particular, fragments of all three genes subject to *trans*-splicing (*nad1*, *nad2*, and *nad5*) are always found on different chromosomes (except in *E. aphyllum*; supplementary table S2,

Supplementary Material online), which differ in copy number by up to three folds (supplementary fig. S6, Supplementary Material online). Notably, however, a 22.6-kb chromosome in the mitogenomes of two closely related species, *G.* sp. and *G. shimizuana*, exists at an abundance of ca. 15 × higher than intragenomic averages (Fig. 2a and supplementary fig. S6, Supplementary Material online), and 45 and 25 × higher, respectively, than the least abundant chromosome. The sequencing depth of this hyper-abundant mitochondrial chromosome slightly exceeds that of plastome in the same individual (Fig. 2a and supplementary fig. S6, Supplementary Material online). The composition of mitochondrial chromosome abundance is largely consistent between two individuals of *G.* sp., as revealed by relative qPCR (supplementary fig. S7, Supplementary Material online).

Identified mitochondrial features of the 22.6-kb chromosome include *cox1*, 26S rRNA (*rrn26*), and a tRNA (*trnfM-CAU*) gene, all of which show ordinary levels of divergence from their homologues (Fig. 2b). To unravel the mechanisms underlying such abundance, we focused on its homology with congeners. All other members of *Gastrodia* Clade 3 possess a mitochondrial chromosome that overlaps >60% of its length (in terms of homologous sequences; Fig. 2b) but exhibiting only modest copy numbers (Fig. 2a and supplementary fig. S6, Supplementary Material online). While <35% of its length, which consists almost solely of *cox1* and *rrn26*, is covered by homologous sequences in the three mitogenomes of Clade 1 and 2. Through sequence alignment between the hyper-abundant chromosome and its homologues, we identified four unique regions (>100 bp; Fig. 2c), each featured with multiple stem-loops as indicated by folding simulations (supplementary fig. S8, Supplementary Material online), but not containing gene fragments or insertions from fungal or plastid genomes.

Mitochondrial Sequence Components of Different Sources

Intracellular plastid insertions (intra-MTPTs) are very rare among *Gastrodia* mitogenomes overall, with the richest (also the highest proportional coverage) being *G. crispa* (4,111 bp; 0.42%). Notably, no intra-MTPTs were found in half of the analyzed *Gastrodia* species, which form a subclade within Clade 3 (Fig. 3a). The intra-MTPT content of other mycoheterotrophic genera, *Epipogium* (ca. 7 kb; 1.6%), though exceeding those of *Gastrodia*, falls behind those of their autotrophic relatives by nearly an order of magnitude (24,625 to 67,243 bp; 3.66% to 9.95%) (supplementary table S1, Supplementary Material online). Plastid gene sequences are denser in intra-MTPT blocks compared to plastomes, especially for *Gastrodia*, where rRNA gene fragments may account for >90%. In some cases, intra-MTPT blocks include entire plastid genes and end near their boundaries, but no trace of splicing is found, while multiple plastid introns are observed (supplementary table S3, Supplementary Material online). Based on 589 repeat pairs assessed (124 for autotrophs and 465 for heterotrophs), most of them are recombinationally inactive, although few exceptions exist which mediate frequent recombination. Overall, recombination ratios in autotrophic and mycoheterotrophic species follow similar distributions (Fig. 3a). Copy numbers of plastomes relative to mitogenomes range by 3.5 folds among mycoheterotrophic orchids, from 5.3 in *G. menghaiensis* and *G. crispa* to 18.7 in *G. elata*, and that of the autotroph *N.*. *mackinnonii* (10.3) is intermediate among them (Fig. 3b).

Analyzing another 19 published mitogenomes of fully heterotrophic plants representing eight major lineages produced a similar result of reduced intra-MTPTs in both parasitic and mycoheterotrophic plants. Despite differences among lineages, intra-MTPT contents in analyzed heterotroph mitogenomes are consistently lower than in their autotrophic relatives, measured either in absolute amount or proportional coverage of mitogenomes, except the mycoheterotroph *Hypopitys monotropa* (Ericaceae: Monotropoideae) which has a similar intra-MTPT content as its autotrophic relatives (Fig. 3a and supplementary table S4, Supplementary Material online). A parallel reduction in plastome size is also found across these lineages, where those of the heterotrophic members are reduced to up to two-third the size of their autotrophic relatives (Fig. 3a and supplementary table S4, Supplementary Material online). Searching orchid mitogenomes against those of fungi revealed 4,931 to 13,523 bp of fungal insertions that accounts for 0.25% to 3.10% the size of each mitogenome, with those of *Gastrodia* being generally around 6,000 bp (Fig. 3c). Analyzed mitogenomes of mycoheterotrophic orchid species do not contain more fungal insertions than those of autotrophic ones. On the contrary, fungal insertions in the *Dendrobium brymerianum* mitogenome (13,525 bp) are over twice more abundant than in the mycoheterotrophs (Fig. 3c and supplementary table S1, Supplementary Material online).

Further investigation on mitochondrial sequence origins focused on *G. elata* and *G. menghaiensis*, using nuclear genome assemblies available. The two mitogenomes differ in size by 2.6 folds (1.51 Mb vs. 580 kb), while sharing only ca. 150-kb syntenic blocks (Fig. 3d). We found plentiful BLAST hits between both mitogenomes and their corresponding nuclear assemblies. These involve 45.6% (687 kb) of the mitogenome of *G. elata* and 32.6% (189 kb) of *G. menghaiensis*, overlapping many non-syntenic regions of the two mitogenomes (Fig. 3d). We tested this using the other nuclear genome assembly for *G. elata* and got a similar proportion (59.3%; supplementary fig. S9, Supplementary Material online). In particular, the ca. 300-bp insertions into the *rps3* gene contain BLAST hits with the nuclear genome assembly in both species (supplementary fig. S4, Supplementary Material online), suggesting a possible nuclear origin of this *Gastrodia*-specific insertion. Still, eight chromosomes (six in *G. elata* and two in *G. menghaiensis*) are left largely uncharacterized with respect to sequence origins. Further analyses unveiled six blocks (9,213 bp in total) present in five of these chromosomes that show strong signs of horizontal transfer, which are verified by PCR or PacBio sequencing reads (supplementary fig. S10 and tables S5 to S9, Supplementary Material online). Five of the six blocks are likely derived from mitochondrial intergenic regions of different angiosperm donors, whereas the other one—a ca. 300-bp fragment at *G. elata* chromosome 49—traces its origin to the near 3′ end of the *psbC* gene of Malvales plastomes (supplementary fig. S10, Supplementary Material online). Online BLAST revealed the absence of homologues of these horizontal gene transfer (HGT) blocks in other assembled orchid mitogenomes (supplementary fig. S10, Supplementary Material online).

Sequence Evolution

Maximum likelihood phylogeny was reconstructed for selected orchid species plus a broad range of angiosperms based

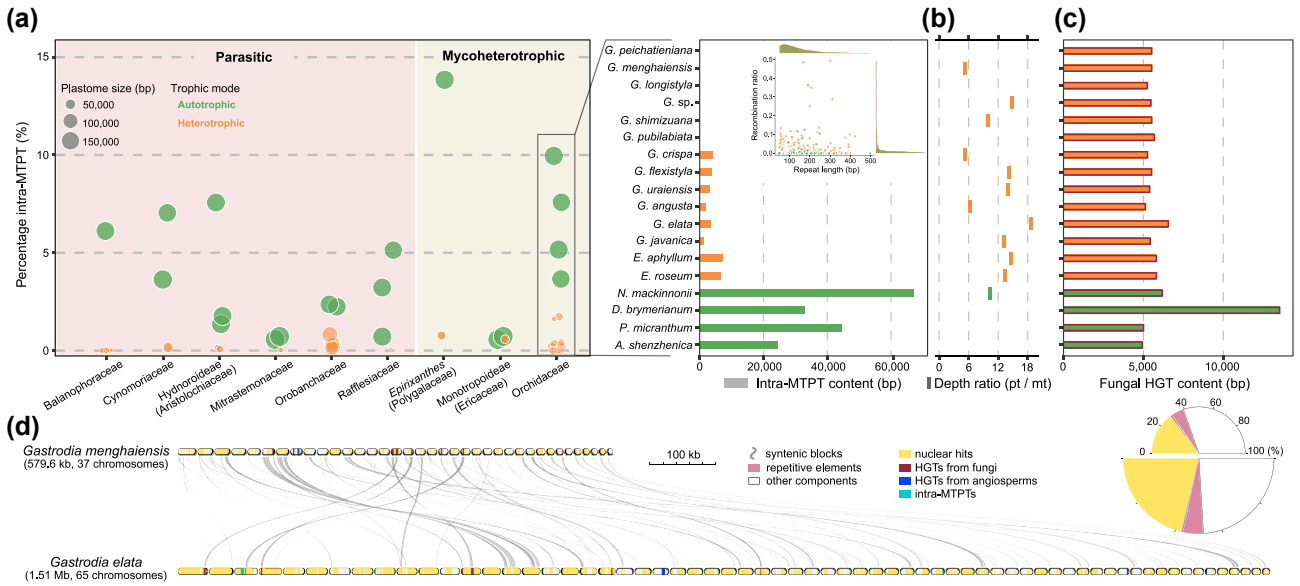

Fig. 3. Unusual collections of transferred sequences in *Gastrodia* mitogenomes. a) Intra-MTPT contents in published heterotrophic plant mitogenomes versus their autotrophic relatives, with a focus on Orchidaceae. Each species is represented by a dot, with dot size scaled by plastome size (except for members of Rafflesiaceae due to plastome loss) and dot color denoting the trophic mode; note that one lineage may involve multiple origins of heterotrophy (Merckx and Freudenstein 2010). Inset at top-right: repeat-mediated homologous recombination ratio versus repeat length in analyzed orchid mitogenomes, with kernel distribution of the two variables drawn at the top and right margins. See supplementary tables S3 and S4, Supplementary Material online for detailed information; *E.*: *Epipogium*, *N.*: *Nervilia*, *D.*: *Dendrobium*, *P.*: *Paphiopedilum*, *A.*: *Apostasia* (same for Fig. 4a). b) Illumina sequencing depth of plastomes relative to mitogenomes. *Gastrodia peichatieniana* and *G. longistyla* were excluded due to large uncertainty in plastome depth (standard deviation > mean), while *G. pubilabiata* and three autotrophic species were not analyzed due to unavailability of required sequencing data (see Methods). c) Contents of fungi-derived HGTs in orchid mitogenomes. d) Distribution of identified sequence components in mitogenomes of *G. menghaiensis* and *G. elata*. Pie chart on the right summarizes the proportional coverage of each category (upper half: *G. menghaiensis*; lower half: *G. elata*). For sites involving more than one category, we computed its contribution as $1/n$ nucleotide to each category, where n is the number of categories involved; the size of the pie is scaled by sequence length under each category. "Nuclear hits" denote sequences with BLAST hits with the nuclear genome; "repetitive contents" are shown only in the pie chart; see supplementary figs. S9 and S10 and table S1, Supplementary Material online for related information.

