Received: 12 February 2024 | Accepted: 20 May 2024

DOI: 10.1002/ajb2.16389

RESEARCH ARTICLE

Botany

All tangled up: Unraveling phylogenetics and reticulate evolution in the vining ferns, *Lygodium* (Schizaeales)

Jessie A. Pelosi[1] | **Bethany A. Zumwalde**[1] | **Weston L. Testo**[1,2,3] | **Emily H. Kim**[1,4] | **J. Gordon Burleigh**[1] | **Emily B. Sessa**[1,5]

[1]Department of Biology, University of Florida, Gainesville, Florida 32611, USA

[2]Botanical Research Institute of Texas, Fort Worth, Texas 76107, USA

[3]The Pringle Herbarium, Department of Plant Biology, University of Vermont, Burlington, Vermont 05405, USA

[4]Department of Microbiology and Cell Sciences, University of Florida, Gainesville, Florida 32611, USA

[5]William and Lynda Steere Herbarium, The New York Botanical Garden, Bronx, New York 10458, USA

Correspondence

Jessie A. Pelosi, Department of Biology, University of Florida, Gainesville, FL, USA 32611.
Email: jessiepelosi@ufl.edu

Abstract

Premise: Reticulate evolution, often accompanied by polyploidy, is prevalent in plants, and particularly in the ferns. Resolving the resulting non-bifurcating histories remains a major challenge for plant phylogenetics. Here, we present a phylogenomic investigation into the complex evolutionary history of the vining ferns, *Lygodium* (Lygodiaceae, Schizaeales).

Methods: Using a targeted enrichment approach with the *GoFlag 408* flagellate land plant probe set, we generated large nuclear and plastid sequence datasets for nearly all taxa in the genus and constructed the most comprehensive phylogeny of the family to date using concatenated maximum likelihood and coalescence approaches. We integrated this phylogeny with cytological and spore data to explore karyotype evolution and generate hypotheses about the origins of putative polyploids and hybrids.

Results: Our data and analyses support the origins of several putative allopolyploids (e.g., *L. cubense, L. heterodoxum*) and hybrids (e.g., *L.* ×*fayae*) and also highlight the potential prevalence of autopolyploidy in this clade (e.g., *L. articulatum, L. flexuosum,* and *L. longifolium*).

Conclusions: Our robust phylogenetic framework provides valuable insights into dynamic reticulate evolution in this clade and demonstrates the utility of target-capture data for resolving these complex relationships.

KEYWORDS

aneuploidy, fern, phylogenomics, polyploidy, reticulate evolution

Reticulate evolution, including hybridization, polyploidy, and horizontal gene transfer, is prevalent throughout vascular plants (e.g., Grant, 1971; Soltis and Soltis, 2009; One Thousand Plant Transcriptomes Initiative, 2019; Wickell and Li, 2020; Stull et al., 2023), creating complex evolutionary histories that are not represented in a typical bifurcating phylogenetic tree (Linder and Rieseberg, 2004). Generating robust phylogenetic frameworks that incorporate reticulate evolution remains a challenge in plant biology. While cytological observations have provided insight into genome evolution and the reticulate histories of plants for decades, recent large-scale genomic datasets offer new promise for revealing conflicting evolutionary histories among genomic regions (e.g., Edelman et al., 2019) and have motivated the development of new tools and methods to disentangle reticulation. These include software designed to handle allelic phasing (Nauheimer et al., 2021; Freyman et al., 2023; Mendez-Reneau et al., 2023; Tiley et al., 2024), infer phylogenetic networks (Than et al., 2008; Solís-Lemus et al., 2017), and implement comparative methods incorporating reticulation (Bastide et al., 2018; Karimi et al., 2020). In this study, we demonstrate how integrating traditional insights from cytology and morphology with new approaches using high throughput sequencing (HTS) data can help elucidate the reticulate evolutionary history of the vining ferns in the genus *Lygodium* (Lygodiaceae, Schizaeales).

Reticulation is particularly prevalent in ferns (Manton, 1950; Barrington et al., 1989; Sigel, 2016), a group of around 10,000 species of varying morphologies and habits (PPG I, 2016).

Although the mechanisms promoting reproductive isolation are poorly understood in ferns (but see Haufler, 2008 for a commentary on speciation mechanisms), based on the sheer frequency of hybridization (around 1200 fern species are known to be of hybrid origin; Knobloch, 1976), prezygotic barriers appear to be relatively weak compared to other lineages of plants. Furthermore, hybridization between deeply diverged (>60 my) species can occur naturally (e.g., ×*Cystocarpium* Fraser-Jenk., Rothfels et al., 2015). Most homoploid fern hybrids are sterile (Knobloch, 1976; Barrington et al., 1989; but see Walker, 1958; Whittier and Wagner, 1971 for examples of fertile hybrids), but fertility can be restored by the coupling of hybridization with a change in ploidy level (i.e., allopolyploidy). This is common in ferns, as an estimated one-third of speciation events are accompanied by a change in ploidy (Wood et al., 2009), and all or nearly all ferns have at least one paleopolyploidy event somewhere in their evolutionary history (Huang et al., 2019; One Thousand Plant Transcriptomes Initiative, 2019; Pelosi et al., 2022). Unraveling the dynamics of reticulate evolution is therefore critical for understanding the biology of ferns, and, more broadly, plants.

Lygodium Sw. (Lygodiaceae), a small clade with current classifications recognizing between 22 and 49 species (Garrison Hanks, 1998; PPG I, 2016), is, with *Salpichlaena* J. Sm. (a small genus in the Blechnaceae), one of only two clades of vining ferns. This habit is facilitated by indeterminate rachis growth, circumnutation (i.e., twining; Figure 1F), and dormant pinnae buds (only present in *Lygodium*, Figure 1G), which allow for the growth of long climbing fronds that can reach up to 30 m and can create dense thickets or mats (Figure 1A). In *Lygodium*, leaflets are dimorphic (Figure 1C–E, I–L), with fertile segments bearing modified margins called sorophores with two rows of sporangia, and leaf morphology is highly variable between species (Figure 1I–L). The genus has a pantropical distribution with centers of diversity in the Asian and American tropics in addition to a few temperate taxa (Garrison Hanks, 1998), and *Lygodium* species have substantial ecological and economic impacts and uses across their range. In its native range, *Lygodium* is used for weaving (Morton, 1966; Rahayu et al., 2020), braided as rope, and made into ceremonial headpieces (Ranker et al., 2022). In the mid-1800s and early 1900s, several *Lygodium* species became widely cultivated as ornamental plants, and escapes of these plants, including *L. japonicum* (Thunb.) Sw. and *L. microphyllum* (Cav.) R.Br., led to established invasions in North America that would become enormously costly (Pemberton and Ferriter, 1998). Between 1998 and 2008, management and research on *L. microphyllum* cost the state of Florida over $15 million (Koop, 2009), and an estimated additional $2 million is spent annually in the state to manage *Lygodium* (Hiatt et al., 2019).

Early work (e.g., Manton and Sledge, 1954; Brownlie, 1961; Mitui, 1965; Walker, 1966) identified several polyploid taxa in *Lygodium* (Table 1). Subsequently, Roy and Manton (1965) determined that *Lygodium* has an aneuploid series of base chromosome numbers with x = 28, 29, 30. Although there has been substantial cytological work on the genus, little is known about the frequency of polyploid speciation or aneuploidy, the relative roles of auto- and allopolyploidy in this clade, or the origins of polyploids in *Lygodium*. The only molecular phylogeny focused on *Lygodium* comes from Madeira et al. (2008), who used two plastid loci to reconstruct the evolutionary relationships among *Lygodium* for the purpose of identifying suitable biocontrol agents for the invasive species in the United States. Plastid markers, however, likely act as a single linkage group (Lynch, 2007) which is maternally inherited in ferns (e.g., Gastony and Yatskievych, 1992; Vogel et al., 1998; Kuo et al., 2018), and therefore, they capture only part of the evolutionary history of an individual. Given the complexities of reticulate evolution and its apparent prevalence in *Lygodium*, multiple independent nuclear loci likely can provide more insight into the evolutionary relationships and the parental origins of polyploids. Here, we use a new target-capture dataset combined with spore morphology, genome size estimates, and chromosome counts to reconstruct the evolutionary relationships of *Lygodium* and identify polyploid taxa and their putative diploid progenitors. A more thorough understanding of the evolution of *Lygodium* that incorporates multiple data types allows us to evaluate the existing taxonomy and classification of the genus and discuss in detail the role that reticulate evolution has played in the group's history.

MATERIALS AND METHODS

Sampling and DNA extraction

We obtained herbarium specimens from the New York Botanical Garden (NY), Field Museum of Natural History (F), Herbarium Pacificum (BISH), Smithsonian's U.S. National Herbarium (US), Gray Herbarium at Harvard University (GH), Pringle Herbarium at the University of Vermont (VT), and the University of Florida Herbarium (FLAS) with the goal of sampling at least two specimens from each taxon. Species identifications were based on the treatment of *Lygodium* by Garrison Hanks (1998) and previous identifications on herbarium sheets (many of which were made by Garrison Hanks herself). Our sampling included a total of 113 samples representing 29 *Lygodium* taxa and four outgroups. We included at least two specimens from each taxon in our dataset, except for seven taxa that were represented by a single accession (Appendix S1: Table S1). We sampled multiple accessions of *L. reticulatum* Schkuhr and *L. versteegii* Christ, but the quantity and quality of DNA from these samples was low and resulted in failed library preparation or low locus recovery. Of the taxa recognized by Garrison Hanks (1998), we were unable to sample only *L. kingii* Copel., an endemic species in New Guinea that may be synonymous with *L. salicifolium* Presl (Holttum 1959). We included several taxa that were treated as synonyms of other names by Garrison Hanks (1998): *L. conforme* C.Chr. (synonymized with *L. circinnatum* (Burm.f.) Sw.), *L. digitatum* D.C.Eaton (*L. longifolium* (Willd.) Sw.), *L. dimorphum* Copel. (*L.trifurcatum* Baker), *L. mexicanum* Presl and *L.*

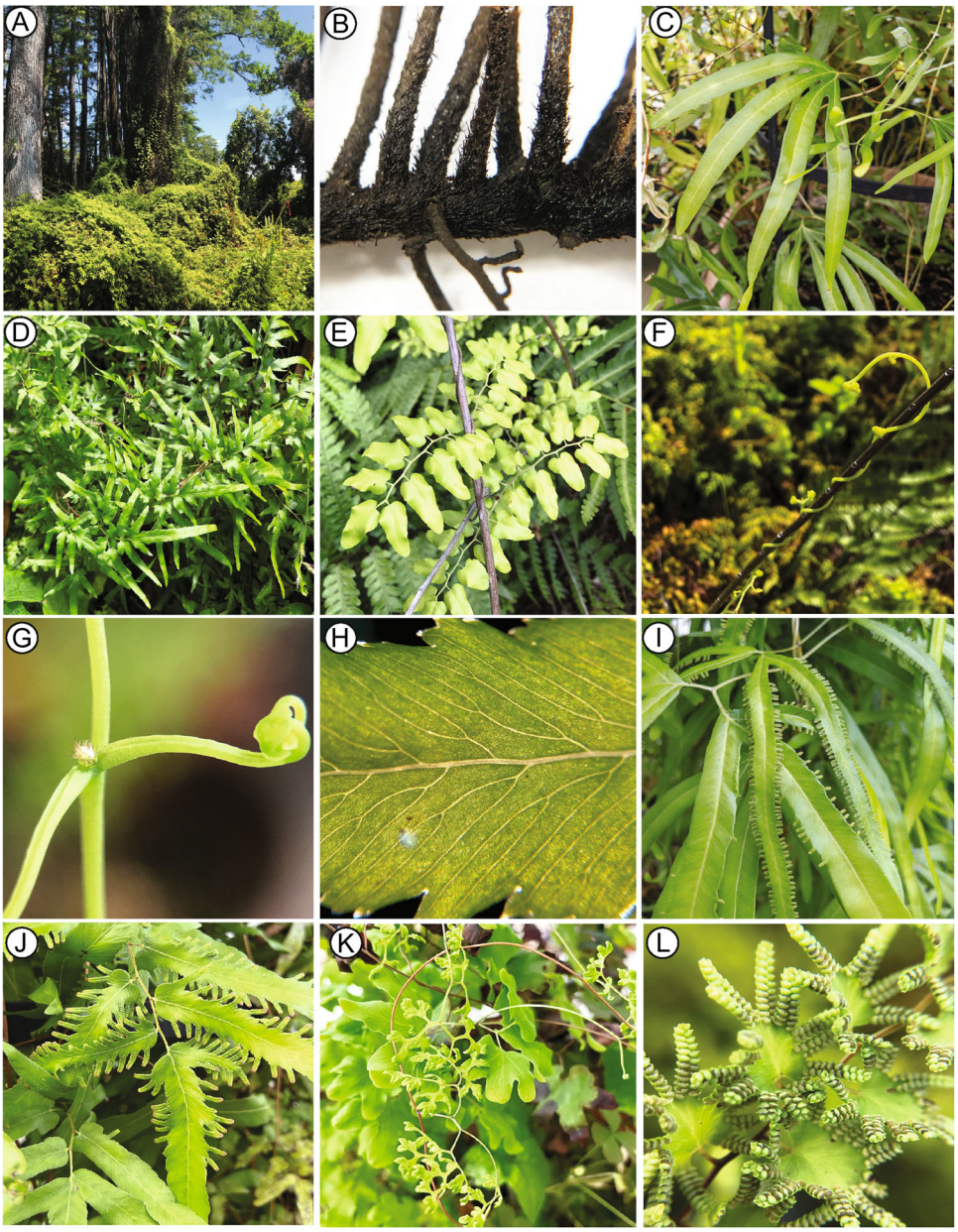

FIGURE 1 Character diversity of *Lygodium*. (A) General vining habit and mat-like growth of *Lygodium microphyllum*. (B) Short creeping rhizome of *L. japonicum* covered in shiny black hairs at 20× magnification. (C) Sterile segments of *L. circinnatum*. (D) Sterile segments of *L. japonicum*. (E) Sterile segments of *L. microphyllum*. (F) Circumnutation (i.e., twining) of a young *L. microphyllum* leaf. (G) Resting bud of *L. microphyllum*. (H) Free veins of *L. volubile* at 10× magnification, the most common venation pattern in *Lygodium*. (I) Fertile segments of *L. circinnatum*. (J) Fertile segments of *L. flexuosum*. (K) Fertile segments of *L. palmatum*. (L) 10× magnification of sorophores of *L. microphyllum*. Images A, B, E–H, L by J. A. Pelosi; C, D, I–K by E. B. Sessa.

polymorphum (Cav.) Kunth (*L. venustum* Sw.), *L. scandens* (L.) Sw. (*L. microphyllum*), and *L. semihastatum* (Cav.) Desv. (*L. auriculatum* (Willd.) Alston), and the putative hybrid *Lygodium* ×*fayae* Jermy & Walker (*L. venustum* × *L. volubile* Sw.).

