



OPEN ACCESS

EDITED BY

Juan Carlos Villarreal A.,
Laval University, Canada

REVIEWED BY

Eric Schuettpehl,
Smithsonian Institution, United States
Tom A. Ranker,
University of Hawaii at Manoa, United States

*CORRESPONDENCE

Li-Yang Kuo
✉ lykuo@life.nthu.edu.tw
Jason R. Grant
✉ jason.grant@unine.ch

RECEIVED 15 September 2023

ACCEPTED 27 December 2023

PUBLISHED 15 January 2024

CITATION

Kuo L-Y, Su H-J, Koubínová D, Xie P-J,
Whitehouse C, Ebihara A and Grant JR (2024)
Organellar phylogenomics of
Ophioglossaceae fern genera.
Front. Plant Sci. 14:1294716.
doi: 10.3389/fpls.2023.1294716

COPYRIGHT

© 2024 Kuo, Su, Koubínová, Xie, Whitehouse,
Ebihara and Grant. This is an open-access
article distributed under the terms of the
[Creative Commons Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/).
The use, distribution or reproduction in other
forums is permitted, provided the original
author(s) and the copyright owner(s) are
credited and that the original publication in
this journal is cited, in accordance with
accepted academic practice. No use,
distribution or reproduction is permitted
which does not comply with these terms.

Organellar phylogenomics of Ophioglossaceae fern genera

Li-Yang Kuo^{1*}, Huei-Jiun Su², Darina Koubínová³,
Pei-Jun Xie¹, Christopher Whitehouse⁴, Atsushi Ebihara⁵
and Jason R. Grant^{3*}

¹Institute of Molecular & Cellular Biology, National Tsing Hua University, Hsinchu, Taiwan,

²Department of Earth and Life Sciences, University of Taipei, Taipei, Taiwan, ³University of Neuchâtel,
Laboratory of Evolutionary Genetics, Neuchâtel, Switzerland, ⁴Phillipskop Mountain Reserve, Western
Cape, Stanford, South Africa, ⁵Department of Botany, National Museum of Nature and Science,
Tsukuba, Japan

Previous phylogenies showed conflicting relationships among the subfamilies and genera within the fern family Ophioglossaceae. However, their classification remains unsettled where contrasting classifications recognize four to 15 genera. Since these treatments are mostly based on phylogenetic evidence using limited, plastid-only loci, a phylogenomic understanding is actually necessary to provide conclusive insight into the systematics of the genera. In this study, we have therefore compiled datasets with the broadest sampling of Ophioglossaceae genera to date, including all fifteen currently recognized genera, especially for the first time the South African endemic genus *Rhizoglossum*. Notably, our comprehensive phylogenomic matrix is based on both plastome and mitogenome genes. Inferred from the coding sequences of 83 plastid and 37 mitochondrial genes, a strongly supported topology for these subfamilies is presented, and is established by analyses using different partitioning approaches and substitution models. At the generic level, most relationships are well resolved except for few within the subfamily Ophioglossoidae. With this new phylogenomic scheme, key morphological and genomic changes were further identified along this backbone. In addition, we confirmed numerous horizontally transferred (HGT) genes in the genera *Botrypus*, *Helminthostachys*, *Mankyua*, *Sahashia*, and *Sceptridium*. These HGT genes are most likely located in mitogenomes and are predominately donated from angiosperm Santalales or non-Ophioglossaceae ferns. By our in-depth searches of the organellar genomes, we also provided phylogenetic overviews for the plastid and mitochondrial MORFFO genes found in these Ophioglossaceae ferns.

KEYWORDS

horizontal gene transfer, mitogenome, MORFFO, Ophioglossaceae, *Rhizoglossum*, phylogenomic, plastome, Santalales

1 Introduction

In ferns, the family Ophioglossaceae (adder's tongues) is the richest in species among the extant eusporangiate lineages, with over 40% of the species (PPG I, 2016; ~65% estimated by Zhang and Zhang, 2022). Members in this family are highly diversified not only in extrinsic morphology (Figure 1) with body size ranging from 1 cm up to 2 m, leaves simple to 4-pinnatifid (Wagner, 1990; Patel and Reddy, 2018), but also intrinsically in their genomic contents with haploid chromosome numbers ranging from 44 to more than 500 (Khandelwal, 1990; Table 1), and genome sizes from 6 to > 90 Gbp/C (Kuo et al., 2021). Despite awareness of this disparity within the fern tree of life, the infra-family phylogeny of Ophioglossaceae remains largely unsettled. Due to extreme scarcity of fossils found in the Ophioglossaceae (Rothwell and Stockey, 1989), our evolutionary understanding of this ancient family is built on the phylogeny of solely extant taxa that are currently classified into four subfamilies (PPG I, 2016; Zhang et al., 2020; Zhang and Zhang, 2022). Among these subfamilies, Ophioglossoideae and Botrychioideae are species rich, whereas the remaining two, Helminthostachyoideae and Mankyuoideae are monotypic, each

composed by a single species, *Helminthostachys zeylanica* (L.) Hook. and *Mankyua chejuense* B.Y.Sun, M.H.Kim & C.H.Ki. These two orphan lineages are deeply rooted in the Ophioglossaceae phylogeny but change positions in different analyses (Figure 2A). *Helminthostachys* had been revealed as the sister of Botrychioideae in many previous phylogenies (Hauk et al., 2003; Sun et al., 2009; Shinohara et al., 2013; Kim and Kim, 2018; Zhang et al., 2020), but in others, as the sister to *Mankyua* + Ophioglossoideae or Botrychioideae + Ophioglossoideae (Shinohara et al., 2013; Shen et al., 2020). Similarly, in these trees, *Mankyua* was placed in different positions that were sister to either Ophioglossoideae, *Helminthostachys* + Ophioglossoideae, *Helminthostachys* + Botrychioideae, or all the remaining subfamilies (Figure 2A).

Notably, previous attempts to infer the molecular phylogeny of Ophioglossaceae are all based on plastid sequences. The majority included limited nucleotide sites from no more than six genetic loci and received weak branch supports regarding to the relationships mentioned above. A few of these studies used plastomic datasets for phylogenetic reconstructions, but still revealed conflicting topologies (Kim and Kim, 2018; Shen et al., 2020; Zhang and



FIGURE 1

Photos of living plants of Ophioglossaceae genera. (A) *Sahashia stricta*. (B) *Botrypus virginianus*. (C) *Helminthostachys zeylanica*. (D) *Cheiroglossa palmata*. (E) *Ophioderma pendula*. (F) *Japanobotrychium lanuginosum*. (G) *Botrychium lunaria*. (H) *Mankyua chejuensis*. (I) *Rhizoglossum bergianum*. (J) *Whittieria engelmannii*. (K) *Sceptridium formosanum*. (L) *Holubiella lunarioides*. (M) *Ophioglossum reticulatum*. (N) *Haukia crotalophoroides*. (O) *Goswamia isanensis*. Photo credits by Li-Yaung Kuo (A, K), Zhi-Xiang Chang (B, F, G), Tian-Chuan Hsu (C), Emily B. Sessa (D, L, N), Pi-Fong Lu (E), Christopher Whitehouse (I), Ponpipat Limpanasitticha (M), and Tassanai Jaruwattanaphan (O). The photo of *Mankyua chejuensis* (H) is modified from Lee et al. (2022).

TABLE 1 Morphological and genomic features of Ophioglossaceae genera.

| Subfamily | Genera (species number) ^a | Tophophore blade | Vascular cambium | Venation | Sporophophore blade | Sporangia | Root | Chromosome basic number ^b | Genome size (Gbp/C) ^c |
|-----------------------|--------------------------------------|------------------|------------------|-----------------------------|---------------------|---------------------------------------------|------------|--------------------------------------|----------------------------------|
| Botrychioideae | <i>Sahashia</i> (1) | Divided | | Free | Divided | Free, longitudinal/subtransverse dehiscence | Non-budded | 44 | 6.0 |
| Botrychioideae | <i>Botrypus</i> (1) | Divided | | Free | Divided | Free, longitudinal/subtransverse dehiscence | Non-budded | 46 (92) | 9.3-11.3 |
| Botrychioideae | <i>Botrychium</i> (35) | Divided | Presence | Free | Divided | Free, horizontal dehiscence | Non-budded | 45 | 10.8-28.6 |
| Botrychioideae | <i>Japanobotrychium</i> (1) | Divided | Presence | Free | Divided | Free, horizontal dehiscence | Non-budded | 45 (90) | 14.1-18.6 |
| Botrychioideae | <i>Holubiella</i> (1) | Divided | Presence | Free | Divided | Free, horizontal dehiscence | Non-budded | 45 | 9.1-9.4 |
| Botrychioideae | <i>Sceptridium</i> (24) | Divided | Presence | Free | Divided | Free, horizontal dehiscence | Non-budded | 45 | 7.8-30.6 |
| Helminthostachyoideae | <i>Helminthostachys</i> (1) | Divided | Absence | Free, seldomly anastomosing | Simple | Free, horizontal dehiscence | Non-budded | 94 | 11.7-13.1 |
| Mankyuoidae | <i>Mankyua</i> (1) | Divided | | Free | Simple or divided | Sunken, basal fused, horizontal dehiscence | Budded | 130 | 12.7 |
| Ophioglossoideae | <i>Ophioglossum</i> (~50) | Simple | Absence | Anastomosing | Simple | Sunken, basal fused, horizontal dehiscence | Budded | 120 | 24.5-128.2 |
| Ophioglossoideae | <i>Haukia</i> (2) | Simple | Absence? | Anastomosing | Simple | Sunken, basal fused, horizontal dehiscence | Non-budded | | 18.54 |
| Ophioglossoideae | <i>Whittieria</i> (1) | Simple | Absence? | Anastomosing | Simple | Sunken, basal fused, horizontal dehiscence | Budded | | 16.29 |
| Ophioglossoideae | <i>Rhizoglossum</i> (1) | Simple | Absence? | Anastomosing | Simple | Sunken, basal fused, horizontal dehiscence | Budded | | |
| Ophioglossoideae | <i>Goswamia</i> (15) | Simple | Absence? | Anastomosing | Simple | Sunken, basal fused, horizontal dehiscence | Budded | 120 (86)? | 18.61 |
| Ophioglossoideae | <i>Cheiroglossa</i> (2) | Simple | Absence? | Anastomosing | Simple | Sunken, basal fused, horizontal dehiscence | Budded | | 53.5 |
| Ophioglossoideae | <i>Ophioderma</i> (6) | Simple | Absence? | Anastomosing | Simple | Sunken, basal fused, horizontal dehiscence | Budded | 370-380 | 98.8-136.7 |

^abased on PPG 1, Zhang et al. (2020) and Zhang and Zhang (2022).