on the coding sequence of each mitochondrial protein-encoding gene (supplementary table S10, Supplementary Material online). In all cases, *Gastrodia* species form a single cluster, showing no signs of horizontal gene transfer. Duplicate gene fragments in some assemblies are always clustered with the intact copy from the same mitogenome (supplementary fig. S11, Supplementary Material online). The two genes with duplicated copies in *N. mackinnonii*, by contrast, each has a native copy and the other clustered with distant eudicots, with that of *ccmC* having moderate support (BS = 0.54) and that of *nad6* strongly supported (BS > 0.90) to originate from rosids (supplementary fig. S11, Supplementary Material online). Gene-specific absolute synonymous (d_S) and non-synonymous (d_N) substitution rates (substitutions per million years) are similar among *Gastrodia* and *Epipogium*, with the medians of *G. javanica* (Clade 1) and *E. aphyllum* slightly exceeding the others, and members of *Gastrodia* Clade 2 being the overall lowest among mycoheterotrophs. Analyzed autotrophic orchids, in general, have much lower substitution rates (2 to 10 folds) than mycoheterotrophic ones (Fig. 4a). Among the 38 mitochondrial protein-encoding genes, mycoheterotrophs peak in the rate of all but one (*nad3*). Such accelerations are less marked when put in the context of the angiosperm phylogeny (supplementary fig. S12, Supplementary Material online), but for numerous protein-encoding genes (*atp4*, *cox3*, *matR*, *mttB*, *nad2*, *nad4*, *nad7*, *rpl5*, *rps1*, *rps2*, *rps3*, *rps7*, *rps11*, *rps12*, *rps14*, and *rps19*), *Gastrodia* is still among lineages with highest substitution rates (supplementary fig. S12, Supplementary Material online). Notably, the ca. 300-bp insertion unique to the *Gastrodia rps3* gene belongs to a prolonged second exon, instead of a novel intron, as it is captured by long non-coding RNA (lncRNA) sequencing (Fig. 4b). The insertion length is always divisible by 3 bp and does not cause frame shifts (supplementary fig. S4, Supplementary Material online). Three-dimensional structural predictions for *Gastrodia* mitochondrial ribosome protein S3 (product of the mitochondrial *rps3* gene) do not produce confident results, with all predicted template modeling scores being around 0.40, and per-atom confidence estimates corresponding to the insertions mostly below 0.50 (supplementary fig. S13, Supplementary Material online). Linking the absolute d_S of each gene with the relative copy number of the chromosome encoding it revealed no correlation between the two (Fig. 4c), which is also the case for different fragments of the three genes subject to *trans*-splicing (supplementary fig. S14, Supplementary Material online). Using published nuclear genome assemblies, the relative mitogenome copy number was estimated as 19.3 and 33.7 for *G. elata* and *G. menghaiensis*, respectively (Fig. 4d).

Based on ancestral sequence reconstruction and inference with syntenic blocks, many of conserved mitochondrial chromosomes have undergone rearrangement events, through which more recently diverged species tend to possess smaller chromosomes. After rearrangements, the evolutionary trajectories of these chromosomes were dominated by nucleotide deletions, although insertions and single nucleotide polymorphisms (SNPs) were also very common. Several of these

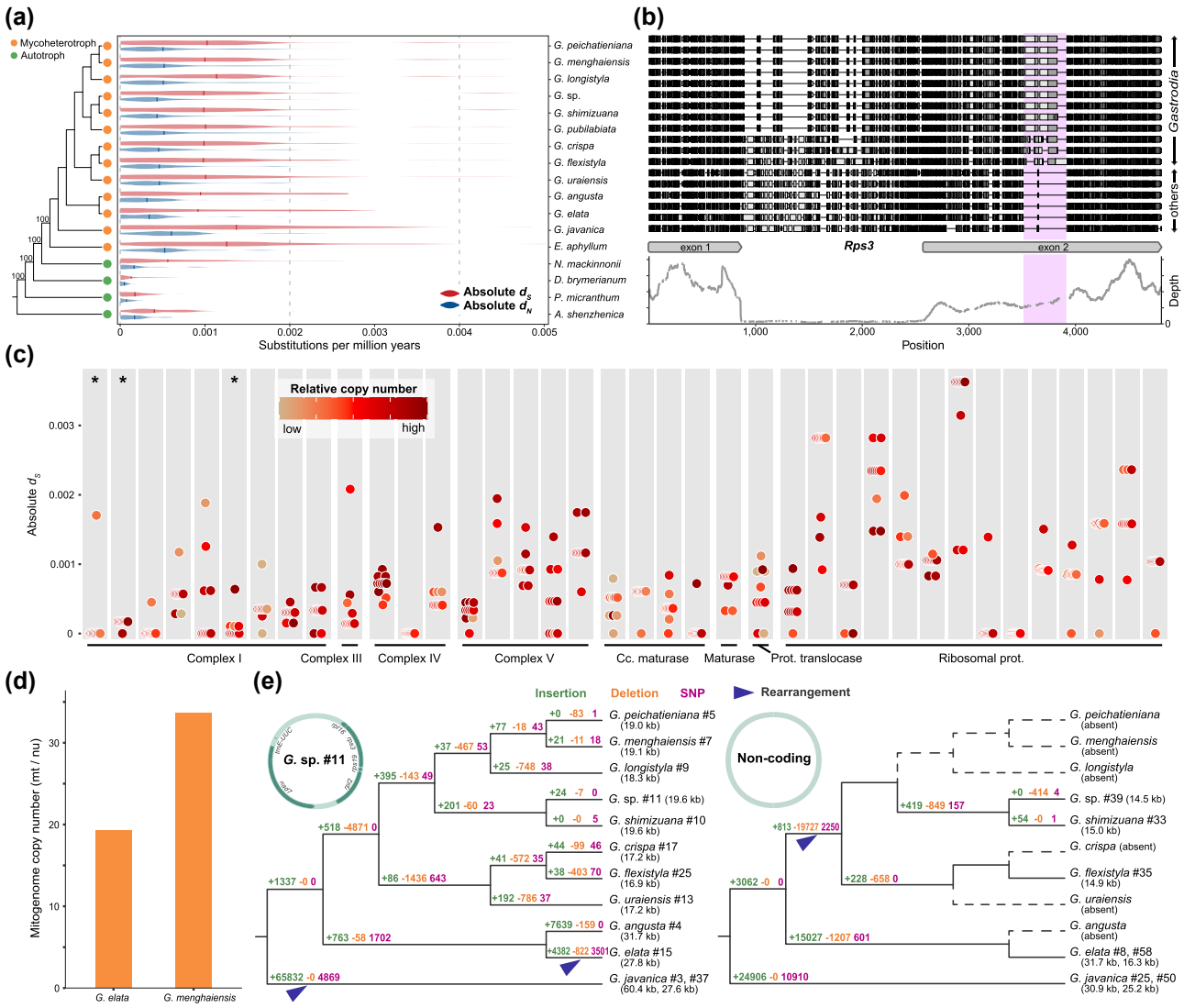

Fig. 4. Mitochondrial sequence evolution in *Gastrodia*. a) Absolute substitution rates of protein-encoding genes in selected orchid mitogenomes. Shaded curves and vertical bars denote kernel distribution of rates and the estimated median rate in each species, respectively; phylogenetic relationships among selected taxa are shown on the left, with percentage bootstrap support labeled for nodes not present in Fig. 1a. b) Alignment of *rps3* gene, highlighting the ca. 300-bp insertion unique to *Gastrodia* (area with purple shading); order of taxa follows panel a, except that the bottom row refers to *Liriodendron tulipifera*, selected for its highly conserved mitochondrial coding sequences (Richardson et al. 2013); see supplementary fig. S4, Supplementary Material online for related information. The curve below the alignment plot denotes lncRNA sequencing depth across the *rps3* gene of *G. elata*, which suggests a prolonged coding sequence due to the insertion. c) Gene-specific absolute d_S in *Gastrodia* mitogenomes and association with copy number. Genes are separated by white borders, and clustered by functional category (solid lines at the bottom). Copy numbers are based on Illumina sequencing depths of chromosomes (or contigs) encoding genes of interest, which are scaled to [0, 1] for each mitogenome and visualized with color densities of dots. For the three genes subject to *trans*-splicing (labeled with asterisks), the longest fragment is shown (see supplementary fig. S14, Supplementary Material online for details). Cc.: cytochrome c, prot.: protein. d) Copy number of mitochondrial relative to nuclear genomes in two *Gastrodia* species. e) Evolutionary trajectory of *Gastrodia* mitochondrial chromosomes, with examples of a gene-dense (left) and a non-coding (right) one. Circles at top-left show encoded genes.

chromosomes, which encode no or only tRNA genes, are entirely lost in some species, while those with protein-encoding ones are retained (Fig. 4e and supplementary fig. S15, Supplementary Material online).