Small fragments of tissue were removed from herbarium sheets, and DNA was extracted using a modified CTAB protocol (Doyle and Doyle, 1987) with 3X CTAB buffer (0.1 M Tris-HCl pH 8.0, 1.4 M NaCl, 20 mM EDTA, 30 mM β-mercaptoethanol, 3% CTAB), followed by a chloroform extraction and two 75% ethanol washes. Pellets were resuspended in 1X TE, and DNA concentration was quantified using a Qubit 2.0 fluorometer (Invitrogen, Waltham, Massachusetts, USA) with the broad-range dsDNA assay.

Ploidy and proxy measurements

We employed several methods to infer the ploidy of *Lygodium* taxa, which is important for estimating the frequency of

TABLE 1 Chromosome counts, genome size estimates, ploidal levels, and spore sizes of *Lygodium* taxa. Chromosome counts and genome size estimates generated in this study are underlined. Bolded chromosome counts and ploidies represent the values most commonly reported in the species.

Taxon	Haploid Chromosome Number (*n*)	Ploidy	1 C Genome Size (pg)	Avg Spore Length (μm)	Avg Spore Width (μm)	Avg Spore Area (μm^2)	Chromosome Count Reference	Genome Size Reference
Lygodium articulatum	**ca. 70**	**4x**	-	98.00	103.08	9005.54	Brownlie (1961)	-
Lygodium auriculatum[1]	**28**	**2x**	-	73.65	78.95	113.58	Nakato and Kato (2001)	-
Lygodium borneense	-	-	-	51.59	53.10	2052.17	-	-
Lygodium circinnatum	**29, 58**	**2x, 4x**	9.34-9.72	81.59	86.59	5595.04	Manton and Sledge (1954), Roy and Manton (1965)	Kuo and Li (2019), Fujiwara et al. (2023)
Lygodium conforme	-	-	-	70.05	73.93	4165.91	-	-
Lygodium cubense	-	-	18.39	95.42	100.05	8114.78	-	This study
Lygodium flexuosum	**56**	**4x**	18.04	66.91	71.59	3783.14	Roy and Manton (1965), Tindale and Roy (2002)	Fujiwara et al. (2023)
Lygodium heterodoxum	-	-	-	70.72	76.49	4317.49	-	-
Lygodium hians	-	-	-	101.23	100.82	8618.17	-	-
Lygodium japonicum	29, **58**	2x, **4x**	11.66-14.38	64.61	68.38	4828.15	Nakato (1990), Manton and Sledge (1954), Mitui (1965), Roy and Holttum (1965), Roy and Manton (1965), Mitui (1968), Pelosi et al. (2023)	Hanson and Leitch (2002), Pelosi et al. (2023)
Lygodium kerstenii	-	-	-	78.88	86.19	6188.20	-	-
Lygodium lanceolatum	-	-	-	86.54	90.93	6863.66	-	-
Lygodium longifolium[2]	**56**	**4x**	10.55	94.68	91.60	6798.79	Roy and Manton (1965)	This study
Lygodium merrillii	-	-	-	87.00	-	-	-	-
Lygodium microphyllum[3]	**30**, 60	**2x, 4x**	5.56-6.20	57.90	60.65	3499.15	Manton and Sledge (1954), Roy and Manton (1965), Lin et al. (2002), Tindale and Roy (2002), Pelosi et al. (2023)	Kuo and Li (2019), Pelosi et al. (2023)
Lygodium oligostachyum	**58**	**4x**	15.73	87.15	93.38	7080.94	This study	This study
Lygodium palmatum	**30**	**2x**	9.96	56.95	49.45	2212.97	Wagner and Wagner (1966), Wagner et al. (1970)	This study
Lygodium polystachyum	-	-	8.24 - 8.99	56.97	58.45	4145.45	-	Fujiwara et al. (2023), This study
Lygodium radiatum	-	-	-	100.02	98.73	8052.63	-	-

TABLE 1 (Continued)

Taxon	Haploid Chromosome Number (*n*)	Ploidy	1 C Genome Size (pg)	Avg Spore Length (μm)	Avg Spore Width (μm)	Avg Spore Area (μm^2)	Chromosome Count Reference	Genome Size Reference
Lygodium reticulatum	**30**	**2x**	-	79.53	85.97	6191.02	Takamiya (1995), Tindale and Roy (2002)	-
Lygodium salicifolium	**28**	**2x**	17.92	49.05	52.00	2305.85	Roy and Manton (1965)	Fujiwara et al. (2023)
Lygodium smithianum	-	-	-	64.64	62.29	3358.47	-	-
Lygodium subareolatum Christ	-	-	-	-	-	-	-	-
Lygodium trifurcatum[4]	28	2x	-	75.47	78.02	5157.69	Roy and Manton (1965)	-
Lygodium venustum[5]	**29**	**2x**	-	63.50	66.23	3638.55	Walker (1985)	-
Lygodium versteegii	-	-	-	90.04	94.43	6514.09	-	-
Lygodium volubile	**29**, 87	**2x**, 3x, 6x	11.33–14.81	80.46	78.40	5008.74	Roy and Manton (1965), Walker (1966, 1985), This study	Clark et al. (2016), This study
Lygodium yunnanense	-	-	-	78.43	82.71	4811.48	-	-
Lygodium × *fayae*	-	-	-	67.11	59.91	3211.05	-	-

Notes: [1]Includes *L. semihastatum;* [2]Includes *L. digitatum;* [3]Includes *L. scandens;* [4]Includes *L. dimorphum;* [5]Includes *L. mexicanum* and *L. polymorphum*, also note that Walker (1985) made a count for *Lygodium venustum* that appears to be two or three cells, rather than a single cell, resulting in a count double that of one cell.

polyploidy and accurately inferring their evolutionary histories. First, spores from a subset of species were collected from herbarium sheets or provided by colleagues, sown on Bold's media (Bold, 1957) with Nitsch's micronutrients (Nitsch, 1951), and germinated at 25°C on a 12:12 light:dark cycle in a Percival growth chamber. To count chromosomes, actively dividing root tips from sporophytes derived from sporophytic selfing (sensu Haufler et al., 2016) were used for chromosome squashes and counts, following the protocol in Pelosi et al. (2023). Additional chromosome counts were compiled from a literature search starting with the Chromosome Counts Database (CCDB, available at http://ccdb.tau.ac.il [accessed May 2023]; Rice et al., 2015).

Second, tissues from haploid gametophytes or diploid sporophytes derived from sporophytic selfing were used to estimate genome size using flow cytometry following Pelosi et al. (2023; see Appendix 1 therein) at the University of Florida Interdisciplinary Center for Biotechnology Research Cytometry Core (RRID:SCR_019119). We used both internal and external genome size standards including *Pisum sativum* L.'Ctirad' (2 C = 9.09 pg; Doležel et al., 1998), *Secale cereale* L. 'Daňkovské' (2 C = 16.19 pg; Doležel et al., 1998), and *Vicia faba* L. 'Inovec' (2 C = 26.90 pg; Doležel et al., 1992).

Spore size is commonly associated with genome size and ploidy in ferns (Wagner, 1974; Schuettpelz et al., 2015; Barrington et al., 1989, 2020). To explore this relationship in *Lygodium*, we collected spores for light microscopy from fertile specimens from herbarium sheets, which were mounted on slides and imaged at 400× with an AmScope T340B-LED microscope and an AmScope MU1000-HS camera. We tried to photograph five to ten spores per sample where possible. Composite images were generated from multiple focal depths. Spore length and width were measured following Barrington et al. (2020), and additional area measurements were made with ImageJ (Schneider et al., 2012) as in Pelosi et al. (2023).

Target capture sequencing and nuclear phylogenomics

DNA from 113 samples (109 *Lygodium* and 4 outgroup taxa; Appendix S1: Table S1) was sent to RAPiD Genomics (Gainesville, Florida, USA) for target capture sequencing using the *GoFlag 408* flagellate land plant probe set (Breinholt et al., 2021a), which targets 408 conserved exons found in 229 low-copy nuclear genes. Library construction, targeted enrichment, and sequencing were completed by RAPiD Genomics following methods described in Breinholt et al. (2021a). The resulting Illumina sequences were demultiplexed and trimmed of adapters and low-quality bases using Trim Galore! version 0.4.4 (https://www.bioinformatics.babraham.ac.uk/projects/trim_galore/). Bases with Phred scores <20 were removed and only read pairs >30 bp were retained (--quality 20 --length 30 --paired). Cleaned, trimmed reads were passed through the GoFlag pipeline (Breinholt et al., 2021b) for assembly of the target-only (conserved exons) and the target + flanking (mostly non-coding) regions. The pipeline includes an iterative baited assembly (IBA) step, which uses USEARCH version 7.0 (Edgar, 2010) to identify significant homology matches to reference sequences, followed by de novo assembly of each locus using the matching reads with BRIDGER version 2014-12-01 (Chang et al., 2015) with 3 iterations, using a k-mer size of 25, and 10× minimum coverage. tBLASTx (Camacho et al., 2009) searches against ten flagellate land plant genomes (see Breinholt et al., 2021a) were then used to assign orthology for sequences if the hit had no additional matches with >95% of the bitscore outside of a 1000 bp window of the target sequence in any of the ten genomes. Sequences with best hits in reference sequences belonging to clades other than ferns were removed as possible contaminants. Sequences for each locus were aligned with MAFFT version 7.490 (Katoh and Standley, 2013). Putative isoforms from the same sample were merged with nucleotide ambiguity codes representing variable sites, and samples with more than one sequence were removed from each locus in the alignment. We also removed alignment columns with fewer than four sequences using a custom Perl script (*deletecol_13Aug.pl*; see Data Availability Statement).

Many aspects of the target enrichment assembly pipeline can affect the resulting phylogenetic datasets. Therefore, we also ran HybPiper version 2.0.1 (Johnson et al., 2016) using the same nucleotide reference sequences as used in the GoFlag pipeline and evaluated where data from the two approaches produced consistent results. Trimmed reads were mapped to the references with BWA version 0.7.17 (Li and Durbin, 2009). Sequences were assembled for each locus with SPAdes version 3.15.3 (Bankevich et al., 2012), and coding sequences extracted with Exonerate version 2.4.0 (Slater and Birney, 2005). For the supercontig dataset, we used the --run_intronerate flag within HybPiper to capture intron sequences that flank the exons. Sequences within loci were aligned with MAFFT version 7.490 (Katoh and Standley, 2013) using the same settings and flags as used in the GoFlag pipeline, and sites with >80% missing samples were removed with trimAl version 1.4.1 (Capella-Gutiérrez et al., 2009). For both pipelines, we excluded fifteen samples due to either low locus recovery (<35 loci, 13 samples) or possible contamination (two samples), and sequence data from libraries derived from the same specimen were merged (four total libraries; two each of *L. hians* E.Fourn. and *L. cubense* Kunth).

The best nucleotide substitution model for each locus was determined with ModelFinderPlus (Kalyaanamoorthy et al., 2017), and maximum likelihood (ML) gene trees were inferred for each locus alignment using IQ-TREE version 2.1.3 (Minh et al., 2020) with 1000 ultrafast bootstraps (Hoang et al., 2018). We also concatenated all loci into a single phylogenetic matrix and partitioned the matrix by locus to generate a ML-based species tree with IQ-TREE. The best substitution model was determined for each locus in the matrix, allowing each locus to have its own evolutionary rate. ML gene trees were used as input to generate species trees under the multi-species coalescent (MSC) with

ASTRAL version 5.7.7 (Zhang et al., 2018b). Although these ML and MSC methods can be effective for inferring phylogenetic trees, they generate only fully resolved bifurcating topologies that do not incorporate the history of reticulation.

Plastid phylogenomics

While linked, uniparentally inherited plastid sequences cannot reveal the full history of reticulation alone, when compared with trees derived from nuclear data, they can help illuminate both areas of possible reticulation in the tree as well as the putative maternal lineage in a hybridization event. To extract off-target plastid reads from the target capture sequences, we first used GetOrganelle version 1.7.3.5 (Jin et al., 2020) with a custom database of fern plastomes downloaded from NCBI (198 plastomes from across a phylogenetic breadth of ferns accessed July 2020; available on our GitHub repository) to generate preliminary plastid gene assemblies for each sample. We fed these assemblies into PhyloHerb version 1.1.2 (Cai et al., 2022) to extract plastid loci. Loci with sequences from at least 20 samples were retained, aligned using MAFFT version 7.490 (Katoh and Standley, 2013) with the '--adjustdirectionaccurately' flag, and sites with >80% missing nucleotides were removed. We generated gene trees for each locus with IQ-TREE version 2.1.3 (Minh et al., 2020) as above and visualized them prior to concatenating. Upon reviewing the trees and alignments, we noticed that several loci had anomalous sequences that impacted tree topology. We manually curated the alignments by pruning sequences that were misaligned. After filtering, we generated a concatenated matrix, partitioned it by locus, and generated a ML tree as above with IQ-TREE version 2.1.3 (Minh et al., 2020).