^breviewed in Shinohara et al., 2013, CCDB (<http://ccdb.tau.ac.il/>).

^creviewed in Kuo et al (2021) unpublished) and Fujiwara et al., 2023.

Morphological characters are based on Hauk et al. (2003) and Zhang and Zhang (2022).

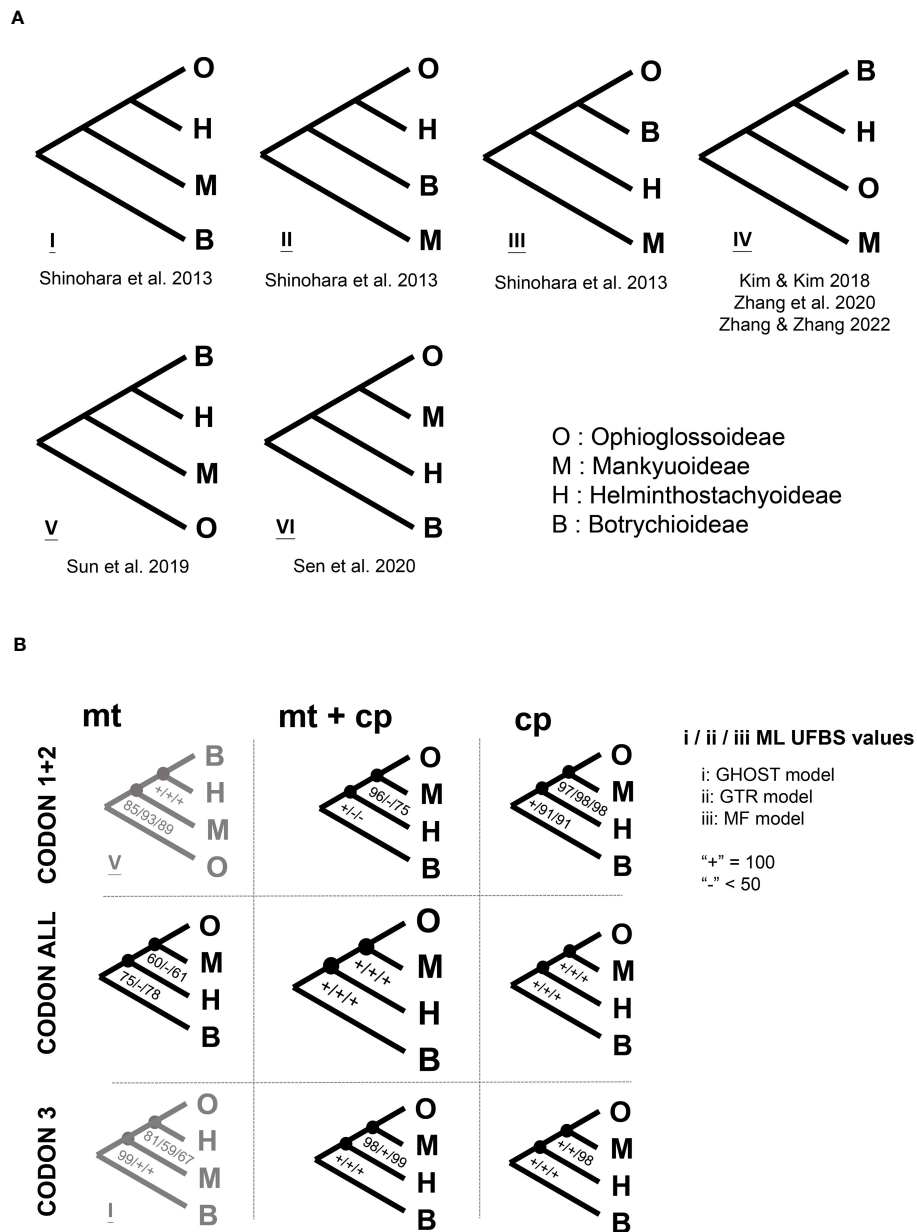


FIGURE 2 The inter-subfamily relationships of Ophioglossaceae. **(A)** Previous phylogenies, the citations for these phylogenies can be found in "References" of the main text. **(B)** The phylogenies inferred in this study using different models and datasets. Details of the different applied models and datasets can be found in "Materials and Methods" of the main text. ML UFBS, maximum likelihood ultrafast bootstrap.

Zhang, 2022). They all applied a rather simplified nucleotide substitution model, which appears to violate the heterogeneous evolution of loci or codon positions, and thus potentially misleads phylogenetic inference (Kapli et al., 2020). Moreover, these phylogenomic analyses included few (only one or two) representatives for the most species rich subfamilies. These incongruences at the subfamily and genus levels also imply a branch attraction issue for these deep divergences in Ophioglossaceae, in which a phylogenetic inference is believed to be sensitive to homoplasy, model violation, and insufficient sampling (Yang and Rannala, 2012; Kapli et al., 2020). In addition to tackling these difficult nodes, supporting evidence

from other genomic features and morphological characters is warranted, and worthy of further investigation.

One of the intriguing findings from Ophioglossaceae is putative intracellular gene transfer (IGT) and horizontal gene transfer (HGT) (Davis et al., 2005; Kim and Kim, 2018; Hao et al., 2022b), and this brings ferns into the earliest foci to study the evolutionary significance of such a genetic mechanism in plants (Li et al., 2014; Li et al., 2018; Wickell and Li, 2020). In particular, angiosperm-to-fern HGTs have been so far discovered only in Ophioglossaceae (Davis et al., 2005). However, for these hypothesized IGTs and HGTs in Ophioglossaceae, their phylogenetic origins or genomic locations remain uncertain. In

the case discovered in *Mankyua*, six open reading frames (ORF) had been identified within the *rps4-trnL* intergenic spacer of the plastome (Kim and Kim, 2018). Some of these ORFs were found in plastomes of other Ophioglossaceae but also mitogenome of Ophioglossaceae and plastome of phylogenetically unrelated ferns (Kim and Kim, 2018). Surprisingly, these ORFs were revealed in different locations in these fern plastomes. In addition, some bacteria genomes were blast-matched with sequences similar to these ORFs. Robison et al. (2018) also found putative HGTs with very similar behaviors from different fern plastomes, and named these genes as Mobile Open Reading Frames in Fern Organelles (MORFFO). Two other HGT cases in Ophioglossaceae are found in *Botrypus* (subfamily Botrychioideae). It was found that two (*matR* and *nad1B-C*) of the four mitochondrial gene regions amplified by PCR were grouped with the Loranthaceae within the angiosperm order Santalales, suggesting the gene sequences might be originated from the root-parasitic Loranthaceae (Davis et al., 2005). The genomic origin and location of these *matR* and *nad1* HGTs in *Botrypus* are not yet confirmed. Lacking comprehensive examination of all genomic parts and close relatives, it remains difficult to identify all these putative IGTs and HGTs in Ophioglossaceae and reconstruct a detailed evolutionary scenario for these genes.

In this study, we first aimed to clarify relationships among the four subfamilies and all 15 genera within Ophioglossaceae (Figure 1). To infer these deepest divergences, a phylogenomic dataset was compiled by sampling additional representatives from these subfamilies and incorporating loci from not only plastomes but also mitogenomes. Moreover, we applied considerable substitution models for phylogenomic inference to minimize systematic errors that may result from a model simplification or a model violation in a phylogenomic dataset. Next, we aimed to infer the phylogenetic origins of putative IGTs and HGTs reported in Ophioglossaceae. We searched their sequences among Ophioglossaceae genome assemblies representing different genomic parts (i.e., plastome, mitogenome, or nuclear genome) in order to verify their genomic origins. By blasting broadly against the nucleotide collection of GenBank and our genome assemblies, we then gathered all highly matched sequences into the phylogenetic surveys of these putative IGTs and HGTs. Moreover, the phylogenomic backbone inferred here provides a solid basis for us to track the evolutionary origins of these putative IGTs and HGTs in Ophioglossaceae as well as the morphological features and genomic changes in this family.

2 Materials and methods

2.1 Sampling and assembling organellar genomes

In total, we included 20 taxa for phylogenomic analyses, covering all four subfamilies and 15 genera in Ophioglossaceae (PPG I, 2016; Zhang and Zhang, 2022) as well as four outgroup taxa. Nine Ophioglossaceae collections were first sequenced in this study, and the others were from previous works (Table S1). To

generate new genome skimming data, the DNAs were first extracted using a modified CTAB protocol (Kuo, 2015), and then fragmented in a Covaris S2 ultrasonicator (Covaris, Woburn, MA, USA) with a size range of 300~400 bp if found to be on average longer than this size. The fragmented DNAs were input for illumina library construction using a NEBNext Ultra II DNA Library Prep Kit for Illumina (New England Biolabs, Ipswich, MA, USA), and following the industry manual. They were finally sequenced using NovaSeq or HiSeq 150 PE (illumina, San Diego, CA, USA) with about 3~10 Gbp per sample.

Except for *Ophioglossum californicum* Prantl and the outgroups, which were already published with both complete plastomes and mitogenomes (Table S1), we assembled both organellar genomes of the other samples. The illumina reads were first trimmed using fastp (Chen et al., 2018) with the default settings, and the trimmed reads were input for organellar genome assembling using NOVOplasty (Dierckxsens et al., 2017; Dierckxsens et al., 2020) with K=39. The plastomes were first assembled with seeds of conspecific *rbcl* sequences downloaded from GenBank. In some samples, other plastid genes were used as seeds for assembly when these genes were missing from the first version of plastome assembly. To fill gaps among the plastome contigs of *Botrypus*, Sanger sequencing was used, and the PCR primers are detailed in Table S2. For the mitogenome assembling of each sample, we used the final plastome assembly from the same collection which could prevent chimeric assemblies with plastid reads (Dierckxsens et al., 2020), and conducted multiple NOVOplasty runs each of which used a different mitochondrial gene as a seed. These seed sequences covered all coding genes (CDS) known from the complete mitogenome sequences of close relatives—*Ophioglossum* and *Psilotum* (Guo et al., 2017). For each sample, we pooled mitogenome assemblies from different NOVOplasty runs, and removed identical sequences prior to annotation.

