Target Capture Methods Offer Insight into the Evolution of Rapidly Diverged Taxa and Resolve Allopolyploid Homeologs in the Fern Genus Polypodium s.s.

Authors: Mendez-Reneau, Jonas, Burleigh, J. Gordon, and Sigel, Erin M.

Source: Systematic Botany, 48(1) : 96-109

Published By: The American Society of Plant Taxonomists

URL: https://doi.org/10.1600/036364423X16758873924135

Systematic Botany (2023), 48(1): pp. 96–109
ISSN 0363-6445 (print); ISSN 1548-2324 (online)

DOI 10.1600/036364423X16758873924135
Date of publication March 28, 2023

Target Capture Methods Offer Insight into the Evolution of Rapidly Diverged Taxa and Resolve Allopolyploid Homeologs in the Fern Genus Polypodium s.s.

Jonas Mendez-Reneau,[1,4] J. Gordon Burleigh,[2] and Erin M. Sigel[3]

[1]Department of Biology, University of Louisiana at Lafayette, 410 East St. Mary Boulevard, Lafayette, Louisiana 70504, USA
[2]Department of Biology, University of Florida, P.O. Box 118525, Gainesville, Florida 32611, USA
[3]Department of Biological Sciences, University of New Hampshire, 38 Academic Way, Durham, New Hampshire 03824, USA
[4]Author for correspondence (jonas.mendez-reneau1@louisiana.edu)

Communicating Editor: Adriano Stinca

Abstract—Like many fern lineages comprising reticulate species complexes, Polypodium s.s. (Polypodiacaeae) has a history shaped by rapid diversification, hybridization, and polyploidy that poses substantial challenges for phylogenetic inference with plastid and single-locus nuclear markers. Using target capture probes for 408 nuclear loci developed by the GoFlag project and a custom bioinformatic pipeline, SORTER, we constructed multi-locus nuclear datasets for diploid temperate and Mesoamerican species of Polypodium and five allotetraploid species belonging to the well-studied Polypodium vulgare complex. SORTER employs a clustering approach to separate putatively paralogous copies of targeted loci into orthologous matrices and haplotype phasing to infer allopolyploid haplotypes across loci, resulting in datasets amenable to both concatenated maximum likelihood and multi-species coalescent phylogenetic analyses. By comparing phylogenies derived from maximum likelihood and multi-species coalescent analyses of unphased and phased datasets, as well as evaluating discordance among gene trees and species trees, we recover support for incomplete lineage sorting within Polypodium s.s., novel relationships among diploid taxa of the Polypodium vulgare complex and its Mesoamerican sister clade, and the placement of several Polypodium species within other genera. Additionally, we were able to infer well-supported phylogenies that identified the hypothesized progenitors of the allotetraploid species, indicating that SORTER is an effective and accurate tool for reconstructing homeolog haplotypes of allopolyploids in fern taxa and other non-model organisms from target capture data.

Keywords—Hyb-seq, incomplete lineage sorting, multi-species coalescent, reticulate evolution.

Once considered a single northern circum-temperate taxon, the Polypodium vulgare complex comprises 10 diploid (Sigel et al. 2014a) and 14 allopolyploid species (Haufler et al. 1995a, 1995b). Containing the type species of Polypodium, Polypodium vulgare L., the complex exhibits many common characteristics of the genus such as an epilithic or epiphytic habit, long creeping rhizomes with knob-like prominences, round ex-indusiate sori, and pinnatifid blades (Haufler et al. 1993). Traditionally, subtle variations in macro- and micro-morphological characters such as venation patterns, the presence of glandular hairs, and rhizome scale morphology were used to delineate taxa (Peterson and Kott 1974; Haufler et al. 1993). Hybridization experiments, cytotaxonomic studies, and isozyme electrophoresis analyses revealed that taxonomic confusion in the group is exacerbated by the presence of both diploid and morphologically intermediate allopolyploid species (Manton 1947, 1950, 1951, 1957; Kott and Britton 1982; Cranfill and Britton 1983; Bryan and Soltis 1987; Haufler and Windham 1991; Haufler and Zhongren 1991; Whitmore and Smith 1991). In combination, prior work established hypotheses about the major clades of diploid species within the P. vulgare complex and the progenitors of the allopolyploid taxa (Fig. 1). Most recently, Sigel et al. (2014a, 2014b) used maternally inherited plastid sequences and a single biparentally inherited nuclear locus to test the monophyly of the P. vulgare complex, infer relationships within and divergence times for four major clades of diploid species (Fig. 1, the A, C, G, and S clades), and con-firm the progenitors of one allotetraploid species, Polypodium hesperium.

Despite our increased understanding of the Polypodium vulgare complex, the relationships among some diploid species and clades in the genus remain unresolved, and the hypothesized progenitors of most allopolyploid species remain untested with phylogenetic analysis using biparentally inherited nuclear sequence data. Furthermore, relationships among the Mesoamerican members of Polypodium s.s. (Fig. 1, the M clade; Sigel et al. 2014a), which comprise the sister group to the P. vulgare complex, and closely related outgroup genera are still largely unresolved (Otto et al. 2009).