Phasing and inferring reticulations

Most pipelines for assembling target-capture data, including the GoFlag pipeline and HybPiper, generate a consensus sequence for a sample from each locus, not individual sequence copies that could represent alleles, homeologs, or even paralogs. Yet these separate copies can help inform the history of reticulation and the parental origins of hybrids and allopolyploids. Phasing methods seek to elucidate each individual sequence from a locus from HTS data. Since phasing short-read HTS data from plants, and especially polyploids, is extremely challenging (e.g., Tiley et al., 2024), we used multiple approaches to test if the resulting sequences converged on a similar evolutionary story of reticulation in *Lygodium*.

First, we phased sequences with PATÉ (Tiley et al., 2024), specifying a maximum of two copies per sample per locus. Briefly, PATÉ maps sequencing reads back to a consensus reference for the loci assembled for each sample with BWA version 0.7.17 (Li and Durbin, 2009), marks and removes PCR duplicates with Picard version 2.9.2 (http://broadinstitute.github.io/picard), and calls SNPs with GATK version 4.1.4.0 (McKenna et al., 2010). We used the variant filtering parameters for GATK in Tiley et al. (2024). Retained biallelic SNPs were then phased with H-PoPG version 0.2.0 (Xie et al., 2016). Sequences were aligned and ML gene trees were constructed as above. Since PATÉ performs phasing for each locus independently and does not assign phased alleles to putative parental subgenomes, we manually inspected each gene tree and assigned alleles to maternal (A) and paternal (B) subgenomes based on their placement in the tree. Knowing to which subgenome a copy belongs enables us to concatenate copies into a single block and infer the phylogeny of the progenitor lineages. An alternative approach using phylogenetic networks does not require this knowledge (see Tiley et al., 2024), but to do this across *Lygodium* is extremely computationally onerous. The maternal subgenome was determined using the plastid phylogeny as a guide (see section Plastid Phylogenomics above). Based on data assembled and generated regarding the ploidy and hybrid origin of taxa, we only retained phased sequence copies for 11 samples that we determined to be hybrids and/or polyploids. For these samples, we replaced the consensus sequences with sequences that we were able to assign to the maternal or paternal subgenomes using the custom python script *distribute_alleles.py* (see Data Availability Statement). Phased sequences which could not be confidently assigned to a subgenome were removed from the alignment. If we determined that copies belonged to the same subgenome (i.e., were sister to each other), we randomly picked only one to include. We used the consensus sequences for the known diploids and putative autopolyploids (i.e., polyploids with phased copies that formed single clades) for alignment and gene and species tree reconstruction.

To evaluate our manual assignments of copies to subgenomes, we also selected from the PATÉ analyses the 20 loci that had the highest taxon occupancy and used these to infer subgenome assignment and placements with homologizer (Freyman et al., 2023) in RevBayes version 1.1.1 (Höhna et al., 2016) for both target-only and target + flanking datasets. We ran homologizer with 50,000 generations and verified that the posterior had an ESS > 200 with Tracer version1.7.2 (Rambaut et al., 2018). The first 10% of generations were discarded as burn-in. We determined that this number of loci was the upper limit for homologizer on our dataset; increasing the number of loci past 20 was not scalable as each run was computationally expensive and very slow even when running on multiple processors.

We phased both the gene and supercontig data from HybPiper, which are equivalent to the target-only and target + flanking datasets from the GoFlag pipeline using HybPhaser version 2.0 (Nauheimer et al., 2021). Unlike PATÉ, which phases reads from samples to their consensus sequences post-assembly, HybPhaser simultaneously maps reads for each sample to a set of consensus sequences representing clades with BBMap version 38.90 (https://sourceforge.net/projects/bbmap/) and then assembles the split reads with HybPiper. For the clade association with

HybPhaser, we picked one diploid, non-hybrid sample per taxon with the highest locus recovery as a reference. After assessing the results from the clade association, we generated phased sequences for all known and suspected polyploids and hybrids and used the two references with the highest mapping rates in the phasing step, unless there were substantial differences between the first and second mapping rates (e.g., 25% and 2% mapping rates). We merged the unphased and phased sequences with seqkit version 2.4.0 (Shen et al., 2016), and then generated alignments, gene trees, and species trees as above.

Sequence matrices for each locus were aligned with MAFFT version 7.490, and sites with >80% missing samples were removed with trimAl version 1.4.1 (Capella-Gutiérrez et al., 2009). ML trees were inferred using IQ-TREE version 2.1.3 (Minh et al., 2020) for each locus with 1000 ultrafast bootstraps (Hoang et al., 2018). The best model for each locus was determined with ModelFinderPlus (Kalyaanamoorthy et al., 2017). A concatenated matrix was constructed and partitioned by locus to generate a ML-based species tree with IQ-TREE. ML gene trees were used as input to generate species trees under the multi-species coalescent (MSC) with ASTRAL version 5.7.7 (Zhang et al., 2018b). A summary phylogeny, largely based on the GoFlag MSC target + flanking phylogeny, was built to visualize the relationships among *Lygodium* and putative reticulation events in the genus. Clades representing taxa were determined for each analysis and compared to the GoFlag MSC targets + flanking regions topology (Figure 2). Substantial topological differences and non-monophyletic taxa are noted on the summary phylogeny.

Comparative analyses of ploidy

We attempted to use the distribution of biallelic variants from the target enrichment sequence data to predict ploidy from the sequencing data. Trimmed reads were first mapped to the locus consensus sequences produced with the GoFlag pipeline for each sample with BWA version 0.7.17 (Li and Durbin, 2009), sorted with samtools version 1.8 (Li et al., 2009), and the distribution of biallelic positions was generated with nQuire version 20180405 (Weiß et al., 2018). Each distribution was denoised, to which Gaussian Mixture Models were fit for diploid, triploid, and tetraploid assumptions (lrdmodel) and compared to empirical data (histotest) in nQuire. Estimates of ploidal level were drawn based on the lowest model Δlog-likelihood from lrdmodel and a high R^2, positive slope, and low standard error from histotest as recommended by the authors.

To determine if ploidy, spore size metrics, and genome sizes were related, we did a sister-clade analysis for four polyploid-diploid taxa pairs (*L. articulatum* A.Rich.- *L. palmatum* (Bernh.) Sw.; *L. flexuosum* (L.) Sw. - *L. salicifolium; L. longifolium* - *L. auriculatum; L. oligostachyum* (Willd.) Desv. - *L. venustum*) for which we have chromosome count data to confirm ploidy, and compared the average spore length, width, and area between the polyploid and diploid taxa with paired Welch's two sample t-tests (e.g., Heilbuth, 2000). Using the chromosome count data, we additionally inferred changes in chromosome number across the phylogeny of *Lygodium* using the GoFlag target + flanking MSC topology and associated branch lengths with ChromEvol version 2.0 (Glick and Mayrose, 2014). Ten default models of chromosome number evolution were evaluated, and the best model was determined by comparing AIC from each. All analyses were conducted on the University of Florida's HiPerGator computing cluster.

RESULTS

Ploidy and proxy measurements

We obtained new chromosome counts for *L. oligostachyum* ($2n = 116$) and *L. volubile* ($2n = 87$), bringing the number of *Lygodium* taxa with chromosome counts to 14 (Table 1, Appendix S2). We also generated new genome size estimates for four taxa: *L. cubense* (1 C = 18.39 pg), *L. longifolium* (1 C = 10.55 pg), *L. oligostachyum* (1 C = 15.73 pg), and *L. palmatum* (1 C = 9.96 pg), bringing the total number of *Lygodium* taxa with genome size estimates to eleven. We report additional estimates for two taxa with previously reported estimates: *L. polystachyum* Wall. ex Moore (1 C = 8.99 pg in this study, 1 C = 8.49 pg from Fujiwara et al., 2023) and *L. volubile* (1 C = 11.33 pg in this study, 1 C = 14.81 pg from Clark et al., 2016).

We measured spores for all taxa except for *L. merrillii* Copel., for which we used a previous measurement from Garrison Hanks (1998). On average we measured eight spores per specimen (range: 3–12); a total of >400 spores were measured across the genus. Average spore length ranged from 49.05 μm in *L. salicifolium* to 101.23 μm in *L. hians*, and width ranged from 49.45 μm in *L. palmatum* to 103.08 μm in *L. articulatum* (Table 1, Figure 3). Spore measurements are provided in our GitHub repository (see Data Availability Statement).

Nuclear assemblies

An average of 1.27×10^6 reads (range: 0.005 to 3.05×10^6) were generated per sample. After excluding samples that had poor locus recovery (<35 loci), we recovered an average of 258 (range: 36–380) loci using the GoFlag pipeline and an average of 296 (range: 62–395) loci using HybPiper. Across all samples, the (unphased) GoFlag and HybPiper pipelines assembled 331 and 402 loci, respectively, for phylogenomics (Table 2). The difference in the number of loci assembled largely reflects the way we dealt with multiple copies per sample that may be retained in the GoFlag pipeline. In cases in which there appears to be greater than simple allelic variation in a sample for a locus, the GoFlag pipeline will retain multiple sequences from a sample. In those cases, to minimize possible paralogy issues, we chose to remove all copies from that

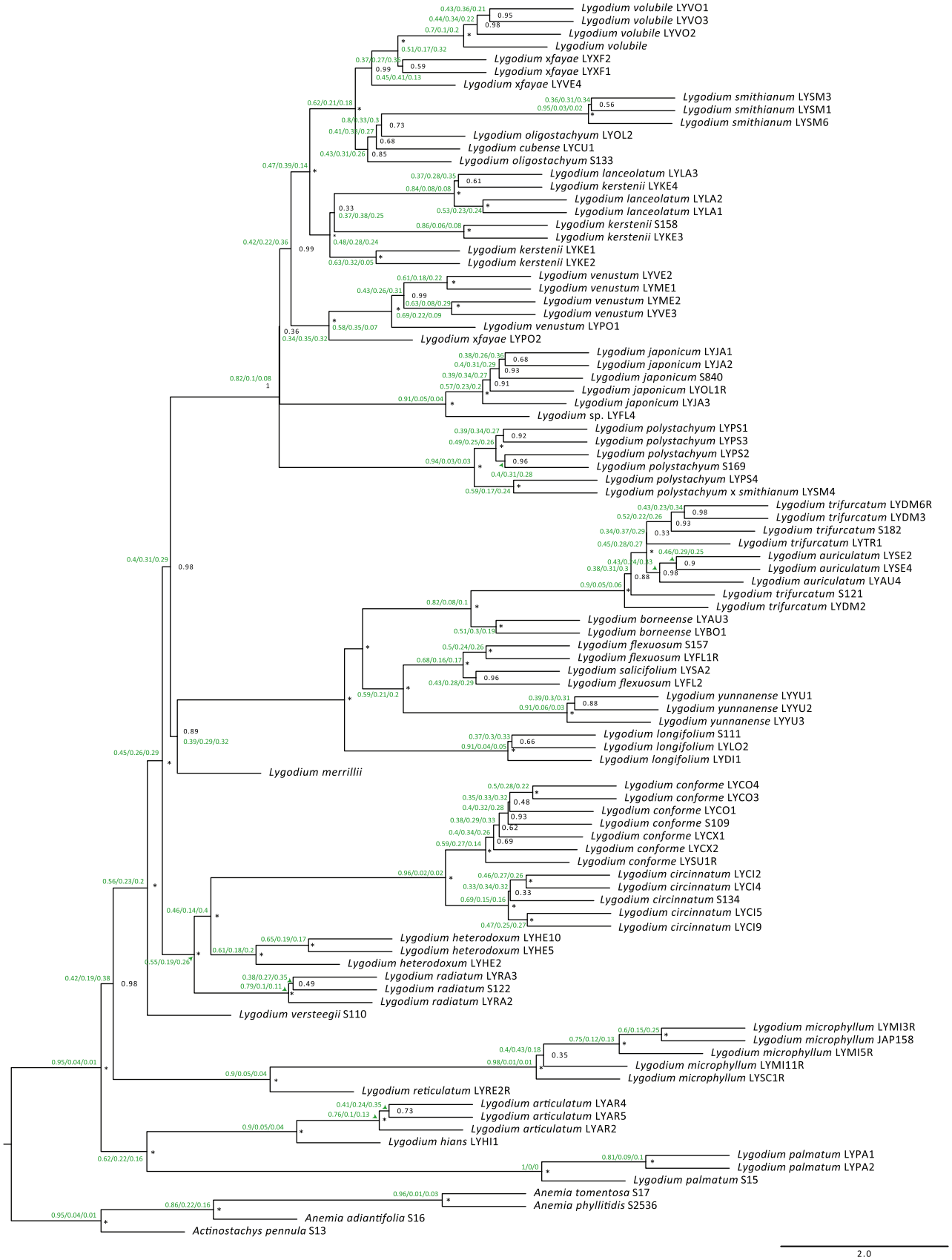

FIGURE 2 Multi-species coalescent phylogeny of *Lygodium* from the unphased target + flanking regions assembled with the GoFlag pipeline. Black node support values are local posterior probabilities; nodes with local posteriors of 1 are denoted with an asterisk. Green node values are the proportion of gene trees that support the primary quartet topology and the two alternative topologies. Branch lengths are proportional to coalescent time. Arrows indicate to which node values belong. Sample codes are based on the resolved identification of herbarium specimens; see Appendix S1, Table S1, column 'NCBI Sample ID'.