For gene annotation of both organellar genomes, we used Geneious (Kearse et al., 2012) and published organellar genomes of *Ophioglossum* and *Psilotum* as references (Table S1). The parameter of similarity was set to 80% and 60% for plastid and mitochondrial genes. Because of the highly repeated nature of fern mitogenomes (e.g., Feng and Wicke, 2022; Zumkeller et al., 2023), an assembling approach using short reads alone was unlikely to assemble a complete fern mitogenome. Therefore, we additionally carried out *de novo* assembling for all Ophioglossaceae samples using SOAPdenovo2 (Luo et al., 2012) and the trimmed reads were used as input with a K-mer setting of 29. tblastn function of blast-2.10.0 + (Camacho et al., 2009) was used to further inspect the presence of mitochondrial CDS among these NOVOplasty and SOAP assemblies.

Each organellar CDS was then aligned under a codon model by MACSE v2.03 (Ranwez et al., 2011) using a setting of “-max_refine_iter 3 -local_realign_init 0.3 -local_realign_dec 0.2”. We manually checked these resulting alignments, made adjustments if necessary, and removed all ambiguously aligned sectors in the ends of these alignments. For each of them, we finally reconstructed a preliminary gene tree to confirm homology/origin of each sequence. The sequences found to phylogenetically behave like potential horizontal transfers (e.g., not grouping with

other Ophioglossaceae sequences) were excluded from the alignments for our phylogenomic analyses.

2.2 Phylogenetic analyses

A total of 37 mitochondrial and 83 plastid CDS alignments were concatenated, and then compiled into nine datasets for our phylogenomic analyses, including mitochondrial (mt), plastid (cp), and mitochondrial + plastid (mt + cp). Each of them contained three datasets: with (1) the first two codon positions (CODON 1 + 2), (2) the third codon position (CODON 3), and (3) all three codon positions (CODON ALL). For the mitochondrial part, *ccm* genes were not included because they were found lacking in the mitogenome assembly of certain Ophioglossaceae lineages (Guo et al., 2017; this study). The *Holubiella* sample was excluded from our mitochondrial datasets, because we recovered only a few mitochondrial CDS genes from its assembly. The reads of this herbarium collection were relatively short due to its highly fragmented DNAs, and thus resulted in a poor assembly for the mitogenome. The presence or absence of these CDS in the datasets was summarized in Table S3. IQ-TREE v2.1.3 was used to perform our phylogenomic analyses (Nguyen et al., 2015; Minh et al., 2020). For every dataset, we conducted three maximum likelihood (ML) analyses with different models: (1) MF: the best partition scheme and substitution rates inferred by ModelFinder (Kalyaanamoorthy et al., 2017) with “rcluster = 100” and Bayesian information criterion, (2) GTR: the finest partitions every with a GTR+F+R10 model, (3) GHOST: the “General Heterogeneous evolution On a Single Topology” model (Crotty et al., 2020) with six unlinked GTR classes (i.e., GTR*H6). In the first two kinds of analyses, the datasets were first partitioned by codon positions by genes. Each of these phylogenetic analyses was conducted with 1000 ultrafast bootstrap (UFBS) replicates (Hoang et al., 2018). In addition, we analyzed gene and site concordance factors (gCF and sCF; Minh et al., 2020) based on the mt + cp CODON ALL dataset.

2.3 Inferring origins of intracellular and horizontal gene transfer

For the sequences suspected to be IGTs or HGTs, we first blasted (blastn and tblastn) them against the nucleotide collection in GenBank and our organellar genome assemblies to explore their possible origins. We then selected sequences that were best matched or the most closely related ones shown in a NCBI BLAST distance tree, and aligned them with our mitochondrial matrices. In addition, we found several mitochondrial HGTs to have an angiosperm origin, especially Santalales. Most other mitochondrial HGTs appeared to have non-Ophioglossaceae fern origins. In order to gain a deeper understanding of their phylogenetic origins, we expanded our mitochondrial matrices to include more representatives from angiosperms, in particular for Santalales (Table S4). Besides including all published mitogenomes of this angiosperm order, we also included sequences of certain

mitochondrial genes (e.g., *matR* and *nad1B-C*) to cover additional families and genera. By these expanded mitochondrial alignments, we reconstructed their individual gene and locus-concatenated ML phylogenies using the MF model (details same as in earlier) using IQ-TREE. In the expanded alignment of *matR*, several sequences appeared to be a pseudogene. Therefore, we realigned these DNA metrics based on a nucleotide mode using MUSCLE (Edgar, 2004) implemented in AliView (Larsson, 2014). As a result, we set only a single partition (i.e., not to partition it with different codon positions) for the phylogenetic reconstruction with this gene. For *nad1B-C*, we used the same approach for its alignment and phylogenetic analysis, because it was composed mostly of the non-coding region–intron 2. In addition, due to the difficulty of aligning sequences across different vascular plant lineages, we included only angiosperm (i.e., close relatives of the donor) sequences for our phylogeny of this HGT.

To further explore whether these HGT-like sequences were located in mitogenome or not, we also evaluated the read coverages of these HGT-like sequences and compared them with those of the mitochondrial CDS genes. In practice, we input the exon sequences of the mitochondrial CDS that were longer than 100 bp, and also these HGT-like sequences for read mapping using BWA v0.7.17 (Li and Durbin, 2009). We calculated the mean depths of all sequences using the SAMtools function “coverage” (Danecek et al., 2021). By our genome skimming approach of < 10 Gbp per sample, sequence depth of a nuclear gene in these Ophioglossaceae taxa was theoretically far less than two due to their considerably large sizes of nuclear genomes (Table 1). Conversely, finding a gene with a higher sequence depth indicated its location in an organellar genome. Once we obtained completed plastome assembly, we could further verify whether such a gene was located in plastome or not. If not found in the plastome, we considered it a mitochondrial gene.

3 Results

3.1 Genome assembly and gene content

We newly assembled ten plastomes in this study, and the information about their GenBank and SRA accessions is provided in Table S1. Each of these plastomes was assembled into a complete and circularized contig, except for that of *Botrypus* and *Holubiella*. These two were composed by one and five linear contigs, respectively. Nonetheless, no plastid genes were missing from their assembly. Notably, compared with the previous plastome by Zhang and Zhang (2022) (GenBank accession: OM897597), our plastome of *Whittieria* was fully assembled without gaps nor sectors of ambiguous bases despite that the same read source was input. Except for the inversion of *trnT*-GGU in *Sceptridium* and *Holubiella*, we found no structural difference among these Ophioglossaceae plastomes, and their gene boundaries between inverted repeat (IR) and large or small single copy regions (LSC or SSC) were also identical. Losses of *psbM*, *trnA*-UGC, *clpP* 2nd intron, and *rpl2* intron were identified in genera of subfamily Ophioglossaceae (Table S5).

Our mitogenome assemblies were rather fragmented and composed of numerous linear contigs from 47 in *Cheiroglossa* to 6,749 in *Ophioderma*. In every sample, mitochondrial genes could be found in multiple contigs that slightly differed in their non-coding sequences (Supplementary File S1). Notably, we have further confirmed the loss of introns in several CDS genes (Table S5) as well as the loss of *ccm* genes in subfamilies Ophioglossoideae, Mankyuoideae, and Helminthostachyoideae (Table S5, Figure 3). Slightly differing from Guo et al. (2017), the presence of *ccmC* gene was also confirmed in *Botrypus* and *Sceptridium*. Assembly with annotations is also supplied in Supplementary File S1.

3.2 Phylogenomic relationships

Our gene alignments (before and after trimming) are provided in Supplementary File S2. We summarized our phylogenomic results in Figures 2, 3, S1. Our reconstructions revealed a solid backbone at the subfamily-level (Figure 2B), and supported well (i.e., ML UFBS values from 90 to 100) the topology of (Botrychioideae, (Helminthostachyoideae, (Mankyuoideae, Ophioglossoideae))). These relationships were only weakly supported in the mt datasets (Figure 2B). Notably, this subfamily topology (i.e., topology VI in Figure 2A) had not been recovered in previous works, except for Shen et al. (2020) (Figure 2A). Within Botrychioideae, one of the two non-monotypic subfamilies, the inter-generic relationships were stable and well resolved with high branch supports (Figure 3; data not shown). These relationships are

also congruent with previous phylogenies (Hauk et al., 2003; Sun et al., 2009; Shinohara et al., 2013; Zhang et al., 2020; Zhang and Zhang, 2022). In contrast, the inter-generic relationships in another non-monotypic subfamily, Ophioglossoideae, remain poorly resolved (Figures 3, S1). The seven Ophioglossoideae genera fell into four clades: *Goswamia* (G), *Rhizoglossum* + *Whittieria* (R-W), *Cheiroglossa* + *Ophioderma* (C-O), *Haukia* + *Ophioglossum* (H-O). Although the monophyly of each of these clades is highly supported, the relationships between them remain unclear (Figures 3, S1). Based on our phylogenomic analyses, we totally identified eight different topologies representing their relationships (Figure S1).