Targeted sequence capture methods combined with high-throughput sequencing (henceforth, TC) are increasingly used for generating multi-locus nuclear sequence datasets for phylogenetic inference of diverse taxa, especially non-model organisms (Mandel et al. 2014; Johnson et al. 2016; Branstetter et al. 2017; Andermann et al. 2019; Faircloth et al. 2020; Beck et al. 2021). Encompassing numerous methodical variations such as target-enrichment (Mamanova et al. 2010) and Hyb-Seq (Weitemier et al. 2014), TC generally refers to hybridizing custom RNA baits or probes to complementary regions of a sample's genomic DNA. Hybridized DNA regions are then captured, often amplified with PCR, and sequenced (Andermann et al. 2020). A significant advantage of TC is that it can simultaneously capture sequence data for hundreds to thousands of biparentally-inherited, unlinked nuclear loci at a relatively low cost compared to other next generation sequencing approaches (McCormack et al. 2013; Smith et al. 2014). When TC probes are designed to capture highly conserved genomic regions, such as exons or Ultra-Conserved Elements (UCEs), the same probe set may be used to obtain orthologous loci for species separated by many millions of years of evolution (e.g. Faircloth et al. 2012; Johnson et al. 2019). Additionally, when sequenced at high read depths, TC can recover sequence variants representing heterozygous alleles, homeologous variants present in hybrid and allopolyploid taxa, and paralogous gene copies. The sequence variation captured with TC has strong potential for inferring the relationships within groups with histories of genomic duplication, reticulate evolution, and rapid diversification (Nicholls et al. 2015; Johnson et al. 2016; Kamneva et al. 2017; Kates et al. 2018; Andermann et al. 2019).

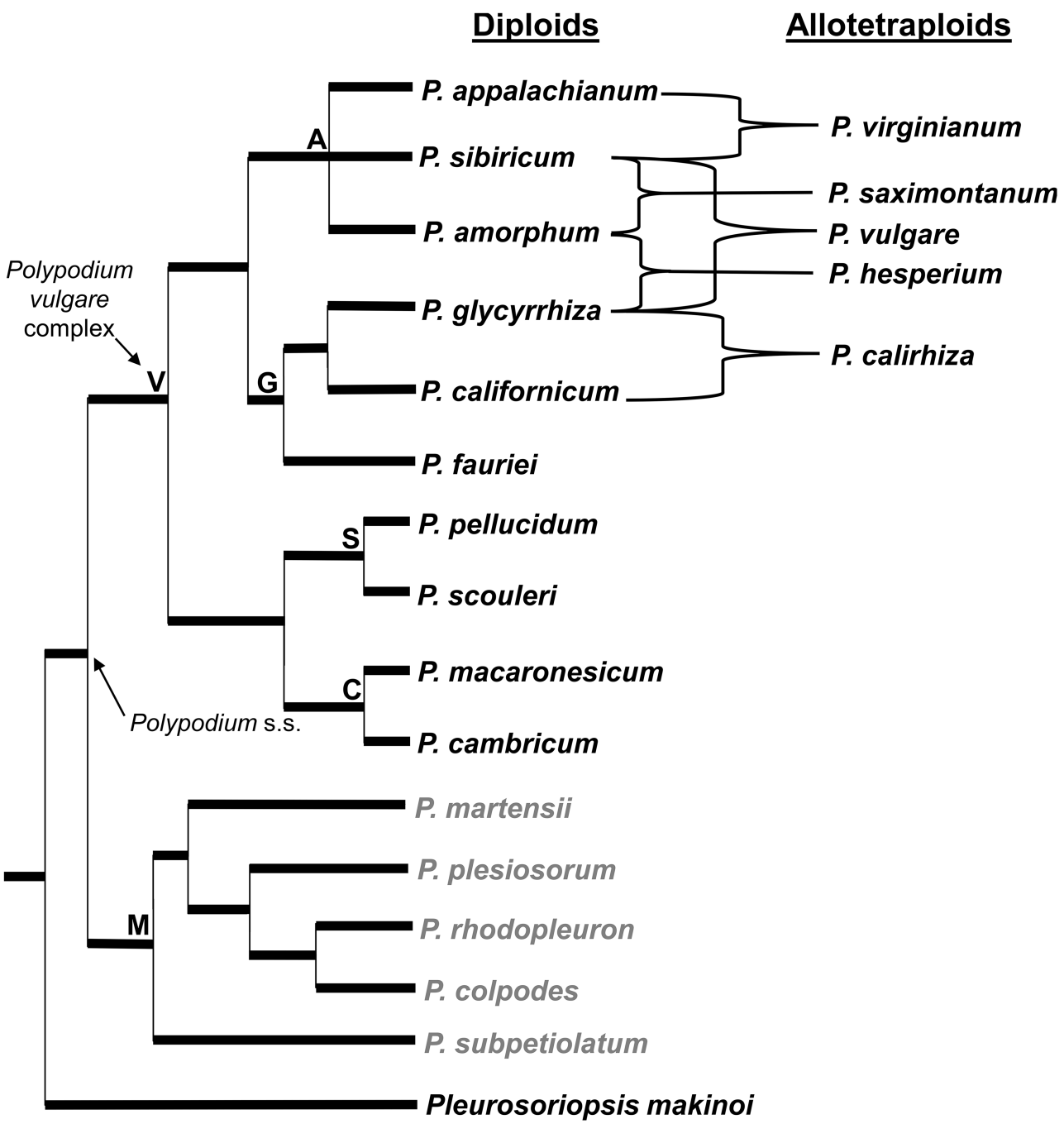

FIG. 1. Relationships among diploid species of Polypodium sensu stricto based on the chloroplast phylogeny of Sigel et al. (2014a). Species of known ploidy (diploid or allotetraploid) are shown in black font, and species of unknown ploidy are shown in gray font. Brackets connect five allotetraploid species to their hypothesized progenitors (Haufler et al. 1995b); Capitalized letters indicate major clades within Polypodium s.s.: V 5 Polypodium vulgare complex; M 5 Mesoamerican clade; A 5 appalachianum clade; G 5 glycyrrhiza clade; C 5 cambricum clade; S 5 scouleri clade.