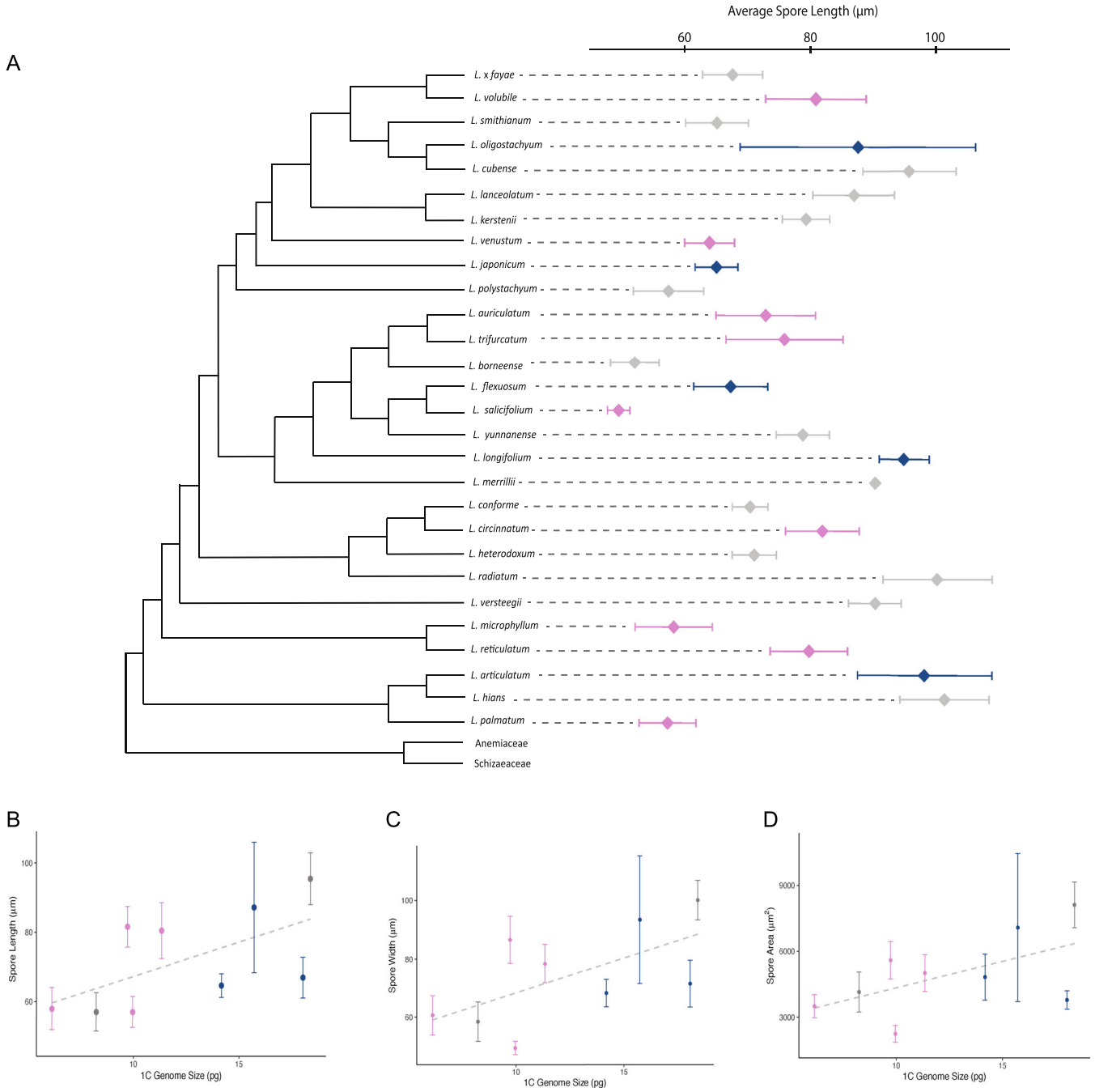

FIGURE 3 Phylogenetic distribution and association of spore and genome size metrics for *Lygodium*. (A) Summary phylogeny of *Lygodium* with spore length for each taxon. The mean spore length (diamond) and ± 1 standard deviation (error bars) are shown for each taxon. The known ploidal level for each taxon (based on chromosome counts) are denoted by color with dark blue = tetraploid, pink = diploid, gray = ploidy unknown. Average spore (B) length, (C) width, and (D) area as it relates to 1 C genome size for *Lygodium* taxa. Error bars are ± 1 standard deviation around the mean (circle). Known ploidal level colored as in panel A.

sample for the locus using a Perl script (*rmall.pl*; see Data Availability Statement). In contrast, when there are multiple sequences, HybPiper chooses among multiple full-length assembled contigs based on depth or percent identity and retains a single copy (Johnson et al., 2016). In the target-only alignment, the matrices consisted of 69,555 to 70,025 bp with an average locus length of 174 to 183 bp including 19,078 to 22,777 parsimony-informative sites (Table 2). The target + flanking matrices contained 216,537 to 262,870 bp with an average locus length of 667 to 882 bp including 77,734 to 102,303 parsimony-informative sites (Table 2). The difference in the number of parsimony-informative sites between the assembly methods is due in large part to the handling of multiple copies recovered for loci, such as the removal of duplicate sequences in the GoFlag pipeline.

TABLE 2 Summary statistics of the datasets generated in this study.

Dataset	Number of Loci	Average Locus Length (bp)	Total Length (bp)	Parsimony-Informative Sites
GoFlag Targets	382	183	69,555	19,078
GoFlag Targets + Flanking	382	882	216,537	102,303
HybPiper Genes	402	174	70,025	22,777
HybPiper SuperContigs	331	667	262,870	77,734
HybPhaser Genes	394	186	73,496	22,945
HybPhaser SuperContigs	394	667	262,900	97,717
PATÉ Targets	145	184	28,856	8069
PATÉ Targets + Flanking	158	666	105,288	39,240
Plastid	43	825	36,672	6657

Nuclear phylogenies

Support values across the inferred phylogenies tended to be lower in the target-only/gene analyses compared to the target + flanking/supercontig datasets (Appendix S3), which may be expected; target + flanking/supercontig datasets contain far more sites than the relatively conserved target regions alone (Faircloth et al., 2012). Although there are numerous differences in the topologies resulting from the target-only/gene alignments and the target + flanking/supercontig alignments, these differences in the multi-species coalescent (MSC) analyses are poorly supported and largely clustered along the backbone of the phylogeny, where there are several short internodes and lower support values. We therefore focus primarily on the MSC analysis of the target + flanking dataset assembled with the GoFlag pipeline for our discussion, which is generally congruent with the other inferred species trees (see Appendix S1: Table S2; Appendix S3 for phylogenies constructed for the other datasets). In general, recognized species were monophyletic and well-supported, with the exceptions of *L. auriculatum/L. trifurcatum* and *L. kerstenii* Kuhn/*L. lanceolatum* Desv. (Figure 2; Appendix S3).

Plastid assembly and phylogeny

We recovered plastid gene assemblies for 90 of the 97 samples used for nuclear phylogenomics. On average, these assemblies were 17,521 bp (range: 547–107,433 bp) and highly fragmented. We assembled sequences for 88 plastid loci using PhyloHerb. After filtering, the concatenated matrix of plastid genes included 43 plastid loci from 86 samples and was 36,672 bp in length (Table 2), with 6.5% of the sites being parsimony informative. The ML tree from the concatenated plastid matrix was largely congruent with the nuclear phylogeny topologies, although there are some key differences, including the placement of *L. articulatum* and *L. hians* sister to the rest of *Lygodium*, and several polyphyletic taxa (Appendix S3).

Phasing

An average of 258 (range: 36–380) loci were recovered for each sample in the target-only phased dataset from PATÉ, with an average of 82 (range: 7–283) loci phased. On average, 75.1% of loci were phased in known hybrids, compared to 53.1% in known polyploids, and 19.6% in diploids. Nearly all loci in all samples were phased as a single block (i.e., the entire locus was phased as a single unit). A similar number of loci were recovered in the target + flanking dataset from PATÉ (average 261, range: 38–382 loci), but there were more phased loci in this dataset with an average of 113 phased loci (range: 7–289). Known hybrids had an average of 84.7% of loci phased, compared to 70.4% in known polyploids, and 27.0% in diploids (Appendix S1: Tables S1, S3). After manually assigning alleles to subgenomes, there were 145 loci in the target-only PATÉ matrix, with an average locus length of 184 bp, consisting of 28,856 bp and 8,069 parsimony-informative sites (Table 2). The target + flanking PATÉ matrix was composed of 158 loci with an average locus length of 665 bp and was 105,288 bp in length with 39,240 parsimony-informative sites (Table 2).

For the samples that could be phased in HybPhaser, we recovered an average of 232 (range: 50–306) loci for the HybPiper genes dataset and 249 (range: 70–335) loci for the HybPiper supercontigs dataset. There were a total of 394 loci in the resulting genes matrix with an average locus length of 186 bp, consisting of 73,496 bp and 22,945 parsimony-informative sites (Table 2). The resulting supercontigs matrix was 262,900 bp in length from 394 loci with an average locus length of 666 bp and a total of 77,734 parsimony-informative sites (Table 2).

Reticulation events

For the phasing analysis using PATÉ, we were able to manually assign alleles with confidence to putative subgenomes for 145 and 158 loci in the target-only and target + flanking datasets, respectively (Table 2), for 11 specimens representing *L. cubense*, *L. heterodoxum* Kunze, *L. oligostachyum*, *L. polystachyum* × *smithianum* Presl, and *L.* ×*fayae* (Appendix S1: Table S3).

The proportion of loci phased, allele divergence (AD), and locus heterozygosity (LH) were high for known polyploids (e.g., *L. articulatum*, *L. longifolium*) and hybrids (e.g., *L.* ×*fayae*), and tended to be greater for the putative allopolyploids (*L. cubense*, *L. heterodoxum*, and *L. oligostachyum*) and hybrids (*L.* ×*fayae* and *L. polystachyum* × *smithianum*) compared to the putative autopolyploids (*L. articulatum*, *L. hians*, *L. radiatum* Prantl, *L. flexuosum* (L.) Sw., *L. longifolium*, *L. japonicum*; Appdendix S1: Tables S1, S3). The progenitors of *L.* ×*fayae* and *L. heterodoxum* were strongly supported in both the target-only and target + flanking datasets (Figure 4). There was mixed support for the progenitors of the putative *L. polystachyum* × *smithianum* hybrid, *L. cubense*, and *L. oligostachyum* (Figure 4). Interestingly, both target-only and target + flanking MSC analyses from PATÉ strongly supported the placement of one *L. polystachyum* × *smithianum* copy with *L. smithianum* and one with *L. polystachyum*, but the concatenated analyses placed both copies with *L. polystachyum* (Figure 4; Appendix S1: Table S2). The homologizer results strongly supported our findings on the assignment of putative parental subgenomes, with the exception of one *L. oligostachyum* sample which had low occupancy in the selected loci and *L. cubense* (Figure 4G; Appendix S1) in the target-only dataset.

In both the gene and supercontig HybPhaser analyses, the putative allopolyploid and hybrid samples resolved in the same clades as in the PATÉ analysis: *L. heterodoxum* copies sister to *L. circinnatum/L. conforme* and *L. merrillii*, *L.* ×*fayae* copies within *L. venustum* and *L. volubile*, *L. cubense* and *L. oligostachyum* copies sister to *L. smithianum* and *L. venustum* (with the exception of one *L. oligostachyum* sample with low locus recovery; Figure 4; Appendix S4). The copies of *L. japonicum*, *L. longifolium*, *L. flexuosum*, and *L. radiatum* all fell into their own clades, suggesting independent putative autopolyploid origins for these taxa. *Lygodium articulatum*, *L. hians*, *L. flexuosum*, and the putative *L. polystachyum* × *smithianum* hybrid had substantial differences in the mapping rates during the clade association, and we were thus unable to phase their copies with HybPhaser. The putative reticulation events in *Lygodium* are summarized in Figure 5.

Topologies

There was strong support for the monophyly of *Lygodium* throughout our datasets and analyses (Appendix S3). However, the subgenera and sections recognized by Garrison Hanks (1998), which were proposed based on morphological and chemical characters, are largely unsupported (Appendix S5). Garrison Hanks (1998) recognized two subgenera: *Lygodium*, which included *L. flexuosum*, *L. japonicum*, *L. kerstenii*, and *L. venustum*, and *Gisopteris* (Bernh.) C.Chr., in which she placed the remainder of the species. Within subgenus *Gisopteris*, Garrison Hanks (1998) recognized three sections: Palmata, Volubilia, and Polystachya (which contains only *L. polystachyum*). Our phylogenies recover a rather more complex set of relationships within *Lygodium*. In nearly all analyses, we recovered a clade composed of *L. articulatum*, *L. hians*, and *L. palmatum* (but see Figure 4A for exceptions) as sister to the rest of the genus, followed by a successive sister clade consisting of *L. microphyllum* and *L. reticulatum*; these two species together are sister to a clade that includes *L. versteegii* as a lone taxon sister to three major clades representing the remainder of *Lygodium* (Figures 2, 5). This topology is consistent with that recovered in Madeira et al. (2008), although their study included fewer taxa.

We recovered a clade of four species that corresponds to a group within section Volubilia that was proposed by Garrison Hanks (1998). This clade includes *L. radiatum* and *L. heterodoxum* as successively sister to a clade with *L. circinnatum* and *L. conforme* (Figures 2, 5; Appendix S5). All of these taxa were represented by multiple accessions and were monophyletic with moderate to strong support (Figures 2, 4). Other taxa placed in section Volubilia by Garrison Hanks (1998) fell in another well-supported clade consisting of *L. merrillii, L. longifolium*, *L. flexuosum*, *L. salicifolium*, *L. borneense* Alderw., *L. trifurcatum*, and *L. auriculatum*. The placement of *L. merrillii* sister to the rest of the "*L. longifolium* clade" was supported across most datasets, although there were some that supported an alternative topology with *L. merrillii* sister to the "*L. longifolium* clade" and "*L. volubile* clade" or successively sister to the remainder of *Lygodium* with *L. versteegii* (Figure 4B). We consistently recovered *L. digitatum* as sister to or nested within *L. longifolium* in a well-supported clade (Appendix S1: Table S2) and the sole sample of *Lygodium salicifolium* that we sampled was nested within the *L. flexuosum* clade, although this may be related to the putative polyploid origin of *L. flexuosum* (see Appendix S6). Both *L. trifurcatum* and *L. dimorphum* were non-monophyletic and resolved in a clade with *L. auriculatum* and *L. semihastatum* (Figure 2). The single sample of *L. auriculatum* in our phylogeny was sister to two accessions identified as *L. semihastatum* (Figure 2; Appendix S1: Table S2).