3.3 Gene sequences with putative IGT or HGT origins

Among six unknown ORFs in the *Mankyua* plastome (Kim and Kim, 2018), we confirmed that three of them are actually homologs of MORFFO genes that were first identified by Robison et al. (2018). ORF 295 belongs to *morffo1*, while ORFs 531 and 187 are parts of *morffo2*. These MORFFO sequences were found in not only the plastomes but also in the mitogenomes of land plants (Figures S2, S3). In addition, we had identified these genes in many other Ophioglossaceae genera, including *Sahashia*, *Botrypus*, *Botrychium*, *Japanobotrychium*, *Whittieria*, *Rhizoglossum*, *Cheiroglossa*, *Goswamia*, and *Helminthostachys*. Interestingly, *morffo2* could also be found in non-fern organellar genomes, including those from

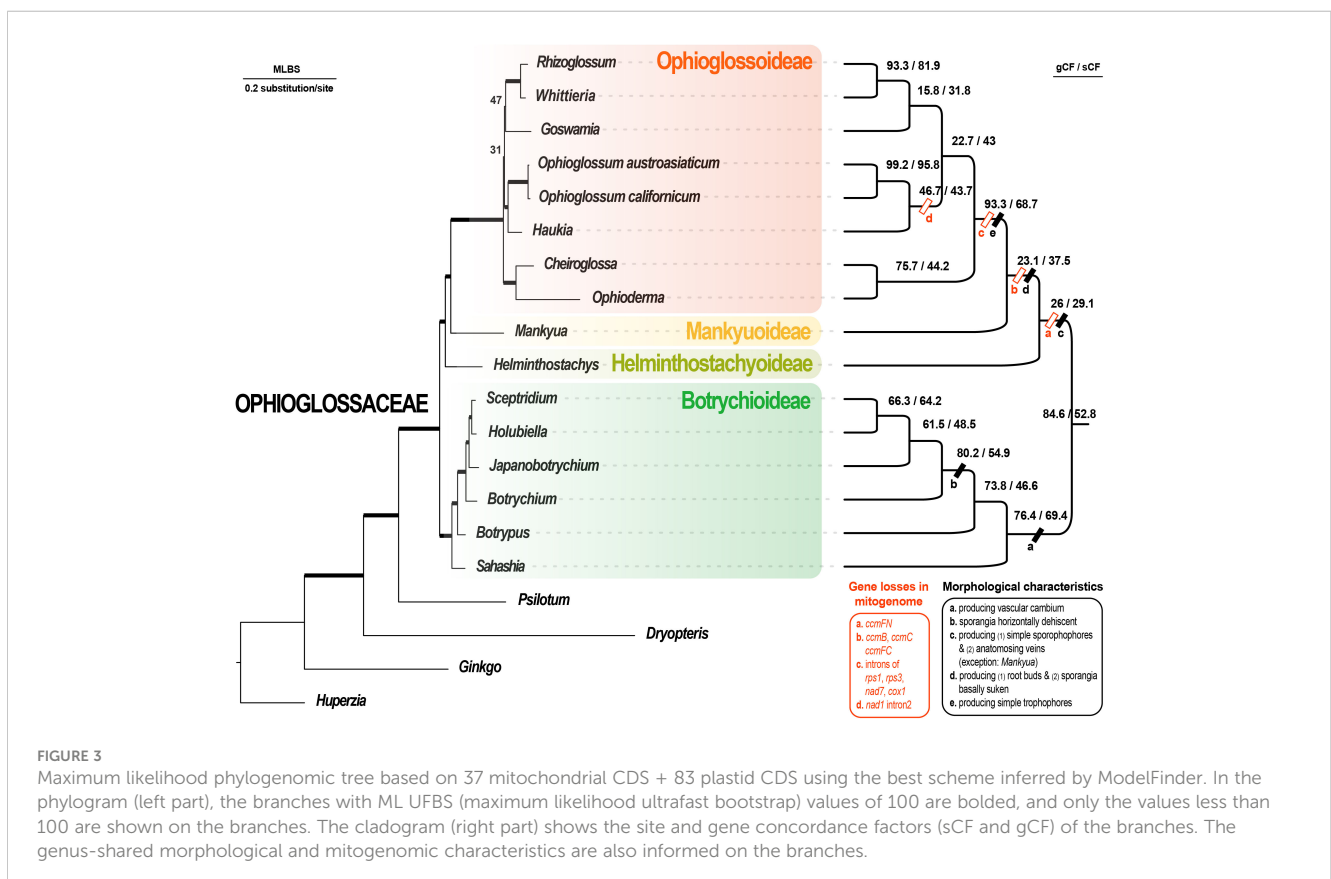


FIGURE 3 Maximum likelihood phylogenomic tree based on 37 mitochondrial CDS + 83 plastid CDS using the best scheme inferred by ModelFinder. In the phylogram (left part), the branches with ML UFBS (maximum likelihood ultrafast bootstrap) values of 100 are bolded, and only the values less than 100 are shown on the branches. The cladogram (right part) shows the site and gene concordance factors (sCF and gCF) of the branches. The genus-shared morphological and mitogenomic characteristics are also informed on the branches.

glaucophytes, green algae, bryophytes, lycophytes, and gymnosperms (Figure S3). The phylogenies of *morfo1* and *morfo2* are shown in Figures 4, 5 with much more details in Figures S2, S3. However, none of Ophioglossaceae species nor any other species had been found with homologs in both its plastome and mitogenome (Figures 4, 5, S2, S3). In other words, no evidence indicating an IGT can be found in our MORFFO phylogenies. On the other hand, MORFFO sequences from different fern families and other plant lineages are usually mixed in our phylogenies (Figures S2, S3). Such mixture patterns imply frequent HGT among these plant lineages, particularly fern species. ORFs 135, 372, and 436 seem to be specific to the plastome of *Mankyua*. Besides its plastome, we found no confident blast hit for these ORFs through the GenBank's nucleotide collection and our genome assemblies, except for ORF-436-like sequence found in the *trnT-trnfM* intergenic region in the plastome of *Ophioglossum californicum*, which is also revealed by Kim and Kim (2018). For ORF 135, we found only BLAST matches of short fragments (< 50% of query length, and < 150 bp) from other organisms.

In addition to the cases of *matR* and *nad1B-C* reported in *Botrypus* (Davis et al., 2005), we found more mitochondrial HGTs that originated from either angiosperm species or ferns based on our phylogenetic surveys (Table 2, Figures 6, S4-23). We also discovered such mitochondrial HGTs in many other Ophioglossaceae genera, including *Helminthostachys*, *Mankyua*, *Sahashia*, and *Scepstridium*, but none of them was from subfamily Ophioglossaceae (Table 2). Moreover, several of these HGTs were identified in clusters on the same contigs, and some were even found to be linked with a host mitochondrial genes (i.e., from the Ophioglossaceae taxon of the sample) (Table 2). In the case of *nad1B-C* (i.e., intron 1 with partial exons 2 and 3) HGT in *Botrypus*, we didn't recover this HGT successfully from our mitogenome nor

from plastome NOVOplasty assemblies, but we did find it in a SOAP contig. The read coverages (i.e., sequence depth) of mitochondrial genes ranged from 1.0 to 29.3 per Gbp (Table S6). Except for the unusually high coverage of *nad7* exon3 HGT in *Sahashia*, the coverage values of these mitochondrial HGTs also fell into this range (Table 2), and support that they are physically located in mitogenomes. However, more than half of these HGTs are pseudogene-like containing frame-shift mutations (Table 2).

Furthermore, we concatenated the sequences of the HGTs from the same contigs in order to reconstruct their phylogenies and to infer their origins with better resolutions. Both cases in *Botrypus* and one in *Helminthostachys* were shown to be nested in the angiosperm order Santalales (Figures 6A–C). Another case in *Helminthostachys* is sister to *Carica* (order Brassicales) (Figure 6D), and the case of *Sahashia* is nested in the angiosperm order Lamiales (Figure 6E).

4 Discussion

4.1 Phylogenomic relationships within Ophioglossaceae

Our previous understanding of Ophioglossaceae phylogeny had been based on molecular phylogenetic analyses using limited plastid loci. Although infrafamilial relationships were inferred by plastomic datasets in many recent phylogenies, those studies sampled limited generic representatives and applied rather simplified substitution models in their phylogenomic analyses (Sun et al., 2009; Kim and Kim, 2018; Shen et al., 2020; Zhang and Zhang, 2022). Those previous phylogenies also revealed conflicting relationships

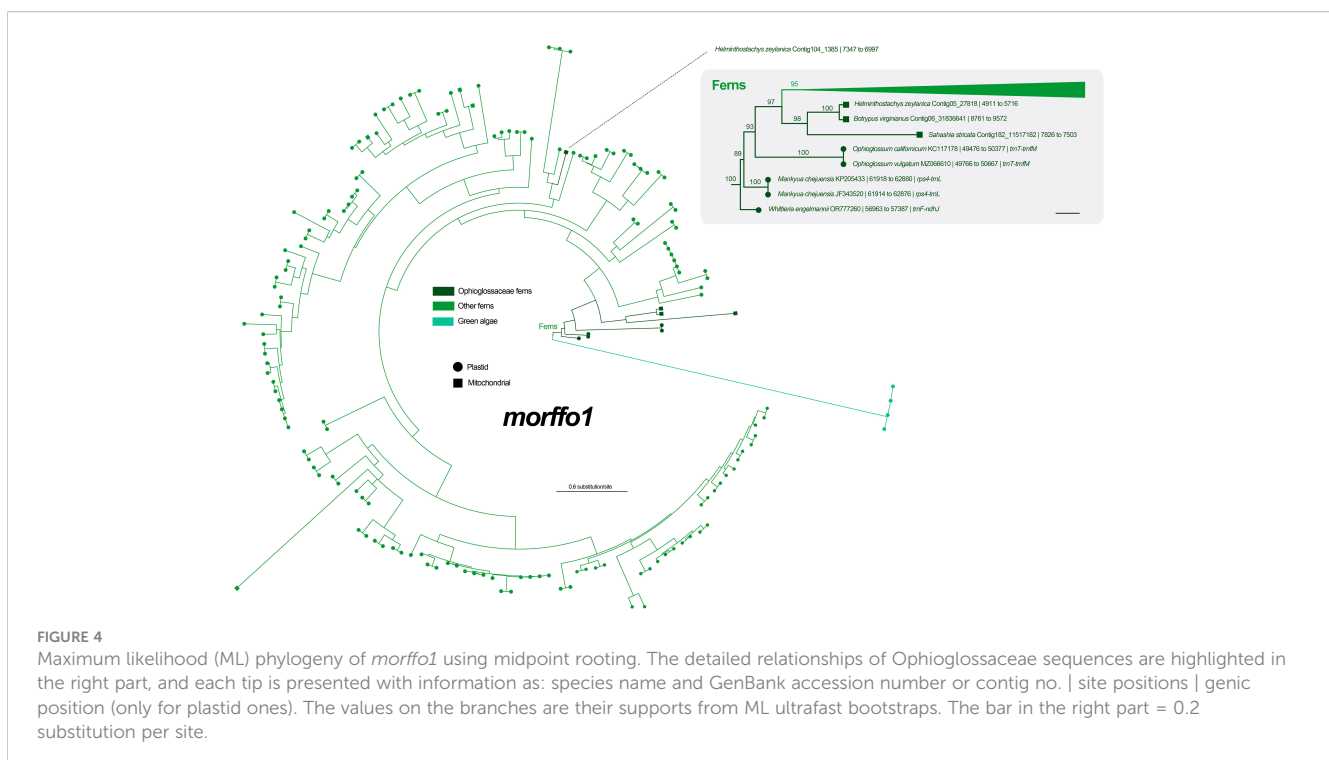


TABLE 2 Mitochondrial HGT genes found in Ophioglossaceae mitogenomes.