Discussion

We present a comparative mitogenomic study for fully mycoheterotrophic plants based on the genus *Gastrodia* and its relatives. These mitogenomes are characterized by a 5-fold size variation with limited homology, peculiar architectures with unequal chromosome abundance, unusual collections of transferred sequences, and elevated evolutionary rates of coding sequences. Placing these findings into the context of angiosperm (heterotroph in particular) organelle genomes, we highlight the diversity of genome-level consequences associated with heterotrophy and demonstrate how they are linked to organelle- or habit-specific characteristics, promoting mechanistic understanding of these peculiar genomic features. We also discuss the potential trajectory of genome size evolution in relation to evolutionary rates.

The Multipartite Architecture Is Likely Associated With Unequally Abundant Mitochondrial Chromosomes

Multi-chromosomality is presumably prevalent among plant mitogenomes, and those with unequally abundant chromosomes

have been reported in distant angiosperms (Alverson et al. 2011; Wu et al. 2015) and other phyla of life (Suga et al. 2008; Phillips et al. 2016). In such mitogenomes, variations of chromosomal copy number are around several folds and correlated with chromosome length, with larger chromosomes usually being more abundant (Alverson et al. 2011; Wu et al. 2015). Although analyzed mitogenomes exist in vivo likely as a diverse collection of alternative conformations (Backert et al. 1997; Gualberto and Newton 2017; Kozik et al. 2019; He et al. 2023), marked intragenomic sequencing depth variations unequivocally point to multipartite mitogenomic architectures. That most of our assemblies exhibit intragenomic copy number variation (2.2 to 6.7 folds) positively correlated with chromosome length (supplementary fig. S6, Supplementary Material online) suggests the potential prevalence of unequal mitochondrial chromosome abundance. Interestingly, the up-to-three-fold copy number difference within a single gene (supplementary fig. S6, Supplementary Material online) implies either post-transcriptional regulations or that gene copy numbers do not correlate proportionally with total transcription rates, to attain proper stoichiometry of messenger RNAs (mRNAs). Although our assemblies were based on flower stalk samples, the overall stable composition of mitochondrial DNA (Zwonitzer et al. 2024; supplementary fig. S7, Supplementary Material online) supports the generality of these findings.

Of particular interest is the hyper-abundant chromosome in *G.* sp. and *G. shimizuana* whose copy numbers likely exceed conspecific plastomes (Fig. 2a), leading to by far the most extreme chromosomal copy number variation observed in any mitogenomes, to our knowledge. It is possibly more abundant than mitochondrial plasmids as well (Palmer et al. 1983), but this has to be tempered by the current unavailability of quantitative data on plasmids inside *Gastrodia* mitogenomes. Assembly reliability of this chromosome is justified by distinct mitochondrial features (*cox1* and *rrn26* genes) and consistently high sequencing depths (supplementary fig. S1, Supplementary Material online). To provide clues into such abundance, we focused on its unique regions (Fig. 2a and supplementary fig. S6, Supplementary Material online). The predicted secondary structures, especially the multiple stem-loops (supplementary fig. S8, Supplementary Material online), suggest their possible functions in chromosome replication (Pearson et al. 1996; Yu et al. 2022), although additional evidence is needed. Alternatively, future genomic or transcriptomic data may reveal whether such copy number variation is due to certain aspects of the nuclear genome, such as genes pertaining to DNA replication, repair, and recombination (Yu et al. 2022). A similar architecture is discovered in two basal metazoan lineages, the rotifer *Brachionus plicatilis* and the cyst nematode *Globodera ellingtonae*, with their mitogenomes both consisting of two similar-sized chromosomes that vary in abundance by more than four folds (Suga et al. 2008; Phillips et al. 2016). Notably, all these hyper-abundant chromosomes encode the large ribosomal RNA subunit. It is therefore intriguing whether such convergence (if not coincidental) is related to differential gene expression—a commonly suggested advantage of genomic fragmentation (Rand 2009).

Transferred Mitochondrial Sequences Highlight Asymmetric Evolutionary Consequences of Heterotrophy

Mixtures of foreign sequences are common components of angiosperm mitogenomes (Goremykin et al. 2009; Cusimano and Wicke 2016; Gandini and Sanchez-Puerta 2017). Intracellular insertions from plastid genomes, specifically, are almost universally present, often accounting for 1% to 10% of the size of mitogenomes (Mower et al. 2012a). All analyzed *Gastrodia* mitogenomes harbor limited intra-MTPT contents (Fig. 3a and supplementary table S3, Supplementary Material online) that contribute to <1% of each mitogenome, and several members of Clade 3, together with some species in the holoparasitic family Balanophoraceae, represent the only known cases in angiosperms where intra-MTPTs are completely absent (supplementary table S4, Supplementary Material online). Acquisition of foreign sequences requires the release of genetic materials from the donor and transmission to the recipient (Mower et al. 2012a), and subsequent incorporation into the recipient genome may be facilitated by homology-mediated recombination. Differential recombination activity, however, is not a likely cause for the paucity of transferred sequences in analyzed mycoheterotrophic orchids (Fig. 3a). Given that plastomes of *Gastrodia* are reduced to approximately one-fifth the size of typical angiosperms' and that intracellular transfers of plastid sequences into *Gastrodia* mitogenomes are likely ancient (Wen et al. 2022; Kim et al. 2023), we propose the contracted plastome hypothesis, where reduction of plastome has led to fewer cases of plastid sequence release and lowered chance of interaction with conspecific mitogenomes. On the other hand, since they hardly produce selective impacts (Mower et al. 2012a; Gandini and Sanchez-Puerta 2017), ancient plastid insertions (when plastomes were less reduced) were gradually eroded by mutations and sequence turnover (Fig. 1). As support, we found clear evidence of intra-MTPT content and plastome size reduction across independent heterotroph lineages as compared to their autotrophic relatives (Fig. 3a and supplementary table S4, Supplementary Material online). Despite variations among individual lineages, this pattern is largely parallel between the two heterotrophic types (Fig. 3a). An alternative explanation, however, concerns plastome expression, if intra-MTPTs were transferred mainly as complementary DNAs (RNAs after retroprocessing), and mycoheterotrophs have reduced expression of plastid genes. While the former is partly supported by high densities of plastid gene sequences inside intra-MTPTs (supplementary table S3, Supplementary Material online), the contribution of such is presumably limited for autotrophic orchids, as no trace of intron splicing is found. Together with a lack of evidence suggesting reduced expression of mycoheterotroph plastomes, this explanation remains ungrounded at present. Given the typically considerable contribution of plastid insertions (Mower et al. 2012a), the contracted plastome hypothesis proposed here may reveal a mechanism through which content and size diversity of angiosperm mitogenomes arose.

Our analyses on horizontal gene transfers, in contrast, unveiled disparate mitogenomic evolution between the two heterotrophic types. Studies on parasitic plants have reported massive mitochondrial sequence acquisition from their hosts, frequently causing functional replacement of native genes (Xi et al. 2013; Sanchez-Puerta et al. 2017). This is partly due to the frequent fusion and fission of plant mitochondria (Arimura et al. 2004; Davis and Xi 2015) and physical contacts that facilitate sequence transmission (Davis and Wurdack 2004; Mower et al. 2010; Lin et al. 2022a). Surprisingly, contents of fungal insertions in *Gastrodia* mitogenomes are quite scarce, only similar to or falling behind that of their autotrophic relatives (Fig. 3c and supplementary table S1, Supplementary Material online), implying that complete nutrient reliance on fungi does not cause

more sequence transfers into mitogenomes, consistent with prior analyses on other mycoheterotrophs (Petersen et al. 2019; Shtratnikoya et al. 2020). Despite the symbiotic relationship, integration of fungal sequences into mycoheterotroph mitogenomes may have been curbed by genetic barriers (but for symbiosis at early life stages that explains the presence of fungal sequences in most extant orchid mitogenomes; see Huang 2013, Sinn and Barrett 2020) due to organellar fusion incompatibility and different organellar division machineries (Richards et al. 2009; Westermann 2010; Rice et al. 2013). In addition, we discovered horizontal origins for duplicated gene copies in *N. mackinnonii* and several blocks of non-coding sequences in *Gastrodia* mitogenomes (Fig. 3d and supplementary figs. S10 and S11, Supplementary Material online). Limited in amount but pointing to diverse donors, these HGTs again sharply contrast those in parasitic plants, where plenty of them mostly originate from a single donor (the host plant) and frequently carry entire genes (e.g. Sanchez-Puerta et al. 2017). Since members of *Gastrodia* and *Nervilia* are not known to form direct contact with other living plants, mechanisms underlying the acquisition of these sequences remain elusive, suggesting potentially unidentified pathways through which mitochondrial HGTs can occur.

Evolutionary Rate Acceleration and Insights into Mitogenome Size Variation

It is generally established that plant mitogenomes evolve rapidly in structure but slowly in sequence (Palmer and Herbon 1988; Cho et al. 2004; Christensen 2021). The repertoire of 38 protein-encoding and three rRNA genes is largely conserved among assembled orchid mitogenomes (e.g. Wang et al. 2023; Yang et al. 2023), although reduced tRNA contents in *Epipogium* and *Gastrodia* mitogenomes indicate a stronger reliance on nucleus-encoded translational apparatus for mycoheterotrophs (supplementary fig. S5, Supplementary Material online) (Schneider 2011). Substitution rates of mitochondrial protein-encoding sequences, however, have undergone marked acceleration in the two genera representing likely independent origins of mycoheterotrophy (Fig. 4a) (Zhang et al. 2023). Even when placed into the angiosperm phylogeny, some genes of *Gastrodia* evolve faster than most other taxa (supplementary fig. S12, Supplementary Material online), as further entrenched by the unique insertion into the coding sequence of their *rps3* gene (Fig. 4b). Across parasitic plant genomes, evolutionary rate increase has been broadly attributed to higher mutation rates and relaxed selective pressures (Bromham et al. 2013; Skippington et al. 2015). Our results, together with studies on nuclear and plastid genomes (e.g. Yuan et al. 2018; Petersen et al. 2019; Wen et al. 2022), extend such acceleration to heterotrophic plants, although future expanded sampling may secure this extrapolation. Rate increase in mycoheterotrophs may be similarly attributed to mutation rates and selective pressures, given their high absolute d_S and d_N/d_S ratios (Fig. 4a). Different from what is typical for angiosperms with extremely fast-evolving mitogenomes (Zwonitzer et al. 2024), however, copy number does not explain rate variation in *Gastrodia* mitogenomes (Fig. 4c and supplementary fig. S14, Supplementary Material online), likely because mitochondrial DNA is overall abundant in this genus to enable high-fidelity replication (Fig. 4d; compare to fig. 2 in Zwonitzer et al. 2024). Factors such as mutagenic local environments and shorter replication cycle of mitochondria are thus candidate drivers for rate acceleration in *Gastrodia* (Birky 2001; Lynch 2007, 2010). Future information on biochemistry and other aspects may illuminate whether the convergent mitochondrial evolutionary rate acceleration across heterotrophic plants is driven by shared mechanisms.