Here we use of the Geneology of Flagellate Plants Project (GoFlag) 408 TC probe set and a custom bioinformatic processing pipeline, SORTER, to infer relationships within the Polypodium vulgare reticulate complex and its Mesoamerican sister lineage. The GoFlag 408 TC probe set was designed to recover 408 highly conserved nuclear exons from 248 low copy nuclear genes across all flagellate land plant lineages (Breinholt et al. 2021), and with sequences from the more variable regions flanking the target exons, it can be useful for resolving lower-level phylogenetic relationships (e.g. Fawcett et al. 2021; Bloesch et al. 2022). The SORTER pipeline uses de novo assembly and sequence similarity clustering algorithms to separate non-orthologous copies of targeted loci when present in diploid taxa and uses paired-end read phasing coupled with UBLAST relative to potential diploid progenitors to infer homeologs in allopolyploids. We use both maximum likelihood (ML) phylogenetic reconstruction on unphased and phased concatenated nuclear datasets and multi-species coalescent (MSC) analysis of phased sequences to elucidate the relationships within and among major clades of the Polypodium vulgare complex, the broader Polypodium s.s., and the parentage of five allopolyploid taxa.

Materials and Methods

Taxon Sampling—We sampled 63 Polypodium herbarium specimens collected between 1939 and 2017 representing 10 diploid species and five allopolyploids belonging to the northern temperate P. vulgare reticulation complex, as well as 17 samples representing 10 Mesoamerican species of Polypodium (Appendix 1). Multiple individuals of some taxa were sampled to represent broad geographic distributions, morphological variants, and hypothesized cryptic taxa (Haufler and Windham 1991; Haufler et al. 1995b; Shalimov et al. 2011; Shalimov and Shmakov 2016). Mesoamerican species were selected without regard to ploidy because of a lack of cytogenetic data for these taxa. Our outgroup comprised one or two species belonging to each of six genera (Pleurosoriopsis, Pleopeltis, Serpocaulon, Pecluma, Phlebodium, and Platycerium) broadly representing the diversity within Polypodiaceae (Appendix 1; Schneider et al. 2004; Schneider 2006; Otto et al. 2009; Assis et al. 2016). Pleurosoriopsis makinoi was included in outgroup sampling based on its position in recent molecular studies as the sister species to Polypodium s.s. (Otto et al. 2009; Sigel et al. 2014a).

DNA Extraction, Library Construction, Target Capture, and High-Throughput Sequencing—DNA was extracted from silica dried material or a sample of an herbarium voucher for each specimen using a modified cetyltrimethylammonium bromide (CTAB) protocol (Doyle and Doyle 1987; see Breinholt et al. 2021). The library construction, targeted sequence capture using the GoFlag 408 probe set, and Illumina sequencing was done by RAPiD Genomics (Gainesville, FL) using protocols described in Fawcett et al. (2021) as part of the GoFlag project. Read data for all specimens is available via the NCBI short read archive (SRA; BioProject: PRJNA841228).

Assembly of Nuclear Target Capture Sequences Using the SORTER Pipeline—A new pipeline, Sorter of Orthologous Regions for Target Enrichment Reads (SORTER), was developed to generate matrices of orthologous sequences from diploid taxa and homeologs of allopolyploid taxa for species tree inference (https://github.com/JonasMendez/SORTER). Briefly, the SORTER pipeline comprises three sets of Python scripts referred to as stages that integrate existing software to accomplish the following: (stage 1) Trim raw pair-ended Illumina reads, de novo assemble consensus allele sequences, and generate alignments of putatively orthologous gene copies for diploid taxa; (stage 2) Generate alignments of phased orthologous alleles for diploid taxa; and (stage 3) Trim raw pair-ended Illumina reads, de novo assemble contigs, and generate alignments with phased homeolog sequences for allopolyploid taxa. Unique features of SORTER include user customizable sequence filtering options (e.g. the number and minimum sequence length of contigs per reference retained) and sequence clustering using identity thresholds to generate sets of orthologous sequences that can be phased into allelic haplotypes. The SORTER pipeline also uses UBLAST to automatically annotate homeologous variants across loci, approximating the parental haplotypes contained within allopolyploids based on diploid sampling without requiring a priori assumptions about parental identity. A detailed description of each stage of the SORTER pipeline and the specific parameters used for this study are provided in Supplemental A (Mendez-Reneau et al. 2022) and illustrated in Fig. 2.