| Taxon | Gene | Read coverage per Gbp (total coverage) | HGT origin | Close relative based on ML-tree | Pseudogene-like | Native homolog |
|-------------------------|------------------------------------------------|----------------------------------------|------------|-----------------------------------------------------------------|-----------------|----------------|
| <i>Botrypus</i> | <i>ccmC</i> ^a | 9.00 (26.05) | angiosperm | Santalales (sister to <i>Tolypanthus/Helicanthus/Taxillus</i>) | yes | yes |
| <i>Botrypus</i> | <i>rpl16</i> ^a | 14.33 (41.45) | angiosperm | Santalales (<i>Santalum/Comandra</i>) | yes | not found |
| <i>Botrypus</i> | <i>rps4</i> ^a | 9.45 (27.34) | angiosperm | Santalales (sister to <i>Tolypanthus/Helicanthus</i>) | yes | yes |
| <i>Botrypus</i> | <i>nad5</i> exon1 ^a | 13.44 (38.89) | angiosperm | unresolved | yes | yes |
| <i>Botrypus</i> | <i>nad5</i> exon2 ^a | 9.76 (28.23) | angiosperm | Santalales | no | yes |
| <i>Botrypus</i> | <i>ccmFN</i> ^a | 12.43 (35.96) | angiosperm | Santalales (week support) | yes | yes |
| <i>Botrypus</i> | <i>nad2</i> exon4 ^{a, b} | 10.04 (29.04) | angiosperm | Santalales (<i>Comandra</i>) | yes | yes |
| <i>Botrypus</i> | <i>nad6</i> ^{a, b} | 7.65 (22.13) | angiosperm | Santalales + <i>Vitis</i> | yes | yes |
| <i>Botrypus</i> | <i>matR</i> ^{a, b} | 10.51 (30.39) | angiosperm | Santalales (sister to <i>Schoepfia</i> + <i>Arjona</i>) | yes | yes |
| <i>Botrypus</i> | <i>rps19</i> ^{a, b} | 18.49 (53.50) | angiosperm | unresolved | no | yes |
| <i>Botrypus</i> | <i>nad1</i> intron2 (with partial exons 2 & 3) | 9.58 (27.72) | angiosperm | Santalales (<i>Schoepfia</i>) | yes | yes |
| <i>Helminthostachys</i> | <i>atp6</i> ^c | 4.95 (22.94) | angiosperm | Santalales | no | yes |
| <i>Helminthostachys</i> | <i>rps4</i> ^c | 4.99 (23.12) | angiosperm | Santalales (<i>Malania</i>) | yes | yes |
| <i>Helminthostachys</i> | <i>nad6</i> ^c | 4.89 (22.67) | angiosperm | unresolved | no | yes |
| <i>Helminthostachys</i> | <i>atp9</i> ^d | 10.88 (50.41) | angiosperm | unresolved | no | yes |
| <i>Helminthostachys</i> | <i>nad5</i> exon1 ^d | 2.33 (10.79) | angiosperm | unresolved | no | yes |
| <i>Helminthostachys</i> | <i>rpl16</i> ^e | 5.12 (23.72) | angiosperm | Santalales (<i>Olax</i>) | yes | yes |
| <i>Helminthostachys</i> | <i>atp9</i> | 11.22 (52.00) | fern | non-Ophioglossaceae ferns | no | yes |
| <i>Mankyua</i> | <i>rpl5</i> ^e | 3.62 (59.75) | fern | Ophioglossaceae ferns (<i>Sceptridium</i>) | no | yes |
| <i>Mankyua</i> | <i>ccmB</i> ^e | 4.11 (67.92) | fern | Ophioglossaceae ferns (<i>Sceptridium</i>) | no | not found |
| <i>Mankyua</i> | <i>nad2</i> exon4 | 5.03 (83.11) | angiosperm | unresolved | yes | yes |
| <i>Mankyua</i> | <i>rps4</i> | 4.29 (70.95) | angiosperm | <i>Liriodendron</i> | no | yes |
| <i>Sahashia</i> | <i>ccmC</i> ^f | 17.92 (62.40) | angiosperm | unresolved | yes | yes |
| <i>Sahashia</i> | <i>rps12</i> ^f | 10.76 (37.46) | angiosperm | Lamiales (<i>Melampyrum</i>) | no | not found |
| <i>Sahashia</i> | <i>nad3</i> ^f | 10.25 (35.69) | angiosperm | Lamiales (<i>Melampyrum</i>) | yes | not found |
| <i>Sahashia</i> | <i>nad5</i> exon2 + exon3 + exon5 + exon6 | 7.92 (27.59) [based on exon2] | angiosperm | Lamiales (<i>Castilleja</i>) | no | yes |
| <i>Sahashia</i> | <i>cox3</i> | 13.34 (46.45) | fern | non-Ophioglossaceae ferns | yes | not found |
| <i>Sahashia</i> | <i>nad7</i> exon1 + exon2 + exon3 | 371.67 (1294.06) [based on exon3] | fern | leptosporangiate ferns | yes | not found |
| <i>Sahashia</i> | <i>nad7</i> exon4 + exon5 | 15.99 (55.68) | fern | leptosporangiate ferns | yes | not found |
| <i>Sceptridium</i> | <i>ccmC</i> [*] | 4.83 (21.78) | angiosperm | unresolved | no | yes |

*Linked with a host mitochondrial gene.

^aon the contig of "ccmC-rpl16-rps4-nad5-ccmFN".

^bon the contig of "nad2-nad6-nad1e4-matR-rps19".

^con the contig of "atp6-rps4-nad6".

^don the contig of "atp9-nad5".

^eon the contig of "rpl5-ccmB".

^fon the contig of "ccmC-rps12-nad3".



FIGURE 6
 Locus-concatenated phylogenies for the HGT genic clusters found in Ophioglossaceae mitogenomes. **(A)** *rps12 + nad3* in *Sahashia*. **(B)** *ccmC + rpl16 + rps4 + nad5 + ccmFN* **(C)** *nad2 + nad6 + matR + rps19* in *Botrypus*. **(D)** *atp9 + nad5* **(E)** *atp6 + rps4 + nad6* in *Helminthostachys*. The values below the branches are their ML UFBS, and only values larger than 50 are shown. The branches with supporting values of maximum likelihood ultrafast bootstrap ≥ 80 are bolded. “+” = 100.

4.2 Evolutionary origins of MORFFO genes

MORFFO genes were first identified by Robison et al. (2018) in many fern plastomes, and *morffo1* and *morffo2* were also discovered in Ophioglossaceae plastomes by Kim and Kim (2018) who reported them as ORF 295 and ORFs 531 and 187, respectively. Kim and Kim (2018) also hypothesized their origins in Ophioglossaceae by HGT and/or IGT. In the current study, we blast-searched these MORFFO sequences through both plastome and mitogenome assemblies across different Ophioglossaceae genera. We indeed recovered mitochondrial MORFFO genes from

our assemblies (Figures 4, 5, S2, S3). Together with blast-matched sequences from the GenBank nucleotide collection, our phylogenies, however, reveal no direct relationships between MORFFO sequences in a plastome and a mitogenome in any certain species (Figures 4, 5, S2, S3). Specifically, for either *morffo1* or *morffo2*, we didn’t find its presence in both organellar genomes in the same Ophioglossaceae collection. In other words, our study provided no supporting evidence for IGT of a MORFFO gene that could switch between organellar genomes in Ophioglossaceae, although such cases are likely true for other genic regions (e.g., plastid-derived DNA fragments in

mitochondrial genomes; Guo et al., 2017; Hao et al., 2022b). However, the mitogenome assemblies of most Ophioglossaceae genera are still incomplete, and thus mitochondrial MORFFO might be overlooked. Besides, IGT might also occur between organellar and nuclear genomes (see below). Complete mitogenome or even nuclear genome sequences from different Ophioglossaceae genera in future may shed light on the origins of these MORFFO genes.

Regardless of mitochondrial or plastid origins, MORFFO sequences from Ophioglossaceae are usually grouped together (Figures 4, 5, S2, S3). Similar phylogenetic patterns can be found in MORFFOs of other fern families, too (Figures S2, S3). In addition, relationships inside these family clusters are sometimes consistent with their species tree (e.g., FTOL; Nitta et al., 2022). However, such closely related MORFFO sequences are usually found in different genic locations. For instance, in the Ophioglossaceae cluster with only plastomic *morffo2*, these sequences are in various plastomic locations, including intergenic regions of *rps4-trnL*, *trnF-ndhI*, and *trnT-trnFM* in LSC, and *ndhF-rpl21* in SSC (cluster III in Figure 5). In addition, some plastomes host two diverged copies, such as *morffo2* in *Helminthostachys* (Figure 5). These patterns imply that MORFFOs can be inherited vertically, and replicate to switch their positions. One plausible scenario to explain their varied genic distributions among close relatives but with a moderate sequence similarity is that these organellar MORFFOs are derived from IGT of different nuclear duplicates. We highlight this possibility for some fern cases because several *morffo2*-like sequences were also found in the nuclear genome of the tree fern *Alsophila spinulosa* (Hook.) R.M.Tryon (Figures 5, S3), which was recently whole-genome-sequenced (Huang et al., 2022).