Across organisms, high mutation rates tend to be associated with smaller genome sizes, given the mostly deleterious nature of mutations (the mutational hazard hypothesis; see Smith 2016 and Lynch et al. 2023 for details). Organelle genomes, in particular, are not expected to be exceptions given their inability to shed mutations by recombination (Lynch 2007). Most heterotroph plastomes analyzed to date, including those of *Gastrodia* (Wen et al. 2022), are characterized by parallel size contraction and evolutionary rate acceleration (reviewed in Graham et al. 2017). In contrast, the size of our analyzed *Gastrodia* mitogenomes spans nearly four folds (Fig. 1a), with many situated well above the typical range of angiosperms' (Wu et al. 2022). This finding, along with that many known fast-evolving angiosperm mitogenomes are unusually large (e.g. Sloan et al. 2012; Zwonitzer et al. 2024; but also see Skippington et al. 2015), appear to contradict the mutational hazard hypothesis. Mitochondrial sequence evolution in angiosperms is generally slow, particularly so when compared to bilateral animals (Lynch 2007). Besides, size variation in angiosperm mitogenomes is mostly due to non-coding (and possibly nonfunctional) components (Mower et al. 2012a). These factors render their expansion less a "mutational liability", but more consequences of sequence birth-death processes (Lynch 2007). However, it does not imply that such processes occurred completely at random. In the context of decreasing mitogenome size from early- to recently-diverged *Gastrodia* species, ancient rearrangements might have generated numerous, fragmented (smaller) chromosomes, with genes distributed to some of them and leaving others "empty". Then, both underwent deletion-inclined mutations, but only ones that encode no or only "dispensable" genes were subject to stochastic loss, presumably due to relaxed selection on their persistence. This fits with the similarly large number of chromosomes with decreasing lengths across Clades 1, 2, and early-diverged species of Clade 3, and decreasing number but similar lengths from early- to recently-diverged Clade 3 members (Table 1 and supplementary fig. S2, Supplementary Material online). This scenario adds details to mitogenome size evolution following a putative "big-bang" (Wu et al. 2015), although the initial phase of this hypothesis—the massive expansion—warrants further evidence.

Conclusions

Our analyses on *Gastrodia* and relatives provide a glimpse into the mitogenomic peculiarities of mycoheterotrophic plants, highlighting the diverse and remarkable genomic outcomes associated with heterotrophy. The multiple aspects of convergent and divergent mitogenome evolution between parasitic and mycoheterotrophic plants uncovered here advance our understanding of how trophic habit can impact the evolutionary trajectory of genomes. We anticipate that these discoveries, instead of only entrenching existing views on the extreme variability of plant mitogenomes, should encourage further scientific endeavors, especially those that employ interdisciplinary approaches, to explore currently elusive aspects of plant mitogenomes such as the evolution of peculiar architectures.

Materials and Methods

Sampling and Sequencing

Living plant materials at flowering time were collected from the wild following national and regional legislations. These

included ten representatives of *Gastrodia* (subfamily Epidendroideae, tribe Gastrodieae) and two species from related genera, *Epipogium* and *Nervilia* (mycoheterotrophic and autotrophic, respectively; both in subfamily Epidendroideae, tribe Nervilieae). For *G.* sp., two individuals were sampled. Voucher specimens were deposited at the Herbarium of the Institute of Botany, the Chinese Academy of Sciences (PE). We extracted total DNA from peduncles with DNAsecure Plant Kits (TIANGEN BIOTECH, Beijing, China) following the manufacturer's protocols. Products were fragmented with Covaris LE220R-plus (Woburn, USA), end-polished, A-tailed, ligated with adaptor, and PCR-amplified, to construct shotgun genomic libraries of 450-bp (for *Epipogium* only) or 350-bp (for others) insert sizes, which were sequenced on Illumina NovaSeq6000 platforms (Novogene, Beijing, China), producing 100-bp (for *Epipogium*) or 150-bp (for others) pair-end sequence reads that totaled ca. 10 Gb for each specimen. A genomic library with 15-kb insert size was also constructed for *G. peichatieniana* and sequenced on a PacBio RS-II Platform, producing ca. 20 Gb long reads. For *G. elata*, we downloaded DNA sequencing reads available at the NCBI SRA repository (BioProject No.: PRJNA632604) and extracted two subsets of 40 Gb (Illumina) and 30 Gb (PacBio), respectively, for subsequent use. Additional lncRNA sequencing of fresh peduncle materials of *G. elata* was performed by Benagen (Wuhan, China) as follows. Total RNA was extracted using TRIzol, the product of which was measured for ratio of absorbance at 260/280 nm and 260/230 nm wavelengths with NanoPhotometer (Implen, Munich, Germany) to assess the level of contamination. The qualified sample was used for library construction with RNA Library Prep Kit following manufacturer's instructions, through which rRNAs were removed by probes and lncRNAs were enriched. Fragmentation by divalent cations was performed in 5X First Strand Synthesis Reaction Buffer, followed by complementary DNA synthesis, NEBNext Adaptor ligation, PCR amplification, and sequencing on Illumina NovaSeq6000 platform to produce ca. 15 Gb pair-end reads. Trimmomatic v0.39 (Bolger et al. 2014) and Canu v2.2 (Koren et al. 2017) were used, both with default settings, to filter or self-correct raw Illumina and PacBio sequencing reads prior to downstream analyses.

Mitogenome Assembly, Annotation, and Search of Syntenic Blocks

For *G. elata* and *G. peichatieniana* where both Illumina and PacBio reads were available, we followed Yu et al. (2022) for mitogenome assembly. For the other species, assembly was performed following Yu et al. (2021). A custom mitogenomic database (supplementary table S10, Supplementary Material online) was used as a reference for the GetOrganelle pipeline (Jin et al. 2020). The last step of both workflows was to disentangle a network of contigs bordered by repetitive sequences when the output GFA file was visualized using Bandage v0.8.1 (Wick et al. 2015; Yu et al. 2022 for detailed descriptions). To minimize subjective impacts, for each repeat pair shorter than the insert size located at joints of contigs, we constructed all (in most cases, two) alternative conformations that included the repeat itself and sequences flanking both sides and mapped Illumina reads to these conformations with Bowtie2 v2.3.4 (Langmead and Salzberg 2012). We then disentangled these contigs according to the dominant conformation as determined by the number of sequencing reads supporting each conformation. For repeats exceeding the insert size (except in *G. elata* and *G. peichatieniana*, where we determined the dominant conformation by a PacBio reads-mapping method similar to Illumina reads), we followed the convention to assemble as multiple possible subcircular chromosomes instead of a "master circle" (Wu et al. 2015). To assess assembly completeness, we calculated the sequencing depth of all assembled mitochondrial sequences with Bowtie2 and Samtools v1.11 (Danecek et al. 2021) (see below). We then re-assembled Illumina reads with SPAdes v3.13.0 (Prjibelski et al. 2020) under a broad range of *k*-mers (57 to 127, whenever applicable), and extracted contigs with similar sequencing depths as assembled mitochondrial sequences. Extracted contigs were compared against assembled mitogenomes using BLASTN (Zhang et al. 2000) to identify potentially missing mitochondrial contigs. The assembled mitogenome of *G. pubilabiata* was downloaded from NCBI GenBank and used in subsequent analyses except those requiring chromosome delimitation, as this assembly was intentionally assembled as a master circle (Kim et al. 2023). Four previously published orchid mitogenomes (three autotrophic and one mycoheterotrophic species) outside of *Gastrodia* were also analyzed (supplementary table S1, Supplementary Material online).

To annotate protein-encoding and rRNA genes, we combined the use of BLASTN and Live Annotate & Predict function implemented in Geneious Prime v2023.1.2 (https://www.geneious.com) against a custom database of plant mitogenomes (supplementary table S10, Supplementary Material online), both with default parameter settings. We additionally searched for unannotated protein-encoding genes with the mitogenome of *Liriodendron tulipifera* (Richardson et al. 2013) as reference under sensitive parameter settings of BLASTN (*-word_size 7 -reward 5 -penalty -4 -gapopen 8 -gapextend 6*), following Skippington et al. (2017). Due to the presence of a unique insertion into the *rps3* gene of *Gastrodia*, we mapped lncRNA sequencing reads of *G. elata* to its mitogenome assembly to determine whether this insertion is retained in mature mitochondrial mRNAs. Annotation of tRNAs was performed with the tRNAscan-SE web server under "other mitochondrial" mode (Lowe and Chan 2016). Finally, we ran MFannot (Lang et al. 2023) and ORFfinder (https://www.ncbi.nlm.nih.gov/orffinder) to check for unannotated genes or mis-annotations.

We searched for syntenic blocks among assembled mitogenomes using BLASTN, with an e-value cutoff of 1e-5, a percent identity of 80, and an alignment length of 250 bp. Syntenic blocks were visualized using custom R scripts adapted from the *RIdeogram* package (Hao et al. 2020).