A total of 23 aligned datasets for diploid and allopolyploid sequences were generated for phylogenetic analyses using the SORTER pipeline (Table 1). To facilitate comparison among datasets and phylogenetic analyses, each dataset included sequences from the same diploid temperate Polypodium, Mesoamerican, and outgroup samples and used loci generated with a 70% identity clustering threshold for stage one of the SORTER pipeline (Supplemental B, Mendez-Reneau et al. 2022). For all datasets, all included loci were filtered with TriFusion software (Silva et al. 2018) to output alignments with $\geq$ 50% sample representation. Each alignment was then imported into Sequence Matrix (Vaidya et al. 2011) to convert flanking gaps into missing data and to generate the final unconcatenated and concatenated alignments (Table 1), keeping only the longest sequence for samples that retained multiple consensus-alleles, putative allelic variants, or homeologs after clustering.

Datasets A, B, and C were generated to compare the effects of ML analyses and MSC analyses, as well as phasing, on the relationships and node support among the diploid temperate members of the P. vulgare complex and the Mesoamerican species of Polypodium (Table 1). Specifically, datasets A and B comprised individual locus alignments that were unphased and phased, respectively, and concatenated together prior to ML analyses. For the phased dataset B, alleles from each sample were randomly concatenated across loci due to lack of genomic haplotype information (Lischer et al. 2014), producing two concatenated sequences per sample. In contrast, dataset C comprised phased individual locus alignments for MSC analysis. Datasets D1–E1 were generated to determine the progenitors of five allotetraploid species belonging to P. vulgare complex (Table 1; Fig. 1). Each dataset incorporated putative homeolog sequences for one of two specimens of an allotetraploid species, one phased concatenated dataset for ML analysis and one phased unconcatenated dataset for MSC analysis for each of two specimens (e.g. P. vulgare 1 and P. vulgare 2), resulting in four datasets per species.

Maximum Likelihood and Bayesian Inference of Concatenated Datasets—Each concatenated dataset (Table 1, datasets A, B, D1–D10) was analyzed with maximum likelihood (ML) using RAxML (Stamatakis 2014) on the CIPRES computing cluster (Miller et al. 2010). Prior to ML analysis, all loci were partitioned and independently assigned the GTRGAMMA model. One independent ML search was performed for each dataset with 20 replicates each (Table 1). One thousand bootstrap pseudo-replicates were performed for each dataset.

For the concatenated alignments that included putative homeolog sequences of allopolyploid taxa (Table 1, datasets D1–D10), we employed an additional concatenation step. Based on the topology of the ML phylogeny with the largest -ln score, we further concatenated homeolog sequences that were sister to one another on the tree using Sequence Matrix (Vaidya et al. 2011) to generate the final homeolog sequences. We then filtered loci using TriFusion software (Silva et al. 2018) to retain loci with two putative homeologs (i.e. loci that had homeolog sequences associated with two putative diploid progenitors) per allopolyploid sample and with $\geq$ 50% sample representation (Supplemental C, Mendez-Reneau et al. 2022).

Maximum Likelihood Reconstruction of Individual Locus Datasets—Individual 'gene' trees were reconstructed for all datasets (Table 1; Supplemental C) using ML methods as described above.

Multi-Species Coalescent Analyses—MSC analyses were performed on locus alignments generated by stages two and three of the SORTER pipeline (Table 1, datasets C, E1–E10; Fig. 2; Supplemental C) using ASTRAL 5.7.7 (Zhang et al. 2018). The software reconstructs the species tree that maximizes the number of quartet topologies in a set of gene trees. Branch lengths output by ASTRAL reflect coalescent units, and local posterior probability support values (LPP) reflect the proportion of quartet topologies in the gene trees that include a given branch (Mirarab 2019).

Gene and Site Concordance Factors—For ML and MSC phylogenies, we included two additional metrics of node support, gene concordance factors (gCFs) and site concordance factors (sCFs) using IQ-TREE 2 (Minh et al. 2020). These metrics represent the percentage of individual locus-specific or 'gene tree' topologies and informative nucleotide sites supporting a particular node, respectively. Concordance factors complement conventional measures of branch support by providing a description of discordance in the underlying data (Minh et al. 2020). We used the unphased and phased ML gene-trees and alignments for individual loci to calculate gCF and sCF values for the phylogenies derived from concatenated datasets A and B (Table 1), respectively. Phased ML gene-trees and alignments from dataset B were used to compute gCF and sCF values for the phylogeny derived from MSC analysis of dataset C, and ML trees and alignments generated by SORTER stage 3 were used to compute gCF and sCF values for each respective allopolyploid dataset (Table 1, datasets D1–E10).

Results

Read Statistics—An average of 518,215 paired-end reads were generated for each specimen (Table S1, Mendez-Reneau et al. 2022). After removing adaptor sequences and trimming low quality reads, an average of 515,385 reads were retained for each specimen. On average 44% of trimmed reads mapped to one of 407 targeted GoFlag loci, with only one of the 408 total targeted loci not retrieving any sequence data. There was a higher average of mapped reads for Polypodium samples (45.83%) than outgroup samples (34.99%). Read coverage of targeted GoFlag loci after clustering averaged 99.76% across all samples. The average read-depth for all samples and all locus-clusters was 147.95, with Polypodium samples having a higher average read depth (154.87) compared to outgroup samples (104.32).