In this study, we also inferred phylogenetic relationships for *morffo1* and *morffo2* by incorporating a broad sequence sampling from the GenBank nucleotide collection (Figures 4, 5, S2, S3). Notably, we found MORFFO sequences not only in fern plastomes, but also in their mitogenomes, organellar genomes of other non-angiosperm green plant lineages, and genomes of viruses and phages. In these MORFFO phylogenies, fern sequences are the most dominant ones, and they also form a major clade (Figures 4, 5). Interestingly, despite the fact that sequences from the same family always clustered together, we didn't find an interfamily relationship highly matching our current understanding of the fern tree of life (Figures S2, S3). Moreover, in some highly supported clades, sequences from a distantly related family can be found nested within. For example, two of three Ophioglossaceae sequence clusters of *morffo2* are placed in the basal positions of the fern-dominant lineage (clusters II and III in Figure 5), but another one is imbedded in sequences of leptosporangiate ferns, and also contains a sequence from the leptosporangiate family Pteridaceae (i.e., *Myriopteris scabra* (C.Chr.) Grusz & Windham; in cluster I in Figure 5). These apparently random interfamily relationships imply the possibility of frequent fern-to-fern HGTs for the MORFFO genes. We also noticed that similar phenomena are revealed in some HGT cases reported previously in ferns (i.e., *PHY3* and *Tma12*) but they are found in nuclear genomes instead (Li et al., 2014; Li et al., 2018).

4.3 Origins of mitochondrial HGTs in Ophioglossaceae

Due to the recent increase in the description of fern mitogenome sequences, more HGT cases have been identified from these mitogenomes (e.g., Zumkeller et al., 2023). In Ophioglossaceae, besides two previously identified cases (i.e., *matR* and *nad1* in *Botrypus*), we identified more mitochondrial HGTs in more genera (Table 2). We also confirmed their presence in multiple conspecific samples (data not shown), so excluding the possibility of individual sample contamination. These HGTs are very likely located in mitogenomes because their read coverages match those of other mitochondrial genes (Table 2, S6). Moreover, some of them are linked with mitochondrial genes from Ophioglossaceae hosts (Table 2). Phylogenetically, these HGTs in Ophioglossaceae mitogenomes are derived from either angiosperm or fern donors. In some species, we can find HGTs with multiple origins (Table 2). Interestingly, among the cases with angiosperm origins, Santalales species were found to be the major donor, and root-parasitic species were most likely ones as HGT sequences are all grouped with root-parasitic lineages such as *Comandra* and *Santalum*, *Arjona* and *Schoefpia*, and *Malania*. In *Sahashia*, two HGT genes are likely derived from a different root-parasitic lineage in the angiosperm order Lamiales—Orobanchaceae (Figure 6A). These findings imply that several Ophioglossaceae taxa live intimately with root-parasitic angiosperms during some stages of their life history. Despite these angiosperm parasites not having been reported to directly host on a fern, these angiosperms are possible to contact with ferns indirectly via mycorrhizal fungi (Brundrett, 2002; Heide-Jørgensen, 2010; de Vega et al., 2011). Such a scenario is plausible especially for Ophioglossaceae ferns because they are always symbiotic with mycorrhizal fungi in their life history that starts from a mycoheterotrophic gametophyte generation to a sporophyte generation relying on mycorrhizae (Lang, 1902; Winther and Friedman, 2007; Whittier, 2015; Chen et al., 2022). In addition, two mitochondrial HGTs in *Mankyua* were found to originate from other Ophioglossaceae ferns, most likely *Sceptridium* in subfamily Botrychioideae; and the others were from non-Ophioglossaceae ferns (Table 2; Figure S4). Fern-to-fern HGTs were also proposed in previous studies, and gametophytes of these plants are suggested to be the most vulnerable stage of being genetically “transformed” (Wickell and Li, 2020). However, it is still difficult to explicitly determine the taxon identity of these donors because fern mitochondrial sequences remain scarce in GenBank and thus only limited fern taxa are analyzed in our mitochondrial phylogenies. Identifying the donors of these HGTs in the future will be very insightful to study mechanisms behind these fern-to-fern and angiosperm-to-fern HGTs.

The unexpectedly common HGTs in Ophioglossaceae raise another question: Do these foreign genes provide any evolutionary advantages or disadvantages to ferns? *PHY3* and *Tma12* are two typical HGT cases in ferns that are usually interpreted to significantly benefit the adaptation of these plants (Li et al., 2014; Li et al., 2018). However, most mitochondrial HGTs presented here appear to be neutral for their Ophioglossaceae hosts, because they are pseudogene-like and/or redundant copies in the

host mitogenomes (Table 2). Nonetheless, two HGTs, the *ccmB* in *Mankyua* and the *rps12* in *Sahashia*, may potentially complement the functional losses of native copies in the host mitogenomes (Table 2). Another intriguing case is the *atp9* copies in the *Helminthostachys* mitogenome. In addition to the native copy, *Helminthostachys* acquired two HGT copies respectively from a non-Ophioglossaceae fern and an angiosperm, and both foreign copies appear to be functional. Further assessment of their expression profile would be helpful to answer whether these HGTs function like beneficial genes.

4.4 Evolutionary trends of organellar genomics in subfamily Ophioglossoidae

One of the notable evolutionary trends in Ophioglossoidae is its acceleration in changes in organellar genomes. From our phylogenomic tree, an obvious rate elevation at the DNA level can be found in Ophioglossoidae, which exhibits longer branches than other lineages (Figure 3). Moreover, their plastomes have evolved to become AT-rich and reduced in size. The most significant case is the genus *Ophioderma*, and its plastome is the smallest known by far in ferns that is ~0.123 Mbp in size (Table S1; Du et al., 2019; Du et al., 2021; Du et al., 2022). We can find similar trends also in the mitogenomes of these Ophioglossoidae genera. Apparently, these taxa have reduced mitogenomes, in which *ccm* genes and several introns are lost (Figure 3; Table S5), and the total DNA amount is also decreased. Mitogenomes of *Ophioglossum* have been fully assembled that are revealed with a single chromosome of a smaller size of 0.37 Mbp (Guo et al., 2017; Hao et al., 2022b). By contrast, mitogenomes from other ferns, including *Psilotum* from the sister family, are typically with multiple chromosomes and total lengths ranging from 0.62 to 1.44 Mbp (Guo et al., 2017; Feng and Wicke, 2022; Zumkeller et al., 2023). In the other Ophioglossaceae subfamilies, the mitogenome sizes seem to be even larger, and that in *Helminthostachys* and *Scepstridium* are estimated at greater than 3.19 and 1.16 Mbp based on our preliminary assemblies (Kuo et al., unpublished data). Additionally, mitogenomes of these other subfamilies are occasionally found to be inserted with pseudogenized HGTs (Table 2). Taken together, these findings imply a selective pressure on the organellar genomes in subfamily Ophioglossoidae to reduce in size, a phenomenon not observed in other subfamilies.

We speculate that this evolutionary trend in Ophioglossoidae is associated with (1) its mixotrophic lifestyle and/or (2) extraordinarily large nuclear genomes. The members in this subfamily produce proliferous and budding roots, that are mycorrhizal, fleshy, and rich with storage materials (Petry, 1914; Chrysler, 1941; Peterson and Brisson, 1977). Importantly, these roots can “parasitically” acquire carbohydrates from nearby plants via mycorrhizal fungi (Suetsugu et al., 2020). Thus, when photosynthetic aboveground parts are seasonally dormant, these non-green but active underground organs remain capable of growing and even asexually reproducing. In other words, Ophioglossoidae species likely behave as autotrophic and heterotrophic lifestyles separately in different seasons or

developmental stages (i.e., root buds vs. latterly foliage stages) during their sporophyte generation. During the heterotrophic stages of the life cycle, organellar genomes that are smaller and AT-rich are favored due to their ability to replicate more rapidly. Consequently, these smaller genomes are likely the result of genic deletions, which are beneficial for the plants during this phase. As the plants transition to the autotrophic stages, where they engage in photosynthesis and assimilation, organellar genomes with deficit characteristics due to deletion-introduction are subsequently selected out. Such selection from the heterotrophic stages can be much severe if the constructional needs of a nuclear genome become greater (e.g., need more nucleotides for its genome replication), and thus it competes for resources with organellar genomes. This hypothesis can better explain why organellar genome downsizing is evident only in Ophioglossoidae which is diagnostic by huge nuclear genomes. Especially for *Ophioderma*, which has the largest nuclear genome size in this family (Table 1), and also exhibits the most diverged organellar genomes (Tables S1, S5; Figure 3).

4.5 Conclusion and future perspectives

Using the phylogenomic analyses that apply both plastomic and mitogenomic datasets, sequences from all genera, and fine-tuned substitution models, we provide strong evidence for well-resolved relationships of the subfamilies in Ophioglossaceae. Importantly, this solid subfamily backbone is consistent with many key morphological and genomic changes in the family and supports the systematic positions of Helminthostachyoideae and Mankyuoideae. At the generic level, some relationships within subfamily Ophioglossoidae remain unclear, and its seven genera are sorted into four highly supported clades. To resolve these difficult nodes within Ophioglossoidae, we expect to utilize a larger phylogenomic dataset in the future that includes more generic representatives as well as more loci from the nuclear genomes, such as transcriptomic sequences. In addition, this study also highlights several evolutionary genomics issues in Ophioglossoidae, including its elevated substitution rates in organellar genomes and a trend toward AT-rich and smaller plastomes. A tentative hypothesis is also provided here to explain these interesting patterns in Ophioglossoidae by considering its mixotrophic lifestyle and huge nuclear genomes. In addition, evolutionary rate heterogeneity and AT-biased codon usages in this fern lineage are likely associated with gene expression levels (Hao et al., 2022a). Therefore, future investigation of rate statics and transcriptomic profiles is also critical to study comparative genomics in Ophioglossoidae.