Depth of Coverage and Copy Number Estimation

Coverage depth of assembled mitogenomes were computed with Illumina sequencing reads using Bowtie2 and Samtools (parameter settings: *–end-to-end –no-discordant –no-mixed*). All circular-mapping chromosomes were linearized to the required input format. Raw coverage data were further analyzed (see below) to produce chromosome (or contig)-specific copy number estimates. We first trimmed data within 500 bp to both ends of linearized chromosomes or those corresponding to intracellular plastid insertions (see below); then, we excluded outliers that fell below 0.05 or above 0.95 quantiles of coverage data for each chromosome. The mean and standard deviation of remaining data were computed as estimates of relative chromosome-specific copy number and associated

uncertainties. We then performed analysis of variance (ANOVA), treating the coverage depth of each nucleotide as individual observations, to test for differences in copy number of different chromosomes against the null hypothesis of equal means. As there were highly significant differences ($P < 0.0001$) in all assembled mitogenomes, we fit estimated copy numbers against chromosome lengths using linear regressions. The same sets of Illumina sequencing reads were also mapped to plastomes (Wen et al. 2022), for our sampled species and *Epipogium roseum* (based on SRA accession SRR1344935), as none of the other autotrophic species in Fig. 3A have publicly available Illumina sequencing data for peduncle DNA. This is for consistency, as plastome abundance can vary considerably across different parts of a plant (Zwonitzer et al. 2024). For *G. elata* and *G. menghaiensis*, sequencing depths of nuclear genomes were additionally estimated by dividing nucleotide count in Illumina sequencing data (read pairs mapping entirely to organelle genomes excluded) by corresponding haploid nuclear genome size reported in literatures (Xu et al. 2021; Jiang et al. 2022). Copy number of mitogenomes relative to plastid and nuclear genomes was then calculated as the ratio of mitogenome sequencing depth (averaged across chromosomes) to those of the latter.

We identified two chromosomes, the fourth of *G.* sp. and *G. shimizuana*, for their exceptional copy number of >15 × higher than intragenomic averages. These hyper-abundant chromosomes are nearly identical in sequence (99.97% overall identity), thus were analyzed as a whole. As all sampled *Gastrodia* species in Clade 3 possess a circular-mapping mitochondrial chromosome that covered >60% (in terms of homologous sequences) of the hyper-abundant one, we aligned these homologues with MAFFT v7.313 (Katoh and Standley 2013) under "normal" mode to identify regions unique to the hyper-abundant chromosome. For each unique region >100 bp, secondary structure was predicted with mfold web server under default settings (Zuker 2003), and palindromic sequences were identified using Palindrome analyser web server, by setting maximum spacer and mismatch to 10 and 5, respectively (Brazda et al. 2016).

Detection of Repeats and Transferred Sequences

We searched for repetitive contents by comparing each mitogenome against itself using BLASTN. Hits between nonidentical locations and with e-values ≤1e-7 were considered as repeat pairs. Transposable elements were annotated using CENSOR (Kohany et al. 2006). To identify transfers from fungi, we used the mitogenome assemblies of *Ustilago maydis* (Sinn and Barrett 2020), as well as *Armillaria* and *Mycena*—fungal taxa known to form symbiotic relationships with *Gastrodia*—as BLASTN references, and searched each assembly with parameters *-word_size 7 -evalue 1e-5 -reward 2 -penalty -3 -gapopen 3 -gapextend 3* (following Sinn and Barrett 2020).

Mitochondrial plastid insertions were detected using BLASTN, with an e-value cutoff of 1e-5, percent identity of 80, and alignment length of 250 bp, with the plastomes of four autotrophic orchids (*Apostasia shenzhenica*, *Dendrobium brymerianum*, *Nervilia mackinnonii*, and *Paphiopedilum micranthum*) as references. As we focus on plastid sequences that result from intracellular transfer (intra-MTPTs), BLAST hits were then filtered to exclude sequences that initially transferred from the plastomes of distant taxa (foreign-MTPTs) following Gandini and Sanchez-Puerta (2017). Intra-MTPT blocks were then screened for plastid gene fragments, based on their homologues in reference plastomes. To further explore the general pattern of intra-MTPTs in heterotrophic plants, we downloaded all currently available mitogenomes for fully heterotrophic plants from GenBank (data retrieved in May 2024) and performed the same workflow under identical parameter settings as specified above. For each lineage that contains heterotrophic members, we also obtained the mitogenomes of their autotrophic or facultatively parasitic (in the case of Santalales only; same for below) relatives (mostly taxa in the same family as the heterotroph, but this was not possible, for example, for Balanophoraceae and Rafflesiaceae; see supplementary table S4, Supplementary Material online), and used the plastomes of autotrophic or facultatively parasitic taxa as BLASTN references.

Finer examinations of transferred mitochondrial sequences in *G. elata* and *G. menghaiensis* were enabled by nuclear genome assemblies. BLASTN was used, with same parameter settings as for plastid sequence detection except a higher identity threshold (90%), to search for mitochondrial sequences that match corresponding nuclear genomes (*G. menghaiensis*: NCBI GenBank No. OP413684-OP413712; *G. elata*: NCBI GenBank No. CM028669-CM028686 and National Genomics Data Center BioProject No. PRJCA005619). We identified eight specific chromosomes (<10% of length covered by syntenic blocks between *G. elata* and *G. menghaiensis*) with largely (>90%) uncharacterized sequence components (not overlapping with nuclear hits, MTPTs, or fungal transfers). These chromosomes were further searched for transfers from distant plants. To do this, we performed online BLASTN (under default settings) with these chromosomes as input, and extracted blocks for which there were hits of >250 bp present in at least five non-orchid taxa. We then downloaded all qualified sequences and reconstructed maximum likelihood phylogenies based on them. Details for phylogenetic inference and criteria for determining horizontal transfers are specified below. Six blocks totaling 9.2 kb showed evidence of horizontal transfer, and their existence was verified by PacBio reads (four of *G. elata*; supplementary tables S5 to S8, Supplementary Material online) and PCR (two of *G. menghaiensis*; see below).

Quantification of Recombination Activity

Frequency of repeat-mediated homologous recombination was assessed on randomly selected identical repeat pairs of ≥50 bp. For analyzed orchid mitogenomes which were published previously (*A. shenzhenica*, *D. brymerianum*, *E. roseum*, and *P. micranthum*), we assembled their mitochondrial contigs using publicly available Illumina sequencing data (SRA accessions: SRR5759388, SRR18635896, SRR1344935, and SRR32292552), with their published assemblies as reference for the GetOrganelle pipeline, and searched for repeats therein. Illumina reads were then mapped to the reference and alternative conformations (see Mitogenome Assembly section for details) associated with each repeat pair using Bowtie2, and reads pairs supporting either conformation were counted. The recombination ratio is calculated as the proportion of repeat pairs supporting alternative conformation.

PCR Assays

Relative qPCR measurement was performed on *G.* sp. Individual 2 to validate the mitochondrial chromosome copy

number estimated with Illumina sequencing depth, and conventional PCR on *G. menghaiensis* for the existence of two mitochondrial blocks of putative horizontal origin from other angiosperms. For the former, we selected five mitochondrial chromosomes that represent the range of sequencing depth variation in this species and assayed four markers of ca. 120 bp on its plastome and each of these mitochondrial chromosomes. The nuclear gene *XY4* was used as a control. Markers were amplified for 40 cycles using TSE501 ArtiCan^{ATM} SYBR qPCR Mix (Tsingke, Beijing, China) with a qTOWER3 Real-time Thermal Cycler System (Analytik Jena, Jena, Germany). For *G. menghaiensis*, markers surrounding blocks of interest were amplified using Taq PCR Master Mix (Bioss, MA, USA) with a Biometra GmbH PowerCycler (Analytik Jena, Jena, Germany). Products were visualized with agarose gel electrophoresis and validated with Sanger sequencing. All primers were designed with Primer3 (Untergasser et al. 2012; supplementary table S9, Supplementary Material online), and reactions were performed in 20-μL volumes.

Substitution Rate Estimation and Structural Prediction

We retrieved a wide range of mitogenome assemblies that covered major angiosperm lineages, especially those with rapidly evolving coding sequences (Sloan et al. 2012; Skippington et al. 2015; Guo et al. 2023), plus a gymnosperm, *Cycas taitungensis* (Cycadaceae). Only one representative was selected for each genus except *Silene* (Caryophyllaceae), given its remarkable intrageneric heterogeneity in substitution rate (Sloan et al. 2012). Protein-encoding sequences of all sampled mitogenomes were extracted and aligned with MAFFT (under "codon" mode) and manually checked, then trimmed with Gblocks v0.91b (Talavera and Castresana 2007) under default settings to remove poorly aligned sites. Maximum likelihood phylogeny was reconstructed for each protein-encoding gene using IQ-TREE v2.2.0.3 (Lam-Tung et al. 2015), with the best-fit substitution model determined by ModelFinder (Kalyaanamoorthy et al. 2017) under the corrected Akaike Information Criterion (AICc). Node supports were based on 1,000 ultrafast bootstrap pseudoreplicates. We determined a gene as horizontally acquired if the corresponding species strongly clusters (BS > 0.70) with distant taxa in a way discordant with established mitochondrial-based phylogenies (Hu et al. 2023). Across our assemblies, genes of potential horizontal origins are found only in *N. mackinnonii*, coexisting with native copies (supplementary fig. S11, Supplementary Material online). We then concatenated all native protein-encoding sequences into a matrix, aligned and trimmed the matrix as described above, and partitioned the matrix according to gene boundaries. A maximum likelihood phylogeny was then generated and used for absolute substitution rate estimation. The phylogeny was first calibrated using r8s v1.81 (Sanderson 2003), with the PL method and optimal smoothing factor determined by the implemented cross validation function, and time constraints on ten nodes (supplementary table S11, Supplementary Material online). Synonymous (d_S) and non-synonymous (d_N) substitution rates for each protein-encoding gene were computed with codeml in the PAML package (Yang 2007), where we assumed branch-specific d_N/d_S ratios and applied the F3 × 4 method for codon frequency calculation. Genes subject to *trans*-splicing were additionally split into fragments and analyzed separately. Absolute substitution rates were then derived by dividing d_S and d_N estimates from each terminal branch by corresponding divergence times, except for the genus *Gastrodia*, where d_S, d_N, and divergence times were summed from its crown node due to very recent divergence of some sampled species (Wen et al. 2022).

Structural Prediction

As *Gastrodia* members harbor a unique insertion inferred to represent part of the coding sequence of *rps3* gene, we used AlphaFold 3 (Abramson et al. 2024), via its web server (https://alphafoldserver.com/), to predict the three-dimensional structure of the protein encoded by this gene, and results were visualized with ChimeraX (Meng et al. 2023).