De Novo Contig Assembly of Consensus Alleles—Samples of diploid species belonging to the P. vulgare complex, Mesoamerican species, and outgroup species generated an average of 533.76 consensus alleles mapping to 407 out of 408 reference

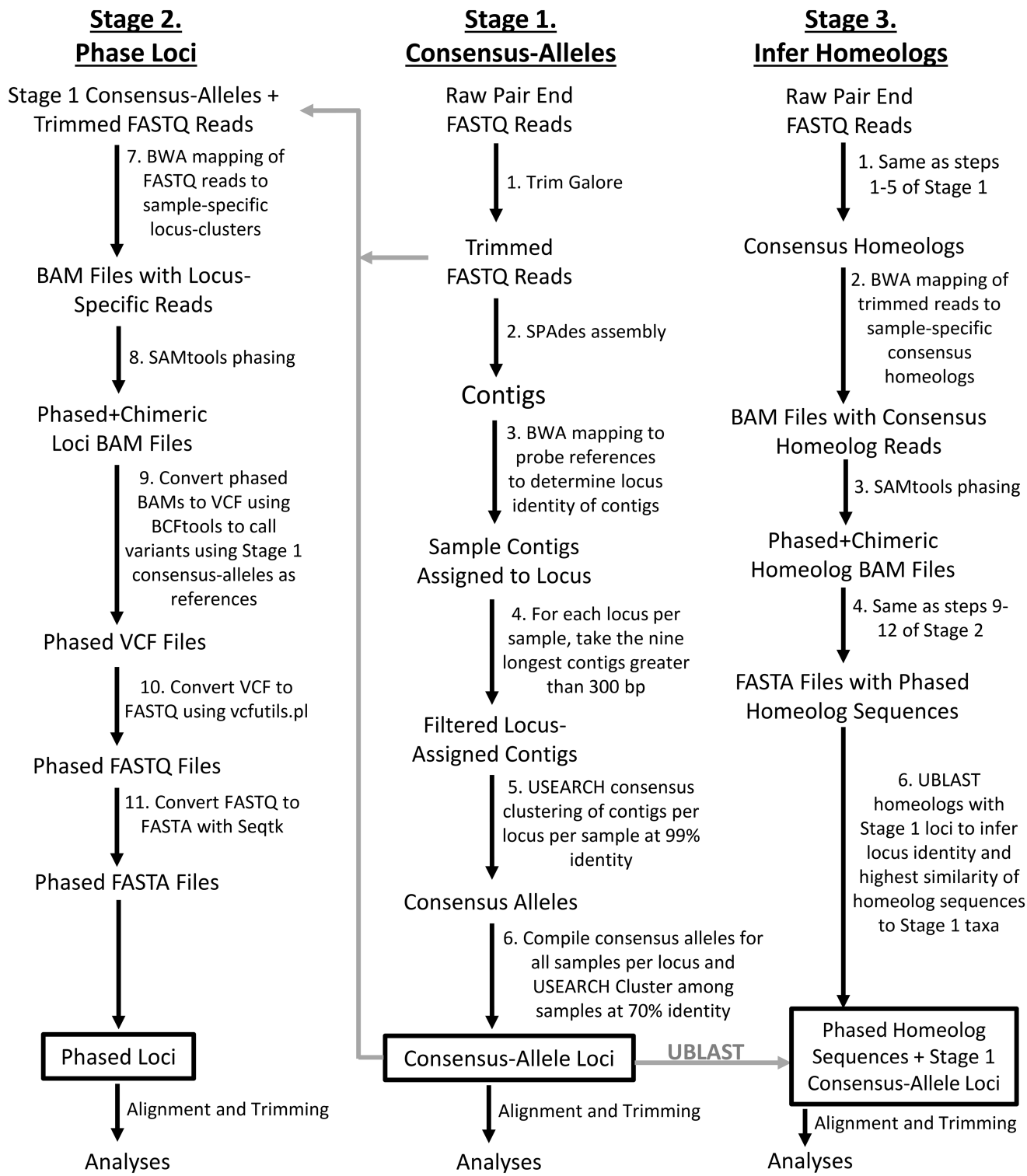

FIG. 2. The three stages of the SORTER pipeline with the contig and clustering settings used in this study. Black arrows indicate data processing steps, with each black arrow pointing to the output for that step. Boxes highlight the final output for each of the three stages of the pipeline. Gray arrows represent the interactions between outputs among different stages of the pipeline. For example, the first step of stage 2 (step 7) utilizes trimmed reads and locus clus-ter sequences generated in stage 1.

loci (min 5 461; max 5 987). After clustering among all diploid samples using a 70% identity clustering threshold in stage one of the SORTER pipeline to generate sets of orthologous sequences, about 3% of consensus alleles were represented by two or more sequences per locus for all samples (Supplemental B; Fig. S1; Mendez-Reneau et al. 2022). Allopolyploid samples generated an average of 842.9 consensus-homeologs that mapped to 407 out of 408 reference loci (min 5 731; max 5 1128).

Datasets and Alignments—The unphased concatenated alignment generated by stage one of the SORTER pipeline (Table 1, dataset A) included 388 loci with $ 50% sample

TABLE 1. Summary of the 23 datasets assembled and analyzed as part of this study. Columns include dataset identifier, dataset descriptor, stage of the SORTER pipeline that generated each dataset, number of loci, total alignment length, percent missing data, percent variable sites, and percent informative sites. The right-most column indicates the phylogenetic analysis performed on each dataset: RaxML maximum likelihood (ML), or ASTRAL Multispecies Coalescent (MSC). The log likelihood value for the 'best' ML tree for each ML analysis is presented in parentheses.