One of our interesting findings here is identifying numerous HGT (or IGT-like) cases in Ophioglossaceae organellar genomes. These HGTs predominantly occur in the mitogenomes of subfamilies Mankyuoideae, Helminthostachyoideae, and Botrychioideae. However, these foreign mitochondrial genes seem to be pseudogenized or functionally redundant for host plants. Notably, the origins of these mitochondrial HGTs can be traced to different vascular plant lineages, and ferns and root-parasitic angiosperms

appear to be the most important donors. These findings also imply that, behind such steady genic “transformations”, these Ophioglossaceae ferns establish intermate and frequent connections with root-parasitic angiosperms and ferns during certain stages in their life history. Studying the underlaid mechanisms of these HGTs is promising in order to explore a novel manner of genic transformation in plants, particularly for their organellar genomes which remains difficult for most plants.

Finally, the current study points out the importance of fern mitogenome sequences in studying various issues of evolutionary genomics in plants, including their phylogenomics and HGT/IGT mechanisms. However, mitogenomes remain understudied across fern diversity. To date, complete or near complete mitogenomic sequences have been described in less than six fern genera (Guo et al., 2017; Song et al., 2021; Feng and Wicke, 2022; Hao et al., 2022b; Zumkeller et al., 2023), and this number is even less than those published for nuclear genomes. This lack of fern mitogenomic data is in part due to the complexity of these genomes in most species, which is rich in long repetitive sequences and even IGT from plastomes. As adopted in the current study, assembling strategies with short reads alone can recover most mitochondrial gene sequences for phylogenomic analyses. However, these approaches cannot produce a complete mitogenome assembly in most ferns. The most promising way to sequence a fern mitogenome is to adapt the long-read NGS approaches. For this, both an efficient protocol of high-molecular-weight DNA extraction (e.g., Xie et al., 2023) and a long-read sequencing strategy (e.g., Nanopore) are indispensable. Therefore, we are looking forward to future attempts of long-read sequencing and assembly of fern mitogenomes, which will facilitate publication of mitogenome sequences across fern diversity.

Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found below: Bioproject accession numbers: PRJNA1019085.

Author contributions

L-YK: Conceptualization, Data curation, Formal Analysis, Funding acquisition, Investigation, Methodology, Project administration, Resources, Software, Supervision, Validation, Visualization, Writing – original draft, Writing – review & editing. H-JS: Conceptualization, Data curation, Formal Analysis, Investigation, Resources, Validation, Visualization, Writing – review & editing. DK: Funding acquisition, Investigation, Software, Writing – review & editing. P-JX: Data curation, Investigation, Writing – review & editing. CW: Resources, Writing – review & editing. AE: Data curation, Resources, Writing – review & editing. JRG: Conceptualization, Funding acquisition, Resources, Supervision, Writing – review & editing.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. The main funding was from National Science and Technology Council in Taiwan (MOST 111-2628-B-007-006-MY3), and the Bioresource Conservation Research Center in College of Life Science from the Higher Education Sprout Project by MOE was granted. JG and DK were funded by the Swiss National Science Foundation grant (31003A_182688). Computational resources used by DK were provided by the e-INFRA CZ project (ID:90254), supported by the Ministry of Education, Youth and Sports of the Czech Republic and by the ELIXIR-CZ project (ID:90255), part of the international ELIXIR infrastructure.

Acknowledgments

We thank Carl Rothfels, Cheng-Wei Chen, Tian-Chuan Hsu for providing their DNA collections of this study and Taiwan Pteridophyte Research Group (TPG) for maintaining DNA collections of this study; Susan Fawcett and the staff in herbarium MICH for the sampling of *Holubiella* for this study; Technology Common (College of Life Science, National Taiwan University, Taiwan) for the assistance with DNA library constructions; Genomics Corp. (New Taipei City, Taiwan) for HiSeq and NovaSeq sequencing; the manuscript editor, Juan Carlos Villarreal A., and three reviewers for comments on the draft of our manuscript; Zhi-Xiang Chang, Tian-Chuan Hsu, Emily B. Sessa, Pi-Fong Lu, Ponpipat Limpanasitticha, Tassanai Jaruwattanaphan for sharing their photos of living plants.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpls.2023.1294716/full#supplementary-material>