The final clade contained a mixture of taxa from both subgenera and had a recalcitrant backbone with short internodes and low support (Figure 2). Some analyses recovered a *L. japonicum* and *L. venustum* clade sister to the remainder of the "*L. volubile* clade" including *L. polystachyum*, whereas other analyses recovered *L. polystachyum* sister to the rest of the clade (Figures 4C, 5; Appendix S3). Interestingly, there was not a consistent pattern in the

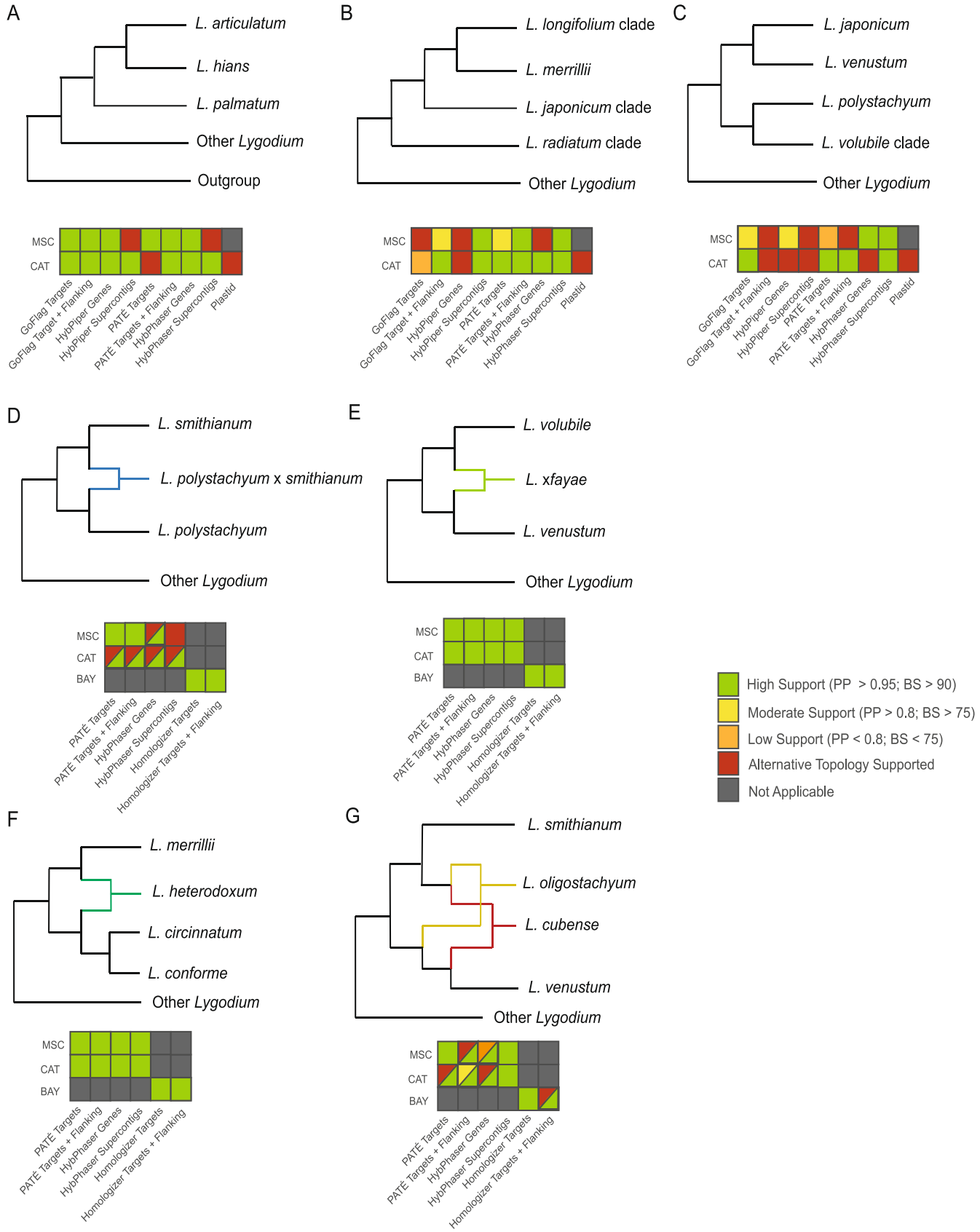

FIGURE 4 (See caption on next page).

placement of these taxa based on dataset or analysis type. Samples of both *L. mexicanum* and *L. polymorphum* were nested with *L. venustum* samples, forming a single, well-supported clade (Appendix 1: Table S2). One substantial difference in our topology is the placement of *L. venustum* nested in a clade of *L. japonicum* and *L. polystachyum* in Madeira et al. (2008), highlighting the difficulties in resolving relationships among taxa involved with reticulation with plastid data alone. *Lygodium kerstenii* and *L. lanceolatum* formed a clade, but the species were not reciprocally monophyletic (Figure 2). Similar to Madeira et al. (2008), we found *L. cubense* and *L. oligostachyum* sister to *L. smithianum*, although we recovered a grade of *L. cubense* and *L. oligostachyum* in some of our nuclear analyses and a clade in others (Figure 4G), while the latter placement was also found by Madeira et al. (2008) in their plastid dataset.

Karyotype evolution

The best ChromEvol model as determined by AIC was the constant rate model ($\Delta AIC = 2.0$). By reconstructing the ancestral haploid chromosome number, we inferred that the ancestral chromosome number for the family was $x = 30$, followed by a transition to $x = 29$ along the backbone, and a second transition to $x = 28$ in the clade composed of *L. auriculatum, L. trifurcatum*, *L. borneense*, *L. flexuosum*, *L. salicifolium*, *L. yunnanense* Ching, and *L. longifolium* (Figure 5). In the sister clade comparisons, we found that the spore area ($t = -2.77$, $df = 3$, $P = 0.035$), length ($t = -4.82$, $df = 3$, $P = 8.51 \times 10^{-3}$), and width ($t = -3.25$, $df = 3$, $P = 0.024$) of the polyploids was significantly greater than the diploids.

On average, 330 (range: 62–1672) biallelic positions were analyzed per sample in nQuire after filtering with the denoise command. We compared the estimated ploidy from nQuire to the known ploidy for a taxon based on chromosome counts, for species with only one known cytotype: in most cases (71.4%) nQuire did not accurately predict the sample's ploidal level.

DISCUSSION

Our analyses of target-capture data, together with new and existing cytological observations, provide strong support for relationships among many recognized species, and valuable insights into the dynamic reticulate evolutionary history of *Lygodium*. Below we synthesize perspectives from several different data types, in the most comprehensive systematic description of *Lygodium* to date.

Taxonomic considerations

Although this study generally follows the classification scheme from Garrison Hanks (1998), which recognizes 26 species of *Lygodium*, there has been little consensus on the number and names of species in *Lygodium*. Taxonomic treatments range from 22 (Prantl, 1881) to 49 (Duek, 1976) extant species, and the current classification from the Pteridophyte Phylogeny Group I (PPG I, 2016) recognizes about 40 unlisted taxa following Kramer (1990). Variation in these treatments is largely due to morphological plasticity and, presumably, reticulating evolution leading to potentially morphologically intermediate, allopolyploid species (Holttum, 1959; Garrison Hanks, 1998). Our new genomic data enable us to assess these existing treatments. We find strong support for the monophyly of *Lygodium*, which is consistent with previous morphological (Prantl, 1881; Reed, 1946; Alston and Holttum, 1959) and molecular analyses (Wikström et al., 2002; Schuettpelz and Pryer, 2007; Madeira et al., 2008; Testo and Sundue, 2016; Pelosi et al., 2022). Most taxa recognized by Garrison Hanks (1998) were recovered as monophyletic (Figures 2, 5).

By incorporating several representatives of synonymized taxa (following Garrison Hanks, 1998) in our sampling, we were able to test whether this classification scheme was appropriate based on the monophyly of the samples. We summarize the findings here, although there are several instances where the support is contradictory or equivocal based on the dataset/analysis (see Appendix S1: Table S2). In all of our datasets, we found strong support for Garrison Hanks's (1998) synonymizing of *L. mexicanum* and *L. polymorphum* with *L. venustum*, as samples of the former two were consistently intermixed and fell within a clade with *L. venustum* (Figure 2; Appendix S1: Table S2). A similar result was found for *L. scandens* and *L. microphyllum*, although in about half of the analyses, *L. scandens* was recovered as sister to *L. microphyllum* and has mixed support of synonymization by Garrison Hanks (1998). Zhang and Garrison Hanks (2013) suggest that the name *L. digitatum* be considered a synonym of *L. longifolium*, for which we found mixed support with some analyses placing *L. digitatum* nested within a clade with *L. longifolium* or as sister (Appendix S1: Table S2).

Some collectors and authors (e.g., Copeland, 1911; Hassler, 2024) recognize *L. dimorphum* as a distinct species from *L. trifurcatum* based on the degree of fertile frond

FIGURE 4 Selected subsets of relationships with varying levels of support from different analyses. Panels A–C show varying levels of conflict among major clades in *Lygodium*. Panels D and E show support for putative hybrids, and panels F and G show support for putative allopolyploids. MSC- species tree generated using the multi-species coalescent; CAT- species tree generated using maximum-likelihood on a concatenated matrix of sequences; BAY- species tree generated using Bayesian inference in homologizer. Green - high support for depicted topology (posterior probability [PP] >0.95, bootstrap [BS] >90); yellow - moderate support (PP > 0.85, BS > 75); orange - low support (PP < 0.85, BS < 75); red - alternative topology supported. For hybrids/allopolyploids (D–G), boxes may be split with part of the topology differently supported.

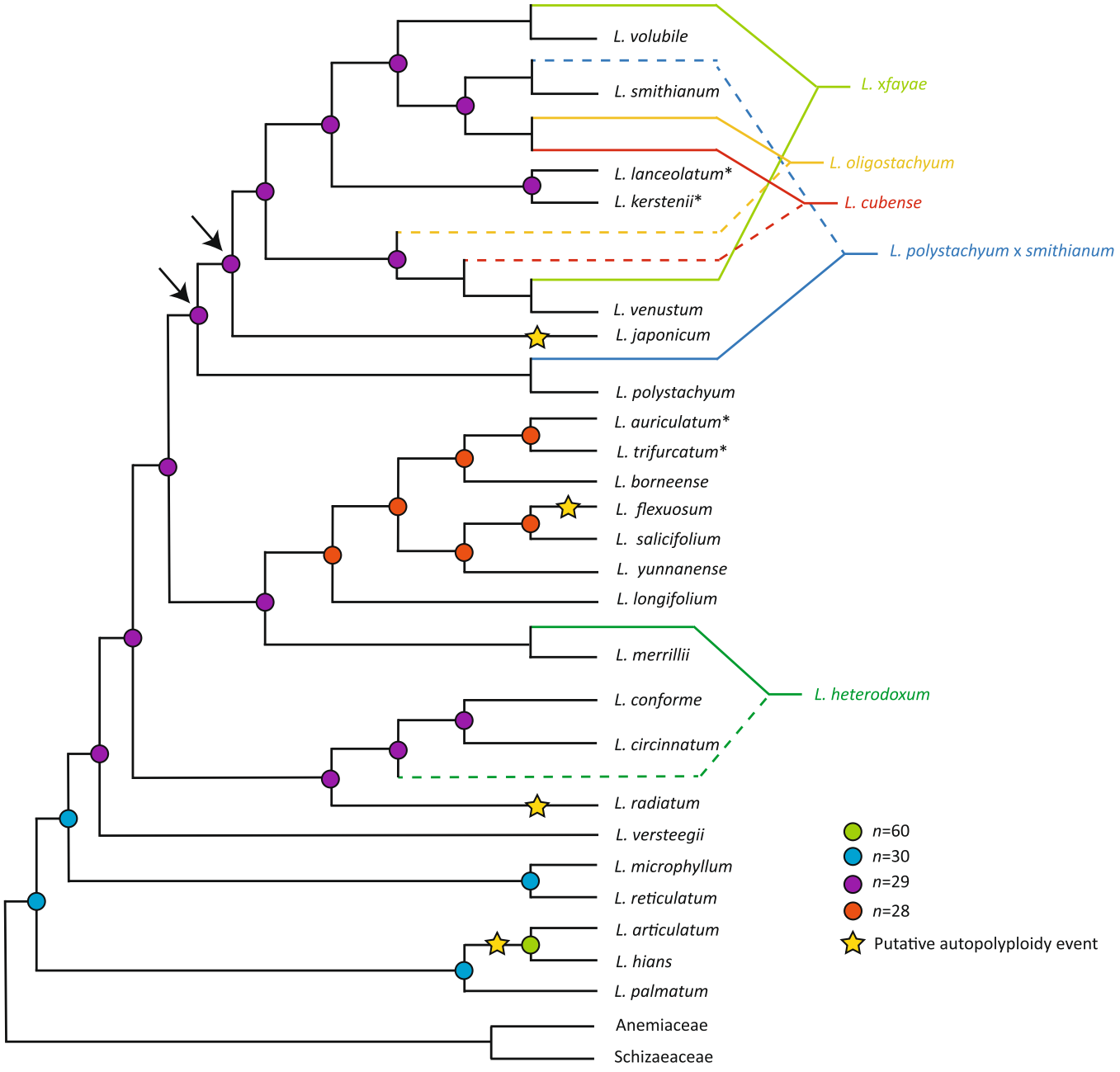

FIGURE 5 Summary species tree with hypothesized reticulation events. Putative autopolyploidy events are denoted with yellow stars. Taxa of hybrid origin are shown with colored branches connecting to both progenitors (solid = maternal parent, dashed = paternal parent). Note that for *Lygodium* ×*fayae* both *L. venustum* and *L. volubile* may be either parent and therefore both are denoted with solid lines. Ancestral reconstruction of haploid chromosome numbers, as inferred with the constant rate model in ChromEvol (Glick and Mayrose, 2014), are given at nodes. Nodes with substantial topological conflicts among analyses are denoted with arrows and taxa with asterisks are consistently polyphyletic across analyses.

dissection, although specimens identified as *L. dimorphum* did not cluster phylogenetically based on frond morphology (Appendix S3). These data support Garrison Hanks's (1998) conclusion that the latter name is a synonym of the former. *Lygodium auriculatum* and *L. semihastatum* present a similar story, with a mix of equivocal support (*L. auriculatum* falling sister to *L. semihastatum*) and strong support (Appendix S1: Table S2) for synonymization of *L. semihastatum* with *L. auriculatum*. In most analyses, there was a clade composed of intermixed *L. trifurcatum* and *L. auriculatum* samples, which are unlikely to be misidentified based on their morphological differences (although phenotypic variation is a possibility). Therefore, future work that includes increased sampling of these taxa and detailed morphological study is needed to clearly delimit species boundaries in this clade.