Dataset ID	Dataset Descriptor	SORTER Stage Output	Number of Loci	Total Alignment Length	Percent Missing Data	Percent Variable Sites	Percent Parsimony Informative Sites	ML
A	Diploid Unphased Concatenated	Stage 1	388	476834	39.11%	52.23%	28.91%	Y (ln 5 23368641.92)
B	Diploids Phased Concatenated	Stage 2	352	459702	41.32%	55.19%	47.22%	Y (ln 5 23824649.30)
C	Diploids MSC	Stage 2	352	459702	41.32%	55.19%	47.22%	N
D1	Diploids 1 P. vulgare 1 Concatenated	Stage 3	235	329574	38.60%	57.09%	48.68%	Y (ln 5 22895337.42)
D2	Diploids 1 P. vulgare 2 Concatenated	Stage 3	132	188909	37.80%	57.74%	49.27%	Y (ln 5 21696281.31)
D3	Diploid 1 P. hesperium 1 Concatenated	Stage 3	211	298696	38.44%	57.14%	48.59%	Y (ln 5 22604283.27)
D4	Diploids 1 P. hesperium 2 Concatenated	Stage 3	176	247964	38.24%	57.21%	48.64%	Y (ln 5 22170374.88)
D5	Diploids 1 P. saximontanum 1 Concatenated	Stage 3	164	229102	37.54%	56.73%	48.26%	Y (ln 5 22020407.89)
D6	Diploids 1 P. saximontanum 2 Concatenated	Stage 3	164	230018	38.18%	57.22%	48.66%	Y (ln 5 22034200.33)
D7	Diploids 1 P. virginianum 1 Concatenated	Stage 3	33	43633	44.09%	61.60%	51.70%	Y (ln 5 2446123.97)
D8	Diploids 1 P. virginianum 2 Concatenated	Stage 3	75	108789	39.81%	59.12%	49.63%	Y (ln 5 21020045.32)
D9	Diploids 1 P. calirhiza 1 Concatenated	Stage 3	95	135545	40.25%	57.63%	48.67%	Y (ln 5 21229178.48)
D10	Diploids 1 P. calirhiza 2 Concatenated	Stage 3	92	134545	38.13%	57.65%	48.76%	Y (ln 5 21229046.22)
E1	Diploids 1 P. vulgare 1 MSC	Stage 3	235	329574	38.60%	57.09%	48.68%	N
E2	Diploids 1 P. vulgare 2 MSC	Stage 3	132	188909	37.80%	57.74%	49.27%	N
E3	Diploid 1 P. hesperium 1 MSC	Stage 3	211	298696	38.44%	57.14%	48.59%	N
E4	Diploids 1 P. hesperium 2 MSC	Stage 3	176	247964	38.24%	57.21%	48.64%	N
E5	Diploids 1 P. saximontanum 1 MSC	Stage 3	164	229102	37.54%	56.73%	48.26%	N
E6	Diploids 1 P. saximontanum 2 MSC	Stage 3	164	230018	38.18%	57.22%	48.66%	N
E7	Diploids 1 P. virginianum 1 MSC	Stage 3	33	43633	44.09%	61.60%	51.70%	N
E8	Diploids 1 P. virginianum 2 MSC	Stage 3	75	108789	39.81%	59.12%	49.63%	N
E9	Diploids 1 P. calirhiza 1 MSC	Stage 3	95	135545	40.25%	57.63%	48.67%	N
E10	Diploids 1 P. calirhiza 2 MSC	Stage 3	92	134545	38.13%	57.65%	48.76%	N

representation for a total alignment length of 476,834 bp (see Table 1 for alignment statistics). Both phased alignments generated by stage two of the SORTER pipeline (Table 1, datasets B, C) comprised 352 loci with $ 50% sample representation for a total alignment length of 459,702 bp. The phased alignments (datasets B, C) had a higher percentage of variable sites and informative sites (55.19% and 47.22%, respectively) compared to the concatenated unphased alignment (dataset A; 52.23% and 28.91%, respectively). The alignments generated from SORTER stage 3 that incorporate putative homeolog sequences from allopolyploid taxa (Table 1, datasets D1–E10) contained between 33–235 loci, an average alignment length of 194,298 bp, an average of 57.91% variable sites, and an average of 49.09% informative sites.

Relationships Among Major Clades of Polypodium s.s. and Outgroup Taxa—MAXIMUM LIKELIHOOD ANALYSES—ML analyses of datasets A and B (those with unphased and phased sequences, respectively, of diploid temperate and Mesoamerican Polypodium species (Table 1; Appendix 1)) yielded nearly identical topologies with minor differences in interspecific relationships (Fig. 3). All phylogenies provide complete support for the monophyly of Polypodium s.s. (MLBS 5 100%; gCF and sCF values are reported for key nodes on Fig. 3), comprising two completely supported sister clades: the temperate Polypodium vulgare complex and a clade of Mesoamerican species allied with P. plesiosorum (M clade). Within the temperate P. vulgare complex, all topologies provide unequivocal support for four major clades of diploid species: the A clade (P. amor-phum (P. sibiricum, P. appalachianum)), the G clade (P. fauriei (P. glycyrrhiza, P. californicum)), the S clade (P. scouleri, P. pellucidum), and the C clade (P. cambricum, P. macaronesicum). All species belonging to the P. vulgare complex and represented by more than one sample (e.g. P. glycyrrhiza) formed completely supported (MLBS 5 100%) monophyletic groups. Relationships among the major clades within the P. vulgare complex were consistently recovered as (G (C (A, S))), with moderate support for the sister pairing of the A and S clades (unphased: MLBS 5 96%; phased: MLBS 5 82%)