Chromosome Evolution Reconstruction

We selected mitochondrial chromosomes that are homologous over ≥90% length in at least three *Gastrodia* species. These homologues, as well as any others harboring syntenic blocks covering ≥10% of their length, were used for analyses. Alternatively, the chromosome is considered lost if <10% of its length is covered in a mitogenome. Ancestral sequences were then reconstructed with maximum likelihood inference implemented in MEGA X (Kumar et al. 2018), using the GTR + G + I model and constrained with the phylogeny generated above. We then compared the ancestral sequence with its immediate descendents' to count indels and SNPs along each branch. Rearrangements, such as chromosome splits and fusions mediated by homologous recombination, were inferred based on reshuffled syntenic blocks. However, given the limited synteny among *Gastrodia* clades, such events, especially for ancient ones, are likely underestimated.

Supplementary Material

Supplementary material is available at *Molecular Biology and Evolution* online.

Acknowledgments

We thank Yajun Wang, Dr. Xuelian Guo, Dongliang Lin, and Chao Ye for their technical support, and Wenchuang He for his comments. This research was supported by the National Key Research and Development Program of China (2022YFF1301704), Linzhi Science and Technology Program (2023-QYCX-02), CAS-ANS0-SDRP-2024-02, and National Natural Science Foundation of China (31870195, 32270214).

Author Contributions

Conceptualization: H.W. and X.J.; sample collection: Z.L., Jw.L., Y.F., B.C., and X.J.; data acquisition and analyses: H.W., D.W., B.S., Jr.L., Z.L., Y.F., Y.W., and S.Q.; preparation of original draft: H.W.; reviewing and editing of manuscript: D.W., M.C., B.C., Z.W., and X.J.; acquisition of fundings: X.J.. All authors read and approved the final manuscript.

Conflict of Interest

The authors declare no competing interests.

Data Availability

All assembled mitogenomes have been deposited at NCBI GenBank, with accession numbers listed in supplementary table S2, Supplementary Material online.

References

Abramson J, Adler J, Dunger J, Evans R, Green T, Pritzel A, Ronneberger O, Willmore L, Ballard AJ, Bambrick J, *et al.* Accurate structure prediction of biomolecular interactions with AlphaFold 3. *Nature*. 2024:630(8016):493–500. https://doi.org/10.1038/s41586-024-07487-w.

Allen JO, Fauron CM, Minx P, Roark L, Oddiraju S, Lin GN, Meyer L, Sun H, Kim K, Wang CY, *et al.* Comparisons among two fertile and three male-sterile mitochondrial genomes of maize. *Genetics*. 2007:177(2):1173–1192. https://doi.org/10.1534/genetics.107.073312.

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

Alverson AJ, Wei XX, Rice DW, Stern DB, Barry K, Palmer JD. Insights into the evolution of mitochondrial genome size from complete sequences of *Citrullus lanatus* and *Cucurbita pepo* (Cucurbitaceae). *Mol Biol Evol*. 2010:27(6):1436–1448. https://doi.org/10.1093/molbev/msq029.

Arimura S, Yamamoto J, Aida GP, Nakazono M, Tsutsumi N. Frequent fusion and fission of plant mitochondria with unequal nucleoid distribution. *Proc Natl Acad Sci U S A*. 2004:101(20):7805–7808. https://doi.org/10.1073/pnas.0401077101.

Backert S, Nielsen BL, Börner T. The mystery of the rings: structure and replication of mitochondrial genomes from higher plants. *Trends Plant Sci*. 1997:2(12):477–483. https://doi.org/10.1016/S1360-1385(97)01148-5.

Barkman TJ, McNeal JR, Lim S-H, Coat G, Croom HB, Young ND, dePamphilis CW. Mitochondrial DNA suggests at least 11 origins of parasitism in angiosperms and reveals genomic chimerism in parasitic plants. *BMC Evol Biol*. 2007:7:248. https://doi.org/10.1186/1471-2148-7-248.

Bergthorsson U, Adams KL, Thomason B, Palmer JD. Widespread horizontal transfer of mitochondrial genes in flowering plants. *Nature*. 2003:424(6945):197–201. https://doi.org/10.1038/nature01743.

Birky CW Jr. The inheritance of genes in mitochondria and chloroplasts: laws, mechanisms, and models. *Annu Rev Genet*. 2001:35(1):125–148. https://doi.org/10.1146/annurev.genet.35.102401.090231.

Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*. 2014:30(15):2114–2120. https://doi.org/10.1093/bioinformatics/btu170.

Braukmann TWA, Broe MB, Stefanovic S, Freudenstein JV. On the brink: the highly reduced plastomes of nonphotosynthetic Ericaceae. *New Phytol*. 2017:216(1):254–266. https://doi.org/10.1111/nph.14681.

Brazda V, Kolomaznik J, Lysek J, Haronikova L, Coufal J, St'astny J. Palindrome analyser—a new web-based server for predicting and evaluating inverted repeats in nucleotide sequences. *Biochem Biophys Res Commun*. 2016:478(4):1739–1745. https://doi.org/10.1016/j.bbrc.2016.09.015.

Bromham L, Cowman PF, Lanfear R. Parasitic plants have increased rates of molecular evolution across all three genomes. *BMC Evol Biol*. 2013:13:126. https://doi.org/10.1186/1471-2148-13-126.

Burger G, Lang BF. Parallels in genome evolution in mitochondria and bacterial symbionts. *IUBMB Life*. 2003:55(4-5):205–212. https://doi.org/10.1080/1521654031000137380.

Cho Y, Mower JP, Qiu Y-L, Palmer JD. Mitochondrial substitution rates are extraordinarily elevated and variable in a genus of flowering plants. *Proc Natl Acad Sci U S A*. 2004:101(51):17741–17746. https://doi.org/10.1073/pnas.0408302101.

Christensen AC. Plant mitochondria are a riddle wrapped in a mystery inside an enigma. *J Mol Evol*. 2021:89(3):151–156. https://doi.org/10.1007/s00239-020-09980-y.

Cusimano N, Wicke S. Massive intracellular gene transfer during plastid genome reduction in nongreen Orobanchaceae. *New Phytol*. 2016:210(2):680–693. https://doi.org/10.1111/nph.13784.

Danecek P, Bonfield JK, Liddle J, Marshall J, Ohan V, Pollard MO, Whitwham A, Keane T, McCarthy SA, Davies RM, *et al.* Twelve years of SAMtools and BCFtools. *GigaScience*. 2021:10(2):giab008. https://doi.org/10.1093/gigascience/giab008.

Davis CC, Wurdack KJ. Host-to-parasite gene transfer in flowering plants: phylogenetic evidence from Malpighiales. *Science*. 2004:305(5684):676–678. https://doi.org/10.1126/science.1100671.

Davis CC, Xi ZX. Horizontal gene transfer in parasitic plants. *Curr Opin Plant Biol*. 2015:26:14–19. https://doi.org/10.1016/j.pbi.2015.05.008.

Fan WS, Liu F, Jia QY, Du HY, Chen W, Ruan JW, Lei JJ, Li D-Z, Mower JP, Zhu AD. *Fragaria* mitogenomes evolve rapidly in structure but slowly in sequence and incur frequent multinucleotide mutations mediated by microinversions. *New Phytol*. 2022:236(2):745–759. https://doi.org/10.1111/nph.18334.

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

Giannakis K, Richards L, Dauda KA, Johnston IG. Connecting species-specific extents of genome reduction in mitochondria and plastids. *Mol Biol Evol*. 2024:41(6):msae097. https://doi.org/10.1093/molbev/msae097.

Goremykin VV, Salamini F, Velasco R, Viola R. Mitochondrial DNA of *Vitis vinifera* and the issue of rampant horizontal gene transfer. *Mol Biol Evol*. 2009:26(1):99–110. https://doi.org/10.1093/molbev/msn226.

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

Gray MW, Burger G, Lang BF. Mitochondrial evolution. *Science*. 1999:283(5407):1476–1481. https://doi.org/10.1126/science.283.5407.1476.

Gualberto JM, Newton KJ. Plant mitochondrial genomes: dynamics and mechanisms of mutation. *Annu Rev Plant Biol*. 2017:68:225–252. https://doi.org/10.1146/annurev-arplant-043015-112232.

Guo X, Wang F, Fang DM, Lin QQ, Sahu SK, Luo LM, Li JN, Chen YW, Dong SS, Chen SS, *et al.* The genome of *Acorus* deciphers insights into early monocot evolution. *Nat Commun*. 2023:14(1):3662. https://doi.org/10.1038/s41467-023-38836-4.

Hao Z, Lv D, Ge Y, Shi J, Weijers D, Yu G, Chen J. *RIdeogram*: drawing SVG graphics to visualize and map genome-wide data on the idiograms. *PeerJ Comput Sci*. 2020:6:e251. https://doi.org/10.7717/peerj-cs.251.

He W, Xiang K, Chen C, Wang J, Wu Z. Master graph: an essential integrated assembly model for the plant mitogenome based on a graph-based framework. *Brief Bioinform*. 2023:24(1):bbac522. https://doi.org/10.1093/bib/bbac522.

Hu HY, Sun PC, Yang YZ, Ma JX, Liu JQ. Genome-scale angiosperm phylogenies based on nuclear, plastome, and mitochondrial datasets. *J Integr Plant Biol*. 2023:65(6):1479–1489. https://doi.org/10.1111/jipb.13455.

Huang J. Horizontal gene transfer in eukaryotes: the weak-link model. *Bioessays*. 2013:35(10):868–875. https://doi.org/10.1002/bies.201300007.

Jiang Y, Hu X, Yuan Y, Guo X, Chase MW, Ge S, Li J, Fu J, Li K, Hao M, *et al.* The *Gastrodia menghaiensis* (Orchidaceae) genome provides new insights of orchid mycorrhizal interactions. *BMC Plant Biol*. 2022:22(1):179. https://doi.org/10.1186/s12870-022-03573-1.