In every dataset, we found that *L. conforme* and *L. circinnatum* were reciprocally monophyletic, disputing

Garrison Hanks's (1998) synonymization of the names (Appendix S1: Table S2). In addition to this molecular support, an examination of morphology these taxa reveals that the fertile pinnae of *L. circinnatum* have a much greater reduction in the lamina compared to accessions identified as *L. conforme*. There was similar unequivocal support that disputes the synonymization of *L. yunnanense* with *L. flexuosum* by Garrison Hanks (1998) and supports a recognition of the species by Zhang and Garrison Hanks (2013).

Ultimately, these molecular data support the taxonomic treatments of *Lygodium* by Garrison Hanks (1998) and Zhang and Garrison Hanks (2013). In addition to the 26 species presented by Garrison Hanks (1998), we further found topological support for the recognition of *L. yunnanense* (recognized by Ching, 1959; Zhang and Garrison Hanks, 2013 and Hassler, 2024) and *L. conforme* (recognized by Hassler, 2024). Two additional species are recognized by Hassler (2024): *Lygodium altum* (C.B.Clarke) Alderw. and *L. boivinii* Kuhn. The former appears to have a single occurrence and has previously been treated as a variety of *L. flexuosum* by Clarke (1880), while the latter is a likely a hybrid between *L. lanceolatum* and *L. kerstenii* with abortive spores (Garrison Hanks, 1998). In total, our work provides molecular support for 28 species and one hybrid taxon (Table 1), and one putative novel hybrid individual.

Ploidy levels

Despite a long history of cytological study (e.g., Manton and Sledge, 1954; Roy and Manton, 1965; Wagner and Wagner, 1966; see Table 1), there are still many open questions regarding the identity, frequency, and origins of polyploid and hybrid taxa in *Lygodium*. We identified several putative hybrids and polyploids and explored their origins using an integrative approach that incorporated spore measurements, genome size estimates, and several assembly and phasing approaches for target-capture data. Here, we summarize the data across the clade, and we provide a detailed discussion on each putative polyploid and hybrid in the Supplemental Text. We identified several putative autopolyploid taxa based on the phylogenetic placement of phased sequence copies; we emphasize that these taxa require further investigation as several fern taxa once hypothesized to be autopolyploids are in fact allopolyploids derived from closely related taxa (e.g., *Dryopteris campyloptera* Clarkson (Sessa et al., 2012) and *Pellaea ternifolia* (Cav.) Link (Windham et al., 2022)). Such "cryptic allopolyploidy" events require more in-depth study of the individual taxa, such as observation of chromosome pairing behavior (but see Barker et al., 2016 for a discussion on the relative abundance and classification of auto- and allopolyploids).

Cytological data are the most direct type of evidence for determining ploidal level. Manton's (1950) foundational work on the cytology of pteridophytes was instrumental in establishing that polyploidy is rampant throughout the ferns in general, and her subsequent work on *Lygodium* (Manton and Sledge, 1954; Roy and Manton, 1965) revealed the prevalence of both polyploidy and aneuploidy in this genus. Of the fourteen taxa with at least one chromosome count (Table 1), there are eight polyploid taxa (one triploid and one hexaploid [*L. volubile* specimens], the rest tetraploids), and four taxa have counts suggesting more than one cytotype. *Lygodium* has three base chromosome numbers (x = 28, 29, 30; Roy and Manton, 1965). Our ChromEvol analysis inferred an ancestral base number of x = 30, a transition to x = 29 along the backbone, at the divergence of *L. versteegii*, and a transition to x = 28 in a single clade (Figure 5). We supplemented these chromosome count data with genome size estimates through flow cytometry, which can provide a fast way to infer ploidy from a variety of plant tissues (Doležel et al., 2007). However, estimating ploidy from genome sizes alone in *Lygodium* is greatly complicated by the variation of genome sizes among known diploids in *Lygodium* (1 C = 5.56 pg in *L. microphyllum* to 11.33 pg in *L. volubile*), and we currently lack sufficient data to robustly assess genome size across the phylogeny. Although some authors have had success in estimating ploidal level from target-capture data alone (e.g., Viruel et al., 2019), we found our estimates from nQuire to be unreliable, which may be due to low coverage and few positions, or perhaps to fundamental differences in the re-diploidization process in ferns compared to other taxa (Haufler and Soltis, 1986).

Spore morphology and size metrics have traditionally been used in fern systematics (Barrington et al., 1986; Tryon, 1990; Smith, 1995), and as a proxy for genome size, and, consequently, ploidal level (Wagner, 1974; Schuettpelz et al., 2015; Barrington et al., 1989, 2020). While the positive relationship between spore and genome size holds for many clades, Garrison Hanks (1998) did not find a relationship between spore size and ploidal level in *Lygodium* (but see Figure 3B–D). Yet by placing spore size in a phylogenetic context, we see evidence that changes in spore size reflect changes in ploidy level. In the four sets of diploid-polyploid pairs, we found that spore size (area, length, and width) was greater in all four known polyploids relative to their diploid partner. Thus, while other factors may be driving much of the overall variation in genome size seen across *Lygodium*, it appears that ploidal level can still affect differences in spore size in comparisons between closely related taxa.

Taken together, these data demonstrate how an integrative approach reveals the high frequency of polyploidy and the labile nature of genomic evolution across *Lygodium*. While the more traditional methods such as cytology, flow cytometry, and spore measurements appear more useful for identifying polyploids (e.g., Pelosi et al., 2023) than trying to infer ploidy from target-capture data alone, these traditional approaches are even more illuminating in the context of the well-resolved phylogeny produced by target-capture sequencing. In particular, knowing the ploidal levels of many taxa enables us to infer the putative origins of the polyploids.

Reticulation

Integrating our understanding of ploidy throughout *Lygodium* with our target-capture data enables us to develop hypotheses about the identity and origins of polyploids and hybrids within *Lygodium*. We employed two different assembly and phasing pipelines (GoFlag and PATÉ, HybPiper and HybPhaser) and used both manual and model-based (homologizer) subgenome assignments. In general, there was a high level of agreement among results from these methods (Figure 4). Phasing alleles from HTS data provides a wealth of information that can be used to identify putative polyploids and hybrids, including statistics such as the proportion of loci phased, allele divergence (AD), locus heterozygosity (LH), and the rate of read-mapping to diploid references (in HybPhaser). Read-mapping statistics hinted at possible modes of origin, with putative allopolyploid and hybrid samples having an approximately evenly split percentage of reads mapping to putative progenitors (e.g., *L. heterodoxum* to *L. merrillii* and *L.conforme/L. circinnatum*), while putative autopolyploids had one reference that received a much higher percentage of mapped reads (e.g., *L. articulatum* to *L. palmatum*). The information garnered from these statistics can aid in the identification of possible reticulation events, but requires further cytological evidence to verify the ploidal level of the taxa involved.

In many cases, the inferences of hybrid and polyploid origins were strongly supported across different analyses (Figure 4), best exemplified by the putative allopolyploid *L. heterodoxum* and known hybrid *L.* ×*fayae* (Figures 4, 5). We did not find that one method was superior to the others, but rather they were complementary. Despite their utility, there are caveats and drawbacks to each approach. Programs like HybPhaser, PATÉ, and homologizer require some knowledge a priori about the ploidy of samples; one cannot rely solely on phasing metrics or sequencing data to definitively identify polyploids. Incorporating additional cytological data is necessary to confirm sample ploidal levels. PATÉ requires users to specify ploidy for each sample; we limited each sample to a maximum of two copies although additional copies may be recovered by increasing this value for known or putative polyploids. Rather than allowing additional copies, we decided to use the ploidy information inferred from chromosome counts, flow cytometry, and spore measurements to decide from which samples we would retain phased copies. An alternative approach could vary the number of copies based on ploidy information (e.g., diploids would have up to two copies, tetraploids up to four copies), which may capture greater variation and heterozygosity, and could aid in reconstructing reticulate evolution in more detail. Unlike HybPhaser, PATÉ does not assign phased alleles to subgenomes. For about one-third of the loci, we were able to manually assign phased alleles to putative subgenomes, but this is not feasible for trees with hundreds of samples. While homologizer uses a model-based approach to assign subgenomes from phased alleles, it is limited in scale, requiring enormous amounts of computational resources and time, even for only twenty loci. However, using large numbers of loci may not be needed to resolve reticulation (e.g., Schuettpelz et al., 2008; Li et al., 2012; Sessa et al., 2012); most of our analyses recovered similar patterns with varying numbers of loci (Table 2), from 394 loci in HybPhaser datasets to just 20 phased loci used with homologizer.

While not employed by this study, phylogenetic network inference methods also can be used to identify reticulation events and potential parental lineages of polyploids. Networks allow hybrid edges between branches that may be due to hybridization, horizontal gene flow, or introgression. Unlike the programs we used here, most phylogenetic network approaches do not require assigning copies to subgenomes, potentially enabling them to use more data than approaches that require subgenome assignments. The loss of these data in our approaches may hinder our ability to assess the complexity of the underlying evolutionary processes such as ongoing gene flow, the relative contributions of the parental lineages, and patterns of gene retention or diploidization. Several software packages have been developed for model-based network inference (e.g., Than et al., 2008; Solís-Lemus et al., 2017; Wen et al., 2018; Zhang et al., 2018a; Flouri et al., 2020), but so far they are computationally burdensome and generally not scalable to even moderately-sized datasets like those presented here. Therefore, at least in the near future, using a combination of methods will be crucial for unraveling reticulate complexes. While target-capture data are a powerful tool for the analysis of reticulate complexes, generating data for hundreds of loci may be unnecessary for assessing hypotheses of reticulate evolution, and alternative approaches include cloning and Sanger sequencing or long-read sequencing of a handful of informative loci can also be effective for addressing reticulation questions. With all molecular methods, a holistic approach integrating cytological and morphological evidence is needed to inform and verify conclusions about reticulate evolution.

CONCLUSIONS

Identifying and resolving the origins of polyploids remains a challenging task for plant evolutionary biology. While recent advances in sequencing techniques and analytical methods offer much promise for resolving complex evolutionary histories, they may be most effective when synthesized with insights from cytology and morphology. By integrating phylogenomic data from the *GoFlag 408* flagellate land probe set, spore morphology, chromosome number, and genome size evolution, we reconstructed the most comprehensive phylogeny of the clade to date, evaluated the existing classification for the group, and generated hypotheses about the putative origins of polyploids and hybrids in *Lygodium*.

AUTHOR CONTRIBUTIONS

J.A.P., W.L.T., G.J.B., and E.B.S. designed the study. J.A.P., B.A.Z, and E.B.S. secured funding. J.A.P. and W.L.T. sampled specimens. J.A.P., B.A.Z., and E.H.K. performed flow cytometry, and B.A.Z. conducted squashes and counted chromosomes. J.A.P. and E.H.K. imaged and analyzed spores. J.A.P. analyzed the data and wrote the manuscript. All authors read, edited, and approved the final version of the manuscript.

ACKNOWLEDGMENTS

We thank Andrea Carmona, Ellen Lake, and Li-Yaung Kuo for providing spores, Jaroslav Doležel for sharing seeds of genome size standards, the Barbazuk lab at the University of Florida for use of computational resources, University of Florida Research Computing for providing computational resources and support, and two anonymous reviewers and editor for their comments. We are grateful for the herbarium curators and staff for their assistance in obtaining specimen loans including: Alan Franck (FLAS), Lisa Fruscella (NY), Barbara Kennedy (BISH), Matthew von Konrat (F), Matthew Pace (NY), and Meghann Toner (US). Funding for this project was provided in part by the Society of Systematic Biologists Graduate Research Grant, International Association for Plant Taxonomy Research Grant, University of Florida Department of Biology Davis Graduate Fellowship and Grinter Fellowship to J.A.P., Torrey Botanical Society Graduate Research Fellowship to B.A.Z., and NSF IOS #1754911 and #2310485 to E.B.S.

CONFLICT OF INTEREST STATEMENT

Emily Sessa is an associate editor of the *American Journal of Botany* but took no part in the peer review and decision-making process for this paper.

DATA AVAILABILITY STATEMENT

Demultiplexed raw sequence data have been uploaded to NCBI under BioProject PRJNA1040242. Alignments, code, and tree files are available at website https://github.com/jessiepelosi/LygoPhylo.