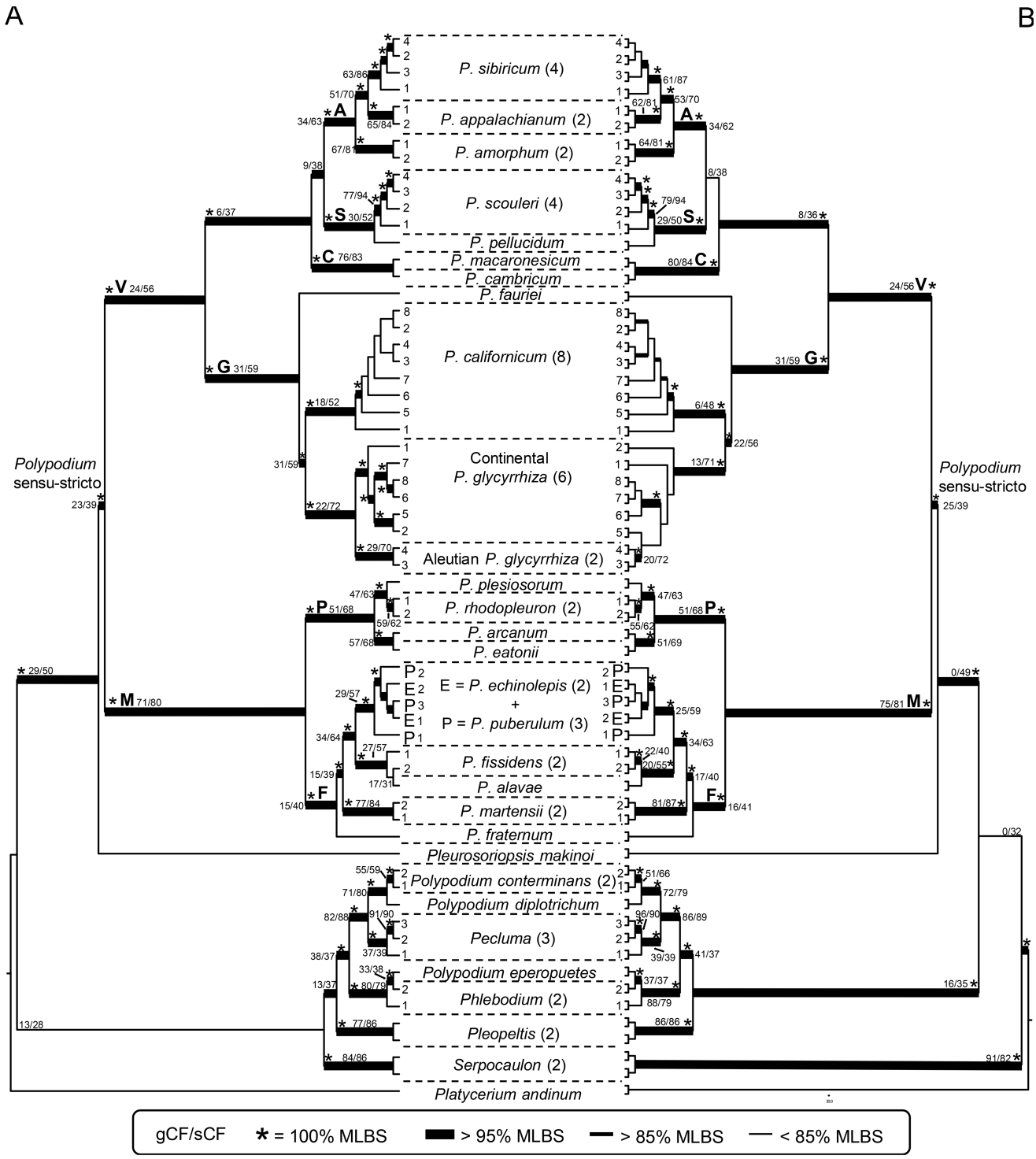

FIG. 3. 'Best' maximum likelihood trees. A. Reconstructed from concatenated unphased dataset A (Table 1). B. Reconstructed from concatenated phased dataset B (Table 1). Branch widths indicate maximum likelihood bootstrap (MLBS) node support values, with asterisks representing 100% MLBS support. Gene concordance factor (gCF) and site concordance factor (sCF) values are presented above most nodes. Bolded letters indicate major clades within Polypodium s.s.: V 5 Polypodium vulgare complex; M 5 Mesoamerican clade; A 5 appalachianum clade; G 5 glycyrrhiza clade; C 5 cambricum clade; S 5 scouleri clade; P 5 plesiosorum clade; F 5 "free-veined" clade. Numbers in parentheses after taxa names represent the number of individuals sampled. For Polypodium species with more than two samples, unique sample identifier numbers are given at tips (Appendix 1). Polyphyletic Polypodium echinolepis and Polypodium puberulum samples are designated with "E" and "P" at tips, respectively.