References

- Brundrett, M. C. (2002). Coevolution of roots and mycorrhizas of land plants. *New Phytol.* 154, 275–304. doi: 10.1046/j.1469-8137.2002.00397.x
- Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K., et al. (2009). BLAST+: architecture and applications. *BMC Bioinf.* 10, 421. doi: 10.1186/1471-2105-10-421
- Chen, K.-H., Xi, Q., Chang, C., and Kuo, L. (2022). Mycobiome detection from a single subterranean gametophyte using metabarcoding techniques. *Appl. Plant Sci.* 10, e11461. doi: 10.1002/aps3.11461
- Chen, S., Zhou, Y., Chen, Y., and Gu, J. (2018). fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics* 34, i884–i890. doi: 10.1093/bioinformatics/bty560
- Chrysler, M. A. (1941). The structure and development of *Ophioglossum palmatum*. *Bull. Torrey Bot. Club* 68, 1–19. doi: 10.2307/2481251
- Crotty, S. M., Minh, B. Q., Bean, N. G., Holland, B. R., Tuke, J., Jermini, L. S., et al. (2020). GHOST: recovering historical signal from heterotachously evolved sequence alignments. *Syst. Biol.* 69, :249–264. doi: 10.1093/sysbio/syz051
- Danecek, P., Bonfield, J. K., Liddle, J., Marshall, J., Ohan, V., Pollard, M. O., et al. (2021). Twelve years of SAMtools and BCFtools. *Gigascience* 10, 1–4. doi: 10.1093/gigascience/giab008
- Davis, C. C., Anderson, W. R., and Wurdack, K. J. (2005). Gene transfer from a parasitic flowering plant to a fern. *Proc. Biol. Sci.* 272, 2237–2242. doi: 10.1098/rspb.2005.3226
- de Vega, C., Arista, M., Ortiz, P. L., and Talavera, S. (2011). Mycorrhizal fungi and parasitic plants: reply. *Am. J. Bot.* 98, 597–601. doi: 10.3732/ajb.1100036
- Dierckxsens, N., Mardulyn, P., and Smits, G. (2017). NOVOPlasty: *de novo* assembly of organelle genomes from whole genome data. *Nucleic Acids Res.* 45, e18. doi: 10.1093/nar/gkw955
- Dierckxsens, N., Mardulyn, P., and Smits, G. (2020). Unraveling heteroplasmy patterns with NOVOPlasty. *NAR Genomics Bioinforma.* 2, 1–10. doi: 10.1093/nargab/lqz011
- Du, X.-Y., Kuo, L.-Y., Zuo, Z.-Y., Li, D.-Z., and Lu, J.-M. (2022). Structural variation of plastomes provides key insight into the deep phylogeny of ferns. *Front. Plant Sci.* 13. doi: 10.3389/fpls.2022.862772
- Du, X., Lu, J., and Li, D. (2019). Advances in the evolution of plastid genome structure in lycophytes and ferns. *Biodivers. Sci.* 27, 1172–1183. doi: 10.17520/biods.2019113
- Du, X., Lu, J., Zhang, L., Wen, J., Kuo, L., Mynssen, C. M., et al. (2021). Simultaneous diversification of Polypodiales and angiosperms in the Mesozoic. *Cladistics* 37, 518–539. doi: 10.1111/cla.12457
- Edgar, R. C. (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* 32, 1792–1797. doi: 10.1093/nar/gkh340
- Feng, Y., and Wicke, S. (2022). New mitochondrial genomes of leptosporangiate ferns allow modeling the mitogenomic inflation syndrome across all land plant lineages. *bioRxiv*. doi: 10.1101/2022.12.23.521604
- Fujiwara, T., Liu, H., Meza-Torres, E. I., Morero, R. E., Vega, A. J., Liang, Z., et al. (2023). Evolution of genome space occupation in ferns: linking genome diversity and species richness. *Ann. Bot.* 131, 59–70. doi: 10.1093/aob/mcab094
- Guo, W., Zhu, A., Fan, W., and Mower, J. P. (2017). Complete mitochondrial genomes from the ferns *Ophioglossum californicum* and *Ptilotum nudum* are highly repetitive with the largest organellar introns. *New Phytol.* 213, 391–403. doi: 10.1111/nph.14135
- Hao, J., Liang, Y., Ping, J., Li, J., Shi, W., Su, Y., et al. (2022a). Chloroplast gene expression level is negatively correlated with evolutionary rates and selective pressure while positively with codon usage bias in *Ophioglossum vulgatum* L. *BMC Plant Biol.* 22, 580. doi: 10.1186/s12870-022-03960-8
- Hao, J., Liang, Y., Su, Y., and Wang, T. (2022b). The complete mitochondrial genome of *Ophioglossum vulgatum* L. Is with highly repetitive sequences: intergenomic fragment transfer and phylogenetic analysis. *Genes (Basel)*. 13, 1287. doi: 10.3390/genes13071287
- Hauk, W. D., Parks, C. R., and Chase, M. W. (2003). Phylogenetic studies of Ophioglossaceae: evidence from *rbcl* and *trnL-F* plastid DNA sequences and morphology. *Mol. Phylogenet. Evol.* 28, 131–151. doi: 10.1016/S1055-7903(03)00032-0
- Heide-Jørgensen, H. (2010). “Chapter 2. Hemiparasitic santalales,” in *Parasitic flowering plants* (Leiden: BRILL), 25–144. doi: 10.1163/ej.9789004167506.i-438.14
- Hoang, D. T., Chernomor, O., vonHaeseler, A., Minh, B. Q., and Vinh, L. S. (2018). UFBoot2: improving the ultrafast bootstrap approximation. *Mol. Biol. Evol.* 35, 518–522. doi: 10.5281/zenodo.854445
- Huang, C. H., Qi, X., Chen, D., Qi, J., and Ma, H. (2020). Recurrent genome duplication events likely contributed to both the ancient and recent rise of ferns. *J. Integr. Plant Biol.* 62, 433–455. doi: 10.1111/jipb.12877
- Huang, X., Wang, W., Gong, T., Wickell, D., Kuo, L.-Y., Zhang, X., et al. (2022). The flying spider-monkey tree fern genome provides insights into fern evolution and arborescence. *Nat. Plants* 8, 500–512. doi: 10.1038/s41477-022-01146-6
- Kalyaanamoorthy, S., Minh, B. Q., Wong, T. K. F., Von Haeseler, A., and Jermini, L. S. (2017). ModelFinder: fast model selection for accurate phylogenetic estimates. *Nat. Methods* 14, 587–589. doi: 10.1038/nmeth.4285
- Kapli, P., Yang, Z., and Telford, M. J. (2020). Phylogenetic tree building in the genomic age. *Nat. Rev. Genet.* 21, 428–444. doi: 10.1038/s41576-020-0233-0
- Kearse, M., Moir, R., Wilson, A., Stones-Havas, S., Cheung, M., Sturrock, S., et al. (2012). Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics* 28, 1647–1649. doi: 10.1093/bioinformatics/bts199
- Khandelwal, S. (1990). Chromosome evolution in the genus *Ophioglossum* L. *Bot. J. Linn. Soc* 102, 205–217. doi: 10.1111/j.1095-8339.1990.tb01876.x
- Kim, H. T., and Kim, K.-J. (2018). Evolution of six novel ORFs in the plastome of *Mankyua chejuense* and phylogeny of eusporangiate ferns. *Sci. Rep.* 8, 16466. doi: 10.1038/s41598-018-34825-6
- Kuo, L.-Y. (2015). *Polyploidy and biogeography in genus Deparia and phylogeography in Deparia lancea*. in *Ph.D. dissertation* (Taipei: National Taiwan University).
- Kuo, L.-Y., Tang, S. K., Kao, T.-T., Ebihara, A., Fawcett, S., Hsiao, M.-C., et al. (2021). A dormant resource for genome size estimation in ferns: C-value inference of the Ophioglossaceae using herbarium specimen spores. *Appl. Plant Sci.* 9, e11452. doi: 10.1002/aps3.11452
- Lang, W. H. (1902). On the prothalli of *Ophioglossum pendulum* and *Helminthostachys zeylanica*. *Ann. Bot.* 16, 23–56. doi: 10.1093/oxfordjournals.aob.a088870
- Larsson, A. (2014). AliView: a fast and lightweight alignment viewer and editor for large data sets. *Bioinformatics* 30, 3276–3278. doi: 10.1093/bioinformatics/btu531
- Lee, J.-W., Yun, H.-G., Hwang, T. Y., and An, J.-B. (2022). Floristic inventory and distribution characteristics of vascular plants in forest wetlands of South Korea. *Biodivers. Data J.* 10, e85848. doi: 10.3897/BDJ.10.e85848
- Li, F.-W., Brouwer, P., Carretero-Paulet, L., Cheng, S., Vries, J., de, Delaux, P. M., et al. (2018). Fern genomes elucidate land plant evolution and cyanobacterial symbioses. *Nat. Plants* 4, 460–472. doi: 10.1038/s41477-018-0188-8
- Li, H., and Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* 25, 1754–1760. doi: 10.1093/bioinformatics/btp324
- Li, F.-W., Villarreal, J. C., Kelly, S., Rothfels, C. J., Melkonian, M., Frangedakis, E., et al. (2014). Horizontal transfer of an adaptive chimeric photoreceptor from bryophytes to ferns. *Proc. Natl. Acad. Sci. U. S. A.* 111, 6672–6677. doi: 10.1073/pnas.1319929111
- Luo, R., Liu, B., Xie, Y., Li, Z., Huang, W., Yuan, J., et al. (2012). SOAPdenovo2: an empirically improved memory-efficient short-read *de novo* assembler. *Gigascience* 1, 18. doi: 10.1186/2047-217X-1-18
- Minh, B. Q., Schmidt, H. A., Chernomor, O., Schrempf, D., Woodhams, M. D., Von Haeseler, A., et al. (2020). IQ-TREE 2: new models and efficient methods for phylogenetic inference in the genomic era. *Mol. Biol. Evol.* 37, 1530–1534. doi: 10.1093/molbev/msaa015
- Nguyen, L. T., Schmidt, H. A., VonHaeseler, A., and Minh, B. Q. (2015). IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol. Biol. Evol.* 32, 268–274. doi: 10.1093/molbev/msu300
- Nitta, J. H., Schuettpelz, E., Ramirez-Barahona, S., and Iwasaki, W. (2022). An open and continuously updated fern tree of life. *Front. Plant Sci.* 13. doi: 10.3389/fpls.2022.909768
- Patel, M., and Reddy, M. N. (2018). Discovery of the world’s smallest terrestrial pteridophyte. *Sci. Rep.* 8, 5911. doi: 10.1038/s41598-018-24135-2
- Peterson, R. L., and Brisson, J. D. (1977). Root cap structure in the fern *Ophioglossum petiolatum*: light and electron microscopy. *Can. J. Bot.* 55, 1861–1878. doi: 10.1139/b77-215
- Petry, L. C. (1914). The anatomy of *Ophioglossum pendulum*. *Bot. Gaz.* 57, 169–192. doi: 10.1086/331260
- PPG I (2016). A community-derived classification for extant lycophytes and ferns. *J. Syst. Evol.* 54, 563–603. doi: 10.1111/jse.12229
- Ranwez, V., Harispe, S., Delsuc, F., and Douzery, E. J. P. (2011). MACSE: multiple alignment of coding sequences accounting for frameshifts and stop codons. *PLoS One* 6, e22594. doi: 10.1371/journal.pone.0022594
- Robison, T. A., Grusz, A. L., Wolf, P. G., Mower, J. P., Fauskee, B. D., Sosa, K., et al. (2018). Mobile elements shape plastome evolution in ferns. *Genome Biol. Evol.* 10, 2558–2571. doi: 10.1093/gbe/evy189
- Rothwell, G. W., and Stockey, R. A. (1989). Fossil ophioglossaceae in the paleocene of western north america. *Am. J. Bot.* 76, 637–644. doi: 10.2307/2444111
- Shen, X.-L., Tong, L., Lei, F.-W., Zhu, Y.-X., Mu, X.-Y., and Jiang, W.-J. (2020). The complete chloroplast genome of a rare fern species from North China, *Botrychium lunaria* (Ophioglossaceae). *Mitochondrial DNA Part B* 5, 1313–1314. doi: 10.1080/23802359.2020.1732243
- Shinohara, W., Nakato, N., Yatabe-Kakugawa, Y., Oka, T., Kun Kim, J., Murakami, N., et al. (2013). The use of *matK* in Ophioglossaceae phylogeny and the determination

of *Mankyua* chromosome number shed light on chromosome number evolution in Ophioglossaceae. *Syst. Bot.* 38, 564–570. doi: 10.1600/036364413X670232

Song, Y. Y., Cui, X. S., Xu, L., Xing, Y. P., Bian, C., Qiao, Y., et al. (2021). The complete mitochondrial genome of *Dryopteris crassirhizoma* Nakai (Dryopteridaceae, *Dryopteris* Adanson). *Mitochondrial DNA Part B Resour.* 6, 2704–2705. doi: 10.1080/23802359.2021.1966344

Suetsugu, K., Taketomi, S., Tanabe, A. S., Haraguchi, T. F., Tayasu, I., and Toju, H. (2020). Isotopic and molecular data support mixotrophy in *Ophioglossum* at the sporophytic stage. *New Phytol.* 228, 415–419. doi: 10.1111/nph.16534

Sun, B. Y., Baek, T. G., Kim, Y. D., and Kim, C. S. (2009). Phylogeny of the family Ophioglossaceae with special emphasis on genus *Mankyua*. *Korean J. Plant Taxon.* 39, 135–142. doi: 10.11110/kjpt.2009.39.3.135

Wagner, W. H. J. (1990). "Ophioglossaceae," in *The families and genera of vascular plants. Vol. I. Pteridophytes and gymnosperms*. Eds. K. U. Kramer and P. S. Green (Berlin Heidelberg: Springer), 193–197.

Whittier, D. P. (2015). Delayed growth in mycoheterotrophic gametophytes of seedless vascular plants. *Am. Fern J.* 105, 1–10. doi: 10.1640/0002-8444-105.1.1

Wickell, D. A., and Li, F. (2020). On the evolutionary significance of horizontal gene transfers in plants. *New Phytol.* 225, 113–117. doi: 10.1111/nph.16022

Winther, J., and Friedman, W. (2007). Arbuscular mycorrhizal symbionts in *Botrychium* (Ophioglossaceae). *Am. J. Bot.* 94, 1248–1255. doi: 10.3732/ajb.94.7.1248

Xie, P.-J., Ke, Y.-T., and Kuo, L.-Y. (2023). Modified CTAB protocols for high-molecular-weight DNA extractions from ferns. *Appl. Plant Sci.* 11, e11526. doi: 10.1002/aps3.11526

Yang, Z., and Rannala, B. (2012). Molecular phylogenetics: principles and practice. *Nat. Rev. Genet.* 13, 303–314. doi: 10.1038/nrg3186

Zhang, L., Fan, X., Petchsri, S., Zhou, L., Pollawatn, R., Zhang, X., et al. (2020). Evolutionary relationships of the ancient fern lineage the adder's tongues (Ophioglossaceae) with description of *Sahashia* gen. nov. *Cladistics* 36, 380–393. doi: 10.1111/cla.12408

Zhang, L., and Zhang, L.-B. (2022). Phylogeny, character evolution, and systematics of the fern family Ophioglossaceae based on Sanger sequence data, plastomes, and morphology. *Mol. Phylogenet. Evol.* 173, 107512. doi: 10.1016/j.ympev.2022.107512

Zhou, W., Soghigian, J., and Xiang, Q.-Y. (2022). A new pipeline for removing paralogs in target enrichment data. *Syst. Biol.* 71, 410–425. doi: 10.1093/sysbio/syab044

Zumkeller, S., Polsakiewicz, M., and Knoop, V. (2023). Rickettsial DNA and a trans-splicing rRNA group I intron in the unorthodox mitogenome of the fern *Haplopteris ensiformis*. *Commun. Biol.* 6, 296. doi: 10.1038/s42003-023-04659-8