ORCID

Jessie A. Pelosi http://orcid.org/0000-0002-2861-3445
Bethany A. Zumwalde http://orcid.org/0000-0002-4430-9603
Weston L. Testo http://orcid.org/0000-0003-3194-5763
Emily H. Kim http://orcid.org/0009-0008-6482-0805
J. Gordon Burleigh http://orcid.org/0000-0001-8120-5136
Emily B. Sessa http://orcid.org/0000-0002-6496-5536

REFERENCES

Alston, A. H. G., and R. E. Holttum. 1959. Notes on taxonomy and nomenclature in the genus *Lygodium* (Schizaeaceae). *Reinwardtia* 5: 11–22.

Bankevich, A., S. Nurk, D. Antipov, A. A. Gurevich, M. Dvorkin, A. S. Kulikov, V. M. Lesin, et al. 2012. SPAdes: A new genome assembly algorithm and its applications to single-cell sequencing. *Journal of Computational Biology* 19: 455–477.

Barker, M. S., N. Arrigo, A. E. Baniaga, Z. Li, and D. A. Levin. 2016. On the relative abundance of autopolyploids and allopolyploids. *New Phytologist* 210: 391–398.

Barrington, D. S., C. H. Haufler, and C. R. Werth. 1989. Hybridization, reticulation, and species concepts in the ferns. *American Fern Journal* 79: 55–64.

Barrington, D. S., C. A. Paris, and T. A. Ranker. 1986. Systematic inferences from spore and stomate size in the ferns. *American Fern Journal* 76: 149–159.

Barrington, D. S., N. R. Patel, and M. W. Southgate. 2020. Inferring the impacts of evolutionary history and ecological constraints on spore size and shape in the ferns. *Applications in Plant Sciences* 8: e11339.

Bastide, P., C. Solís-Lemus, R. Kriebel, K. William Sparks, and C. Ané. 2018. Phylogenetic comparative methods on phylogenetic networks with reticulations. *Systematic Biology* 67: 800–820.

Bold, H. C. 1957. Morphology of plants. Harper and Row, New York, New York, USA.

Breinholt, J. W., S. B. Carey, G. P. Tiley, E. C. Davis, L. Endara, S. F. McDaniel, L. G. Neves, et al. 2021a. A target enrichment probe set for resolving the flagellate land plant tree of life. *Applications in Plant Sciences* 9: e11406.

Breinholt, J. W., S. B. Carey, G. P. Tiley, E. C. Davis, L. Endara, S. F. McDaniel, L. G. Neves, et al. 2021b. Target enrichment probe set for resolving the flagellate land plant tree of life. *Dryad Dataset*, website: https://doi.org/10.5061/dryad.7pvmcvdqg

Brownlie, G. 1961. Additional chromosome numbers – New Zealand ferns. *Transactions of the Royal Society of New Zealand, Botany* 1: 1–4.

Cai, L., H. Zhang, and C. C. Davis. 2022. PhyloHerb: A high-throughput phylogenomic pipeline for processing genome skimming data. *Applications in Plant Sciences* 10: e11475.

Camacho, C., G. Coulouris, V. Avagyan, N. Ma, J. Papadopoulos, K. Bealer, and T. L. Madden. 2009. BLAST+: Architecture and applications. *BMC Bioinformatics* 10: 421.

Capella-Gutiérrez, S., J. M. Silla-Martínez, and T. Gabaldón. 2009. trimAl: A tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics* 25: 1972–1973.

Chang, Z., G. Li, J. Liu, Y. Zhang, C. Ashby, D. Liu, C. L. Cramer, and X. Huang. 2015. Bridger: A new framework for de novo transcriptome assembly using RNA-seq data. *Genome Biology* 16: 30.

Ching, R. C. 1959. *Lygodium yunnanense*. *In* S.-S. Chien and W.-Y. Chun, Flora Reipublicae Popularis Sinicae, vol. 2 (Pteridophyta, Ophioglosseae-Oleandreae, 345–346. Typis Academiae Scientiarum Sinicae, Pekini [Beijing], China.

Clark, J., O. Hidalgo, J. Pellicer, H. Liu, J. Marquardt, Y. Robert, M. Christenhusz, et al. 2016. Genome size evolution of ferns: Evidence for relative statis of genome size across the fern phylogeny. *New Phytologist* 210: 1072–1082.

Clarke, C. B. 1880. A review of the ferns of Northern India. *Transactions of the Linnean Society of London (Botany)*, Ser. 2, 1: 425–611.

Copeland, E. B. 1911. Paupan ferns collected by the Reverand Copeland King. *Philippine Journal of Science, Section C, Botany* 6: 65–92.

Doležel, J., J. Greilhuber, S. Lucretti, A. Meister, M. A. Lysák, L. Nardi, and R. Obermayer. 1998. Plant genome size estimation by flow cytometry: Inter-laboratory comparison. *Annals of Botany* 82: 17–26.

Doležel, J., J. Greilhuber, and J. Suda. 2007. Estimation of nuclear DNA content in plants using flow cytometry. *Nature Protocols* 2: 2233–2244.

Doležel, J., S. Sgorbati, and S. Lucretti. 1992. Comparison of three DNA fluorochromes for flow cytometric estimation of nuclear DNA content in plants. *Physiologia Plantarum* 85: 625–631.

Doyle, J. J., and J. L. Doyle. 1987. A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochemical Bulletin* 19: 11–15.

Duek, J. J. 1976. Contribution of the Flora of Cuba: Osmundaceae, Schizaeaceae and Gleicheniaceae (Pteridophyta). *Feddes Repertorium* 87: 325–360.

Edelman, N. B., P. B. Frandsen, M. Miyagi, B. Clavijo, J. Davey, R. B. Dikow, G. García-Accinelli, et al. 2019. Genomic architecture and introgression shape a butterfly radiation. *Science* 366: 594–599.

Edgar, R. C. 2010. Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* 26: 2460–2461.

Faircloth, B. C., J. E. McCormack, N. G. Crawford, M. G. Harvey, R. T. Brumfield, and T. C. Glenn. 2012. Ultraconserved elements anchor thousands of genetic markers spanning multiple evolutionary timescales. *Systematic Biology* 61: 717–726.

Flouri, T., X. Jiao, B. Rannala, and Z. Yang. 2020. A Bayesian implementation of the multispecies coalescent model with introgression for phylogenomic analysis. *Molecular Biology and Evolution* 37: 1211–1223.

Freyman, W. A., M. G. Johnson, and C. J. Rothfels. 2023. homologizer: Phylogenetic phasing of gene copies into polyploid subgenomes. *Methods in Ecology and Evolution* 14: 1230–1244.

Fujiwara, T., H. Liu, E. I. Meza-Torres, R. E. Morero, A. J. Vega, Z. Liang, A. Ebihara, et al. 2023. Evolution of genome space occupation in ferns: Linking genome diversity and species richness. *Annals of Botany* 131: 59–70.

Garrison Hanks, J. 1998. A monographic study of Lygodium Swartz (Pteridophyta: Lygodiaceae). Ph.D. dissertation, City University of New York, New York, USA.

Gastony, G. J., and G. Yatskievych. 1992. Maternal inheritance of the chloroplast and mitochondrial genomes in cheilanthoid ferns. *American Journal of Botany* 79: 716–722.

Glick, L., and I. Mayrose. 2014. ChromEvol: Assessing the pattern of chromosome number evolution and the inference of polyploidy along a phylogeny. *Molecular Biology and Evolution* 31: 1914–1922.

Grant, V. 1971. Plant speciation. Columbia University Press, New York, New York, USA.

Hanson, L., and I. J. Leitch. 2002. DNA amounts for five pteridophyte species fill phylogenetic gaps in C-value data. *Botanical Journal of the Linnean Society, Linnean Society of London* 140: 169–173.

Hassler, M. 2024 World ferns. Synonymic checklist and distribution of ferns and lycophytes of the world, version 19.2. Last update 17 April 2024. Website: www.worldplants.de/ferns/. (accessed 25 April 2024).

Haufler, C. H. 2008. Species and speciation. In T. A. Ranker, and C. H. Haufler [eds.], The biology and evolution of ferns and lycophytes, 303–331. Cambridge University Press, Cambridge, UK.

Haufler, C. H., K. M. Pryer, E. Schuettpelz, E. B. Sessa, D. R. Farrar, R. Moran, J. J. Schneller, et al. 2016. Sex and the single gametophyte: Revising the homosporous vascular plant life cycle in light of contemporary research. *Bioscience* 66: 928–937.

Haufler, C. H., and D. E. Soltis. 1986. Genetic evidence suggests that homosporous ferns with high chromosome numbers are diploid. *Proceedings of the National Academy of Sciences, USA* 83: 4389–4393.

Heilbuth, J. C. 2000. Lower species richness in dioecious clades. *American Naturalist* 156: 221–241.

Hiatt, D., K. Serbesoff-King, D. Lieurance, D. R. Gordon, and S. L. Flory. 2019. Allocation of invasive plant management expenditures for conservation: Lessons from Florida, USA. *Conservation Science and Practice* 1: e51.

Hoang, D. T., O. Chernomor, A. von Haeseler, B. Q. Minh, and L. S. Vinh. 2018. Ufboot2: Improving the ultrafast bootstrap approximation. *Molecular Biology and Evolution* 35: 518–522.

Höhna, S., M. J. Landis, T. A. Heath, B. Boussau, N. Lartillot, B. R. Moore, J. P. Huelsenbeck, and F. Ronquist. 2016. RevBayes: Bayesian phylogenetic inference using graphical models and an interactive model-specification language. *Systematic Biology* 65: 726–736.

Holttum, R. E. 1959. Schizaeaceae. *In* C. G. G. J. Van Steenis, and R. E. Holttum [eds.], Flora Malesiana, II Pteridophyta, 37–62. Martinus Nijhoff, Boston, Massachusetts, USA.

Huang, C.-H., X. Qi, D. Chen, J. Qi, and H. Ma. 2019. Recurrent genome duplication events likely contributed to both the ancient and recent rise of ferns. *Journal of Integrative Plant Biology* 62: 433–455.

Jin, J.-J., W.-B. Yu, J.-B. Yang, Y. Song, C. W. dePamphilis, T.-S. Yi, and D.-Z. Li. 2020. GetOrganelle: A fast and versatile toolkit for accurate de novo assembly of organelle genomes. *Genome Biology* 21: 241.

Johnson, M. G., E. M. Gardner, Y. Liu, R. Medina, B. Goffinet, A. J. Shaw, N. J. C. Zerega, and N. J. Wickett. 2016. HybPiper: Extracting coding sequence and introns for phylogenetics from high-throughput sequencing reads using target enrichment. *Applications in Plant Sciences* 4: 1600016.

Kalyaanamoorthy, S., B. Q. Minh, T. K. F. Wong, A. von Haeseler, and L. S. Jermiin. 2017. ModelFinder: Fast model selection for accurate phylogenetic estimates. *Nature Methods* 14: 587–589.

Karimi, N., C. E. Grover, J. P. Gallagher, J. F. Wendel, C. Ané, and D. A. Baum. 2020. Reticulate evolution helps explain apparent homoplasy in floral biology and pollination in Baobabs (Adansonia; Bombacoideae; Malvaceae). *Systematic Biology* 69: 462–478.

Katoh, K., and D. M. Standley. 2013. MAFFT multiple sequence alignment software, version 7: Improvements in performance and usability. *Molecular Biology and Evolution* 30: 772–780.

Knobloch, I. W. 1976. Pteridophyte hybrids. *Publications of the Museum, Michigan State University, Biological Series* 5(4): 273-352.

Koop, A. L. 2009. *Lygodium microphyllum* (Old world climbing fern), *Lygodium japonicum* (Japanese climbing fern), and *Lygodium flexuosum*. *In* A. Jackson [ed.]. Weed Risk Assessment. APHIS, U. S. Department of Agriculture, Raleigh, North Carolina, USA.

Kramer, K. U. 1990. Schizaeaceae. *In* K. U. Kramer [ed.], Pteridophytes and gymnosperms, 258- 263. Springer Verlag, Berlin, Germany.

Kuo, L.-Y., and F.-W. Li. 2019. A roadmap for fern genome sequencing. *American Fern Journal* 109: 212–223.

Kuo, L.-Y., T.-Y. Tang, F.-W. Li, H.-J. Su, W.-L. Chiou, Y.-M. Huang, and C.-N. Wang. 2018. Organelle genome inheritance in *Deparia* ferns (Athyriaceae, Aspleniineae, Polypodiales). *Frontiers in Plant Science* 9: 486.

Li, F.-W., K. M. Pryer, and M. D. Windham. 2012. *Gaga*, a new fern genus segregated from *Cheilanthes* (Pteridaceae). *Systematic Botany* 37: 845–860.

Li, H., and R. Durbin. 2009. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* 25: 1754–1760.

Li, H., B. Handsaker, A. Wysoker, T. Fennell, J. Ruan, N. Homer, G. Marth, et al. 2009. The Sequence Alignment/Map format and SAMtools. *Bioinformatics* 25: 2078–2079.

Lin, S.-J., H.-R. Zhuang, K. Iwatsuki, and H.-S. Lu. 2002. Cytotaxonomic study of ferns from China II. Species of Fujian. *Japanese Journal of Botany* 77: 129-138.

Linder, C. R., and L. Rieseberg. 2004. Reconstructing patterns of reticulate evolution in plants. *American Journal of Botany* 91: 1700–1708.

Lynch, M. 2007. The origins of genome architecture. Sinauer Associates, Inc., Sunderland, Massachusetts, USA.

Madeira, P. T., R. W. Pemberton, and T. D. Center. 2008. A molecular phylogeny of the genus *Lygodium* (Schizaeaceae) with special reference to the biological control and host range testing of *Lygodium microphyllum*. *Biological Control* 45: 308–318.