Jin J-J, Yu W-B, Yang J-B, Song Y, dePamphilis CW, Yi T-S, Li D-Z. GetOrganelle: a fast and versatile toolkit for accurate de novo assembly of organelle genomes. *Genome Biol*. 2020:21(1):241. https://doi.org/10.1186/s13059-020-02154-5.

Kalyaanamoorthy S, Bui Quang M, Wong TKF, von Haeseler A, Jermiin LS. ModelFinder: fast model selection for accurate phylogenetic

estimates. *Nat Methods*. 2017:14(6):587–589. https://doi.org/10.1038/nmeth.4285.
Katoh K, Standley DM. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol Biol Evol*. 2013:30(4):772–780. https://doi.org/10.1093/molbev/mst010.
Kim YK, Jo S, Cheon S-H, Hong J-R, Kim K-J. Ancient horizontal gene transfers from plastome to mitogenome of a nonphotosynthetic orchid, *Gastrodia pubilabiata* (Epidendroideae. Orchidaceae). *Int J Mol Sci*. 2023:24(14):11448. https://doi.org/10.3390/ijms241411448.
Kohany O, Gentles AJ, Hankus L, Jurka J. Annotation, submission and screening of repetitive elements in Repbase: RepbaseSubmitter and Censor. *BMC Bioinformatics*. 2006:7:474. https://doi.org/10.1186/1471-2105-7-474.
Koren S, Walenz BP, Berlin K, Miller JR, Bergman NH, Phillippy AM. Canu: scalable and accurate long-read assembly via adaptive *k*-mer weighting and repeat separation. *Genome Res*. 2017:27(5):722–736. https://doi.org/10.1101/gr.215087.116.
Kozik A, Rowan BA, Lavelle D, Berke L, Schranz ME, Michelmore RW, Christensen AC. The alternative reality of plant mitochondrial DNA: one ring does not rule them all. *PLoS Genet*. 2019:15(8):e1008373. https://doi.org/10.1371/journal.pgen.1008373.
Krause K. From chloroplasts to "cryptic" plastids: evolution of plastid genomes in parasitic plants. *Curr Genet*. 2008:54(3):111–121. https://doi.org/10.1007/s00294-008-0208-8.
Kumar S, Stecher G, Li M, Knyaz C, Tamura K. MEGA X: molecular evolutionary genetics analysis across computing platforms. *Mol Biol Evol*. 2018:35(6):1547–1549. https://doi.org/10.1093/molbev/msy096.
Lam-Tung N, Schmidt HA, von Haeseler A, Bui Quang M. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol Biol Evol*. 2015:32(1):268–274. https://doi.org/10.1093/molbev/msu300.
Lang BF, Beck N, Prince S, Sarrasin M, Rioux P, Burger G. Mitochondrial genome annotation with MFannot: a critical analysis of gene identification and gene model prediction. *Front Plant Sci*. 2023:14:1222186. https://doi.org/10.3389/fpls.2023.1222186.
Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. *Nat Methods*. 2012:9(4):357–359. https://doi.org/10.1038/nmeth.1923.
Lin QS, Banerjee A, Stefanovi S. Mitochondrial phylogenomics of *Cuscuta* (Convolvulaceae) reveals a potentially functional horizontal gene transfer from the host. *Genome Biol Evol*. 2022a:14(6):evac091. https://doi.org/10.1093/gbe/evac091.
Lin YX, Li P, Zhang YC, Akhter D, Pan RH, Fu ZX, Huang MQ, Li XB, Feng YL. Unprecedented organelle genomic variations in morning glories reveal independent evolutionary scenarios of parasitic plants and the diversification of plant mitochondrial complexes. *BMC Biol*. 2022b:20(1):49. https://doi.org/10.1186/s12915-022-01250-1.
Liu F, Fan WS, Yang J-B, Xiang C-L, Mower JP, Li D-Z, Zhu AD. Episodic and guanine-cytosine-biased bursts of intragenomic and interspecific synonymous divergence in Ajugoideae (Lamiaceae) mitogenomes. *New Phytol*. 2020:228(3):1107–1114. https://doi.org/10.1111/nph.16753.
Liu X, Luo H, Liu Z, Yang B. Mitochondrial genome characteristics reveal evolution of *Danxiaorchis yangii* and phylogenetic relationships. *Int J Mol Sci*. 2025:26(2):562. https://doi.org/10.3390/ijms26020562.
Lowe TM, Chan PP. tRNAscan-SE On-line: integrating search and context for analysis of transfer RNA genes. *Nucleic Acids Res*. 2016:44(W1):W54–W57. https://doi.org/10.1093/nar/gkw413.
Lynch M. *The origins of genome architecture*. Sunderland (MA): Sinauer Associates; 2007.
Lynch M. Evolution of the mutation rate. *Trends Genet*. 2010:26(8):345–352. https://doi.org/10.1016/j.tig.2010.05.003.
Lynch M, Ali F, Lin T, Wang Y, Ni J, Long H. The divergence of mutation rates and spectra across the Tree of Life. *EMBO Rep*. 2023:24(10):e57561. https://doi.org/10.15252/embr.202357561.
Meng EC, Goddard TD, Pettersen EF, Couch GS, Pearson ZJ, Morris JH, Ferrin TE. UCSF chimeraX: tools for structure building and analysis. *Protein Sci*. 2023:32(11):e4792. https://doi.org/10.1002/pro.4792.
Merckx V, Freudenstein JV. Evolution of mycoheterotrophy in plants: a phylogenetic perspective. *New Phytol*. 2010:185(3):605–609. https://doi.org/10.1111/j.1469-8137.2009.03155.x.
Michaud M, Cognat V, Duchêne A-M, Maréchal-Drouard L. A global picture of tRNA genes in plant genomes. *Plant J*. 2011:66(1):80–93. https://doi.org/10.1111/j.1365-313X.2011.04490.x.
Mower JP. Variation in protein gene and intron content among land plant mitogenomes. *Mitochondrion*. 2020:53:203–213. https://doi.org/10.1016/j.mito.2020.06.002.
Mower JP, Jain K, Hepburn NJ. The role of horizontal transfer in shaping the plant mitochondrial genome. In: Maréchal-Drouard L, editors. *Mitochondrial genome evolution*. Oxford, UK: Academic Press; 2012a. p. 41–69.
Mower JP, Sloan DB, Alverson AJ. Plant mitochondrial genome diversity: the genomics revolution. In: Wendel JF, Greilhuber J, Dolezel J, Leitch IJ, editors. *Plant genome diversity volume 1: plant genomes, their residents, and their evolutionary dynamics*. Vienna, Austria: Springer; 2012b. p. 123–144.
Mower JP, Stefanovic S, Hao W, Gummow JS, Jain K, Ahmed D, Palmer JD. Horizontal acquisition of multiple mitochondrial genes from a parasitic plant followed by gene conversion with host mitochondrial genes. *BMC Biol*. 2010:8:150. https://doi.org/10.1186/1741-7007-8-150.
Palmer JD, Herbon LA. Plant mitochondrial DNA evolves rapidly in structure, but slowly in sequence. *J Mol Evol*. 1988:28(1-2):87–97. https://doi.org/10.1007/BF02143500.
Palmer JD, Shields CR, Cohen DB, Orton TJ. An unusual mitochondrial DNA plasmid in the genus *Brassica*. *Nature*. 1983:301(5902):725–728. https://doi.org/10.1038/301725a0.
Pearson CE, Zorbas H, Price GB, ZannisHadjopoulos M. Inverted repeats, stem-loops, and cruciforms: significance for initiation of DNA replication. *J Cell Biochem*. 1996:63(1):1–22. https://doi.org/10.1002/(SICI)1097-4644(199610)63:1%3C1::AID-JCB1%3E3.0.CO;2-3.
Petersen G, Darby H, Lam VKY, Pedersen H, Merckx VSFT, Zervas A, Seberg O, Graham SW. Mycoheterotrophic *Epirixanthes* (Polygalaceae) has a typical angiosperm mitogenome but unorthodox plastid genomes. *Ann Bot*. 2019:124(5):791–807. https://doi.org/10.1093/aob/mcz114.
Phillips WS, Brown AMV, Howe DK, Peetz AB, Blok VC, Denver DR, Zasada IA. The mitochondrial genome of *Globodera ellingtonae* is composed of two circles with segregated gene content and differential copy numbers. *BMC Genomics*. 2016:17(1):706. https://doi.org/10.1186/s12864-016-3047-x.
Plants of the World Online (POWO). 2023. Genus *Gastrodia* R. Br. [accessed 2023 November 29]. https://powo.science.kew.org/.
Prjibelski A, Antipov D, Meleshko D, Lapidus A, Korobeynikov A. Using SPAdes *de novo* assembler. *Curr Protoc Bioinform*. 2020:70(1):e102. https://doi.org/10.1002/cpbi.102.
Rand DM. 'Why genomes in pieces?' revisited: sucking lice do their own thing in mtDNA circle game. *Genome Res*. 2009:19(5):700–702. https://doi.org/10.1101/gr.091132.109.
Rice DW, Alverson AJ, Richardson AO, Young GJ, Sanchez-Puerta MV, Munzinger J, Barry K, Boore JL, Zhang Y, dePamphilis CW, *et al*. Horizontal transfer of entire genomes via mitochondrial fusion in the angiosperm *Amborella*. *Science*. 2013:342(6165):1468–1473. https://doi.org/10.1126/science.1246275.
Richards TA, Soanes DM, Foster PG, Leonard G, Thomton CR, Talbot NJ. Phylogenomic analysis demonstrates a pattern of rare and ancient horizontal gene transfer between plants and fungi. *Plant Cell*. 2009:21(7):1897–1911. https://doi.org/10.1105/tpc.109.065805.
Richardson AO, Rice DW, Young GJ, Alverson AJ, Palmer JD. The "fossilized" mitochondrial genome of *Liriodendron tulipifera*: ancestral gene content and order, ancestral editing sites, and extraordinarily low mutation rate. *BMC Biol*. 2013:11:29. https://doi.org/10.1186/1741-7007-11-29.
Sanchez-Puerta MV, García LE, Wohlfeiler J, Ceriotti LF. Unparalleled replacement of native mitochondrial genes by foreign homologs in a

holoparasitic plant. *New Phytol*. 2017:214(1):376–387. https://doi.org/10.1111/nph.14361.