Manton, I. 1950. Problems of cytology and evolution in the Pteridophyta. Oxford University Press, Oxford, UK.

Manton, I., and W. A. Sledge. 1954. Observations on the cytology and taxonomy of the pteridophyte flora of Ceylon. *Philosophical Transactions of the Royal Society, B, Biological Sciences* 238: 127–185.

McKenna, A., M. Hanna, E. Banks, A. Sivachenko, K. Cibulskis, A. Kernytsky, K. Garimella, et al. 2010. The genome analysis toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Research* 20: 1297–1303.

Mendez-Reneau, J., J. Gordon Burleigh, and E. M. Sigel. 2023. Target capture methods offer insight into the evolution of rapidly diverged taxa and resolve allopolyploid homeologs in the fern genus *Polypodium* s.s. *Systematic Botany* 48: 96–109.

Minh, B. Q., H. A. Schmidt, O. Chernomor, D. Schrempf, M. D. Woodhams, A. von Haeseler, and R. Lanfear. 2020. IQ-TREE 2: New models and efficient methods for phylogenetic inference in the genomic era. *Molecular Biology and Evolution* 37: 1530–1534.

Mitui, K. 1965. Chromosome studies on Japanese ferns. *Journal of Japanese Botany* 40: 117–124.

Mitui, K. 1968. Chromosomes and speciation in ferns. *Science Reports of the Tokyo Kyoiku Daigaku, Section B, Zoology and Botany* 13: 285-333.

Morton, C. 1966. The use of climbing fern, *Lygodium*, in weaving. *American Fern Journal* 56: 79–81.

Nakato, N. 1990. Notes on chromosomes of Japanese pteridophytes. *Journal of Japanese Botany* 65: 204-209.

Nakato, N., and M. Kato. 2001. Chromosome numbers of ten species of ferns from Guam, S. Mariana Isls. *Acta Phytotaxonomica et Geobotanica* 52: 125–133.

Nauheimer, L., N. Weigner, E. Joyce, D. Crayn, C. Clarke, and K. Nargar. 2021. HybPhaser: A workflow for the detection and phasing of hybrids in target capture data sets. *Applications in Plant Sciences* 9: e11441.

Nitsch, J. P. 1951. Growth and development in vitro of excised ovaries. *American Journal of Botany* 38: 566–577.

One Thousand Plant Transcriptomes Initiative. 2019. One thousand plant transcriptomes and the phylogenomics of green plants. *Nature* 574: 679–685.

Pelosi, J. A., E. H. Kim, W. B. Barbazuk, and E. B. Sessa. 2022. Phylotranscriptomics illuminates the placement of whole genome duplications and gene retention in ferns. *Frontiers in Plant Science* 13: 882441.

Pelosi, J. A., B. A. Zumwalde, O. Hornych, K. Wheatley, E. H. Kim, and E. B. Sessa. 2023. *Lygodium japonicum* (Lygodiaceae) is represented by a tetraploid cytotype in Florida. *American Fern Journal* 113: 43–55.

Pemberton, R. W., and A. P. Ferriter. 1998. Old World Climbing Fern (*Lygodium microphyllum*), a dangerous invasive weed in Florida. *American Fern Journal* 88: 165–175.

PPG I. 2016. A community-derived classification for extant lycophytes and ferns. *Journal of Systematics and Evolution* 54: 563–603.

Prantl, K. 1881. Schizaeaceae. Untersuchungen zur Morphologie der Gefässkrptogamen, 7–85. Engelmann Verlag, Leipzig, Germany.

Rahayu, M., E. Sri Kuncari Mahdawia, and M. Setiawan. 2020. Short Communication: Ethnobotanical study of *Lygodium circinnatum* and its utilization in crafts weaving in Indonesia. *Biodiversitas* 21: 617–621.

Rambaut, A., A. J. Drummond, D. Xie, G. Baele, and M. A. Suchard. 2018. Posterior summarization in Bayesian phylogenetics using Tracer 1.7. *Systematic Biology* 67: 901–904.

Ranker, T. A., M. J. Balick, G. M. Plunkett, K. D. Harrison, J.-P. Wahe, and M. Wahe. 2022. Ethnobotany and vernacular names of the lycophytes and ferns of Tafea province, Vanuatu. *American Fern Journal* 112: 143–177.

Reed, C. F. 1946. The phylogeny and ontogeny of the Pteropsida. I. Schizeales. *Boletim da Sociedade Broteriana* 21: 71–197.

Rice, A., L. Glick, S. Abadi, M. Einhorn, N. M. Kopelman, A. Salman-Minkov, J. Mayzel, et al. 2015. The Chromosome Counts Database (CCDB) - a community resource of plant chromosome numbers. *New Phytologist* 206: 19–26.

Rothfels, C. J., A. K. Johnson, P. H. Hovenkamp, D. L. Swofford, H. C. Roskam, C. R. Fraser-Jenkins, M. D. Windham, and K. M. Pryer. 2015. Natural hybridization between genera that diverged from each other approximately 60 million years ago. *American Naturalist* 185: 433–442.

Roy, S. K., and I. Manton. 1965. A new base number in the genus *Lygodium*. *New Phytologist* 64: 286–292.

Schneider, C. A., W. S. Rasband, and K. W. Eliceiri. 2012. NIH Imagine to ImageJ: 25 years of image analysis. *Nature Methods* 9: 671-675.

Schuettpelz, E., A. Grusz, M. D. Windham, and K. M. Pryer. 2008. The utility of gapCp in resolving polyploid fern origins. *Systematic Botany* 33: 621–629.

Schuettpelz, E., and K. M. Pryer. 2007. Fern phylogeny inferred from 400 leptosporangiate species and three plastid genes. *Taxon* 56: 1037–1050.

Schuettpelz, E., K. M. Pryer, and M. D. Windham. 2015. A unified approach to taxonomic delimitation in the fern genus *Pentagramma* (Pteridaceae). *Systematic Botany* 40: 629–644.

Sessa, E. B., E. A. Zimmer, and T. J. Givnish. 2012. Unraveling reticulate evolution in North American *Dryopteris* (Dryopteridaceae). *BMC Evolutionary Biology* 12: 104.

Shen, W., S. Le, Y. Li, and F. Hu. 2016. SeqKit: A cross-platform and ultrafast toolkit for FASTA/Q file manipulation. *PLoS One* 11: e0163962.

Sigel, E. M. 2016. Genetic and genomic aspects of hybridization in ferns. *Journal of Systematics and Evolution* 54: 638–655.

Slater, G. S. C., and E. Birney. 2005. Automated generation of heuristics for biological sequence comparison. *BMC Bioinformatics* 6: 31.

Smith, A. R. 1995. Non-molecular phylogenetic hypotheses for ferns. *American Fern Journal* 85: 104–122.

Solís-Lemus, C., P. Bastide, and C. Ané. 2017. Phylonetworks: A package for phylogenetic networks. *Molecular Biology and Evolution* 34: 3292–3298.

Soltis, P. S., and D. E. Soltis. 2009. The role of hybridization in plant speciation. *Annual Review of Plant Biology* 60: 561–588.

Stull, G. W., K. K. Pham, P. S. Soltis, and D. E. Soltis. 2023. Deep reticulation: The long legacy of hybridization in vascular plant evolution. *Plant Journal* 114: 743–766.

Takamiya, M. 1995. Chromosomal studies of ferns and fern-allies in the Republics of Fiji and Vanuatu, South Pacfic I. *Acta Phytotaxonomica and Geobotanica* 46: 137–145.

Testo, W., and M. Sundue. 2016. A 4000-species dataset provides new insight into the evolution of ferns. *Molecular Phylogenetics and Evolution* 105: 200–211.

Than, C., D. Ruths, and L. Nakhleh. 2008. PhyloNet: A software package for analyzing and reconstructing reticulate evolutionary relationships. *BMC Bioinformatics* 9: 322.

Tiley, G. P., A. A. Crowl, P. S. Manos, E. B. Sessa, C. Solís-Lemus, A. D. Yoder, and J. G. Burleigh. 2024. Benefits and limits of phasing alleles for network inference of allopolyploid complexes. *Systematic Biology* 2024: syae024.

Tindale, M. D., and S. K. Roy. 2002. A cytotaxonomic survey of the Pteridophyta of Australia. *Australian Systematic Botany* 15: 839–937.

Tryon, A. F. 1990. Fern spores: Evolutionary levels and ecological differentiation. *In* M. Hesse, and F. Ehrendorfer [eds.], Morphology, development, and systematic relevance of pollen and spores, Plant systematics and evolution, Supplementum 5:71–79. Springer Verlag, Vienna, Austria.

Viruel, J., M. Conejero, O. Hidalgo, L. Pokorny, R. F. Powell, F. Forest, M. B. Kantar, et al. 2019. A target capture-based method to estimate ploidy from herbarium specimens. *Frontiers in Plant Science* 10: 937.

Vogel, J. C., S. J. Russell, F. J. Rumsey, J. A. Barrett, and M. Gibby. 1998. Evidence for maternal transmission of chloroplast DNA in the genus *Asplenium* (Aspleniaceae, Pteridophyta). *Botanica Acta* 111: 247–249.

Wagner, W. H. 1974. Structure of spores in relation to fern phylogeny. *Annals of the Missouri Botanical Garden* 61: 332–353.

Wagner, W. H., D. R., Farrar, and B. W. McAlpin. 1970. Pteridology of the Highlands Biological Station area, southern Appalachians. *Journal of the Elisha Mitchell Scientific Society* 86: 1–27.

Wagner, W. H., and F. S. Wagner. 1966. Pteridophytes of the Mountain Lake Area, Giles Co., Virginia: Biosystematic Studies 1964-65. *Castanea* 31: 121–140.

Walker, T. G. 1958. Hybridization in some species of *Pteris* L. *Evolution* 12: 82–92.

Walker, T. G. 1966. IX.—A cytotaxonomic survey of the pteridophytes of Jamaica. *Transactions of the Royal Society of Edinburgh* 66: 169–237.

Walker, T. G. 1985. Cytotaxonomic studies of the ferns of Trinidad II. *Bulletin of the British Museum (Natural History), Botany* 13: 149–249.

Weiß, C. L., M. Pais, L. M. Cano, S. Kamoun, and H. A. Burbano. 2018. nQuire: A statistical framework for ploidy estimation using next generation sequencing. *BMC Bioinformatics* 19: 122.

Wen, D., Y. Yu, J. Zhu, and L. Nakhleh. 2018. Inferring phylogenetic networks using phylonet. *Systematic Biology* 67: 735–740.

Whittier, P., and W. H. Wagner. 1971. The variation in spore size and germination in *Dryopteris* taxa. *American Fern Journal* 61: 123–127.

Wickell, D. A., and F.-W. Li. 2020. On the evolutionary significance of horizontal gene transfers in plants. *New Phytologist* 225: 113–117.

Wikström, N., P. Kenrick, and J. C. Vogel. 2002. Schizaeaceae: A phylogenetic approach. *Review of Palaeobotany and Palynology* 119: 35–50.

Windham, M. D., L. Huiet, J. S. Metzgar, T. A. Ranker, G. Yatskievych, C. H. Haufler, and K. M. Pryer. 2022. Once more unto the breach, dear friends: Resolving the origins and relationships of the *Pellaea wrightiana* hybrid complex. *American Journal of Botany* 109: 821–850.

Wood, T. E., N. Takebayashi, M. S. Barker, I. Mayrose, P. B. Greenspoon, and L. H. Rieseberg. 2009. The frequency of polyploid speciation in vascular plants. *Proceedings of the National Academy of Sciences, USA* 106: 13875–13879.

Xie, M., Q. Wu, J. Wang, and T. Jiang. 2016. H-PoP and H-PoPG: Heuristic partitioning algorithms for single individual haplotyping of polyploids. *Bioinformatics* 32: 3735–3744.

Zhang, C., H. A. Ogilvie, A. J. Drummond, and T. Stadler. 2018a. Bayesian inference of species networks from multilocus sequence data. *Molecular Biology and Evolution* 35: 504–517.

Zhang, C., M. Rabiee, E. Sayyari, and S. Mirarab. 2018b. ASTRAL-III: Polynomial time species tree reconstruction from partially resolved gene trees. *BMC Bioinformatics* 19: 153.

Zhang, X., and J. Garrison Hanks. 2013. Lygodiaceae. *In* Z. Y. Wu, P. H. Raven, and D. Y. Hong [eds.], Flora of China, 2-3 (Pteridophytes), 118–121. Missouri Botanical Garden Press, St. Louis, Missouri, USA.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

Appendix S1:

Table S1. Voucher information for samples used in this study.

Table S2. Molecular evidence for/against the synonymization of *Lygodium* taxa.

Table S3. Average phasing statistics for putative polyploid and hybrid taxa.

Appendix S2. Chromosome squashes of *Lygodium oligostachyum* and *L. volubile*.

Appendix S3. Phylogenies constructed from the datasets generated in this study. Sample IDs are based on Table S1 column 'Sample'.

Appendix S4. Annotated ASTRAL phylogeny of phased locus copies from the supercontig HybPhaser dataset. Locus copies from individual samples are connected by thin black lines. Node support values are given as local posterior probabilities, with * denoting an LPP of 1.

Appendix S5. Summary phylogeny of *Lygodium* with classifications from Garrison Hanks (1998) depicted. Black squares denote membership of that taxon to the subgenus/section.

Appendix S6. Discussion of hypothesized polyploids, hybrids, and putative origins of *Lygodium* taxa.

How to cite this article: Pelosi, J. A., B. A. Zumwalde, W. L. Testo, E. H. Kim, J. G. Burleigh, and E. B. Sessa. 2024. All tangled up: Unraveling phylogenetics and reticulate evolution in the vining ferns, *Lygodium* (Schizaeales). *American Journal of Botany* 111(9): e16389. https://doi.org/10.1002/ajb2.16389