Sanderson MJ. R8s: inferring absolute rates of molecular evolution and divergence times in the absence of a molecular clock. *Bioinformatics*. 2003:19(2):301–302. https://doi.org/10.1093/bioinformatics/19.2.301.

Schneider A. Mitochondrial tRNA import and its consequences for mitochondrial translation. *Annu Rev Biochem*. 2011:80:1033–1053. https://doi.org/10.1146/annurev-biochem-060109-092838.

Shtratnikoya VY, Schelkunov MI, Penin AA, Logacheva MD. Mitochondrial genome of the nonphotosynthetic mycoheterotrophic plant *Hypopitys monotropa*, its structure, gene expression and RNA editing. *PeerJ*. 2020:8:e9309. https://doi.org/10.7717/peerj.9309.

Sinn BT, Barrett CF. Ancient mitochondrial gene transfer between fungi and the orchids. *Mol Biol Evol*. 2020:37(1):44–57. https://doi.org/10.1093/molbev/msz198.

Skippington E, Barkman TJ, Rice DW, Palmer JD. Miniaturized mitogenome of the parasitic plant *Viscum scurruloideum* is extremely divergent and dynamic and has lost all *nad* genes. *Proc Natl Acad Sci U S A*. 2015:112(27):E3515–E3524. https://doi.org/10.1073/pnas.1504491112.

Skippington E, Barkman TJ, Rice DW, Palmer JD. Comparative mitogenomics indicates respiratory competence in parasitic *Viscum* despite loss of complex I and extreme sequence divergence, and reveals horizontal gene transfer and remarkable variation in genome size. *BMC Plant Biol*. 2017:17(1):49. https://doi.org/10.1186/s12870-017-0992-8.

Sloan DB, Alverson AJ, Chuckalovcak JP, Wu M, McCauley DE, Palmer JD, Taylor DR. Rapid evolution of enormous, multichromosomal genomes in flowering plant mitochondria with exceptionally high mutation rates. *PLoS Biol*. 2012:10(1):e1001241. https://doi.org/10.1371/journal.pbio.1001241.

Sloan DB, Wu Z, Sharbrough J. Correction of persistent errors in *Arabidopsis* reference mitochondrial genomes. *Plant Cell*. 2018:30(3):525–527. https://doi.org/10.1105/tpc.18.00024.

Smith DR. The mutational hazard hypothesis of organelle genome evolution: 10 years on. *Mol Ecol*. 2016:25(16):3769–3775. https://doi.org/10.1111/mec.13742.

Suga K, Welch DBM, Tanaka Y, Sakakura Y, Hagiwarak A. Two circular chromosomes of unequal copy number make up the mitochondrial genome of the rotifer *Brachionus plicatilis*. *Mol Biol Evol*. 2008:25(6):1129–1137. https://doi.org/10.1093/molbev/msn058.

Talavera G, Castresana J. Improvement of phylogenies after removing divergent and ambiguously aligned blocks from protein sequence alignments. *Syst Biol*. 2007:56(4):564–577. https://doi.org/10.1080/10635150701472164.

Untergasser A, Cutcutache I, Koressaar T, Ye J, Faircloth BC, Remm M, Rozen SG. Primer3-new capabilities and interfaces. *Nucleic Acids Res*. 2012:40(15):e115. https://doi.org/10.1093/nar/gks596.

Wang J, Kan S, Liao X, Zhou J, Tembrock LR, Daniell H, Jin S, Wu Z. Plant organellar genomes: much done, much more to do. *Trends Plant Sci*. 2024:29(7):754–769. https://doi.org/10.1016/j.tplants.2023.12.014.

Wang M, Yu W, Yang J, Hou Z, Li C, Niu Z, Zhang B, Xue Q, Liu W, Ding X. Mitochondrial genome comparison and phylogenetic analysis of *Dendrobium* (Orchidaceae) based on whole mitogenomes. *BMC Plant Biol*. 2023:23(1):586. https://doi.org/10.1186/s12870-023-04618-9.

Ward BL, Anderson RS, Bendich AJ. The mitochondrial genome is large and variable in a family of plants (Cucurbitaceae). *Cell*. 1981:25(3):793–803. https://doi.org/10.1016/0092-8674(81)90187-2.

Warren JM, Salinas-Giegé T, Triant DA, Taylor DR, Drouard L, SloanDB. Rapid shifts in mitochondrial tRNA import in a plant lineage with extensive mitochondrial tRNA gene loss. *Mol Biol Evol*. 2021:38(12):5735–5751. https://doi.org/10.1093/molbev/msab255.

Wen YY, Qin Y, Shao BY, Li JW, Ma CB, Liu Y, Yang BY, Jin XH. The extremely reduced, diverged and reconfigured plastomes of the largest mycoheterotrophic orchid lineage. *BMC Plant Biol*. 2022:22(1):448. https://doi.org/10.1186/s12870-022-03836-x.

Westermann B. Mitochondrial fusion and fission in cell life and death. *Nat Rev Mol Cell Biol*. 2010:11(12):872–884. https://doi.org/10.1038/nrm3013.

Wick RR, Schultz MB, Zobel J, Holt KE. Bandage: interactive visualization of *de novo* genome assemblies. *Bioinformatics*. 2015:31(20):3350–3352. https://doi.org/10.1093/bioinformatics/btv383.

Wicke S, Naumann J. Molecular evolution of plastid genomes in parasitic flowering plants. In: Chaw SM, Jansen RK, editors. *Plastid genome evolution*. Oxford, UK: Academic Press; 2018. p. 315–347.

Wu ZQ, Cuthbert JM, Taylor DR, Sloan DB. The massive mitochondrial genome of the angiosperm *Silene noctiflora* is evolving by gain or loss of entire chromosomes. *Proc Natl Acad Sci U S A*. 2015:112(33):10185–10191. https://doi.org/10.1073/pnas.1421397112.

Wu Z-Q, Liao X-Z, Zhang X-N, Tembrock LR, Broz A. Genomic architectural variation of plant mitochondria-a review of multichromosomal structuring. *J Syst Evol*. 2022:60(1):160–168. https://doi.org/10.1111/jse.12655.

Xi ZX, Wang YG, Bradley RK, Sugumaran M, Marx CJ, Rest JS, Davis CC. Massive mitochondrial gene transfer in a parasitic flowering plant clade. *PLoS Genet*. 2013:9(2):e1003265. https://doi.org/10.1371/journal.pgen.1003265.

Xu YX, Lei YT, Su ZX, Zhao M, Zhang JX, Shen GJ, Wang L, Li J, Qi JF, Wu JQ. A chromosome-scale *Gastrodia elata* genome and large-scale comparative genomic analysis indicate convergent evolution by gene loss in mycoheterotrophic and parasitic plants. *Plant J*. 2021:108(6):1609–1623. https://doi.org/10.1111/tpj.15528.

Yang J-X, Dierckxsens N, Bai M-Z, Guo Y-Y. Multichromosomal mitochondrial genome of *Paphiopedilum micranthum*: compact and fragmented genome, and rampant intracellular gene transfer. *Int J Mol Sci*. 2023:24(4):3976. https://doi.org/10.3390/ijms24043976.

Yang Z. PAML 4: phylogenetic analysis by maximum likelihood. *Mol Biol Evol*. 2007:24(8):1586–1591. https://doi.org/10.1093/molbev/msm088.

Yu R, Chen X, Long L, Jost M, Zhao R, Liu L, Mower JP, dePamphilis CW, Wanke S, Jiao Y. De novo assembly and comparative analyses of mitochondrial genomes in Piperales. *Genome Biol Evol*. 2023:15(3):evad041. https://doi.org/10.1093/gbe/evad041.

Yu R, Sun C, Liu Y, Zhou R. Shifts from *cis*-to *trans*-splicing of five mitochondrial introns in *Tolypanthus maclureis*. *PeerJ*. 2021:9:e12260. https://doi.org/10.7717/peerj.12260.

Yu R, Sun C, Zhong Y, Liu Y, Sanchez-Puerta MV, Mower JP, Zhou R. The minicircular and extremely heteroplasmic mitogenome of the holoparasitic plant *Rhopalocnemis phalloides*. *Curr Biol*. 2022:32(2):470–479.e5. https://doi.org/10.1016/j.cub.2021.11.053.

Yuan Y, Jin XH, Liu J, Zhao X, Zhou JH, Wang X, Wang DY, Lai CJS, Xu W, Huang JW, *et al.* The *Gastrodia elata* genome provides insights into plant adaptation to heterotrophy. *Nat Commun*. 2018:9(1):1615. https://doi.org/10.1038/s41467-018-03423-5.

Zhang G, Hu Y, Huang M-Z, Huang W-C, Liu D-K, Zhang D, Hu H, Downing JL, Liu Z-J, Ma H. Comprehensive phylogenetic analyses of Orchidaceae using nuclear genes and evolutionary insights into epiphytism. *J Integr Plant Biol*. 2023:65(5):1204–1225. https://doi.org/10.1111/jipb.13462.

Zhang Z, Schwartz S, Wagner L, Miller W. A greedy algorithm for aligning DNA sequences. *J Comput Biol*. 2000:7(1-2):203–214. https://doi.org/10.1089/10665270050081478.

Zhao Z, Li Y, Zhai J-W, Liu Z-J, Li M-H. Organelle genomes of *Epipogium roseum* provide insight into the evolution of mycoheterotrophic orchids. *Int J Mol Sci*. 2024:25(3):1578. https://doi.org/10.3390/ijms25031578.

Zuker M. Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res*. 2003:31(13):3406–3415. https://doi.org/10.1093/nar/gkg595.

Zwonitzer KD, Tressel LG, Wu Z, Kan S, Broz AK, Mower JP, Ruhlman TA, Jansen RK, Sloan DB, Havird JC. Genome copy number predicts extreme evolutionary rate variation in plant mitochondrial DNA. *Proc Natl Acad Sci U S A*. 2024:121(10):e2317240121. https://doi.org/10.1073/pnas.2317240121.