Interspecific relationships within the Mesoamerican M clade were consistent and well-supported across all phylogenies (average unphased: MLBS 5 95%; average phased: MLBS 5 100%; Fig. 3). The M clade is split into sister groups comprising ((P. eatonii, P. arcanum) (P. plesiosorum, P. rhodopleuron)) and (P. fraternum (P. martensii ((P. alavae, P. fissidens) (P. puberulum, P. echinolepis))), respectively referred to as the plesiosorum P clade and the free-veined F clade. The two samples of P. fissidens were united in a completely supported clade in all topologies (MLBS 5 100%) but with P. alavae nested

within P. fissidens in the unphased topology (Table 1, dataset A; Fig. 3) Alternatively, the phased topology (Table 1, dataset B; Fig. 3) supported P. alavae as sister to P. fissidens (MLBS = 100%). Three P. puberulum and two P. echinolepis samples were not recovered as monophyletic, but their samples were united in an entirely supported clade (MLBS = 100%).

In general, outgroup genera formed well-supported clades (average unphased: MLBS = 100%; phased: MLBS = 100%; Fig. 3), and both datasets yielded nearly identical topologies, except for the placement of Serpocaulon (Table 1, datasets A–B; Fig. 3). ML analysis of unphased dataset A placed Serpocaulon in a moderately supported clade (MLBS = 84%) with several genera including Pecluma, whereas ML analysis of phased dataset B provided complete support (MLBS = 100%) for the placement of Serpocaulon as sister to all taxa except Platycerium. Pleurosoriopsis makinoi was consistently and completely supported as sister to Polypodium s.s. Notably, three previously unsequenced Polypodium species, P. conterminans, P. diplotrichum, and P. eperopeutes, are resolved as closely allied to the outgroup genera Pecluma and Phlebodium.

Multiple Species Coalescence Analyses—The species tree generated from the MSC analysis of dataset C (Table 1; Fig. 4) was very similar to those generated by ML analyses of concatenated datasets A and B (Table 1; Fig. 3), with complete support for the monophyly of Polypodium s.s. and for the sister pairing of the Polypodium vulgare complex and the M clade (LPP = 1.0; gCF and sCF values are reported for key nodes on Fig. 4). Relationships among the A, C, G, and S clades were the same as in the ML phylogenies (Fig. 3) but with low support for the sister relationship of the S and A clades (LPP = 0.76; Fig. 4). Inter-species relationships within each of the four major clades of the P. vulgare complex (the A, C, G, and S clades) and the Mesoamerican M clade were well-supported (average LPP = 1.0) and identical to the ML phylogenies (Fig. 3). Outgroups relationships were similar to those in the ML topology derived from dataset A (Table 1; Fig. 3) with complete support (LPP = 1.0) for the monophyly of most outgroup genera (Fig. 4).

Intraspecific Relationships Within Polypodium s.s.—

Maximum Likelihood Analyses—Topology and node support for intraspecific relationships for most species were consistent between trees generated from concatenated datasets A and B (Table 1; Fig. 3), with a few exceptions. For example, datasets A and B (Table 1) yielded different intraspecific relationships and support for the eight included specimens of P. glycyrrhiza. In the topology derived from unphased dataset A, the P. glycyrrhiza specimens from the Aleutian Islands are well supported as sister to the other P. glycyrrhiza samples, whereas in dataset B there is little support for this relationship.

Multiple Species Coalescence Analyses—For many species intraspecific relationships and support varied substantially among the MSC phylogeny (Table 1, datasets A, C; Fig. 4) and the concatenated ML phylogenies (Table 1, datasets A, B; Fig. 3). For example, for P. glycyrrhiza we consistently recovered a strongly supported Aleutian Island clade nested among the Continental North American specimens but with variable support.

Phylogenetic Placement of Allopolyploid Homeologs—For each of the five included allotetraploid species, topologies obtained from ML analyses of the concatenated datasets (e.g. Table 1, datasets D1–D10) were generally congruent with the topologies obtained from MSC analyses (e.g. datasets E1–E10). All datasets and analyses yielded support for the monophyly of the A, G, C, and S clades of the P. vulgare complex and the monophyly of individual diploid species, allowing for the reliable identification of allopolyploid progenitors by assessing the placement of allopolyploid homeolog sequences relative to sequences from diploid taxa. Our results support each of the hypotheses about the diploid progenitors of the five allopolyploid species illustrated in Fig. 1. We recovered strong support for P. glycyrrhiza and P. sibiricum as the progenitors of P. vulgare (Fig. 5A, average MLBS = 100%, average LPP = 0.99), P. glycyrrhiza and P. amorphum as the progenitors of P. hesperium (Fig. 5B, average MLBS = 100%, average LPP = 1.0), and for P. amorphum and P. sibiricum as the progenitors of P. saximontanum (Fig. 5C, average MLBS = 100%, average LPP = 1.0). Finally, P. sibiricum and P. appalachianum are strongly supported at the progenitors of P. virginanum (Fig. 5D, average MLBS = 80.5%, average LPP = 1.0), and P. glycyrrhiza and P. californicum are strongly supported at the progenitors of P. calirhiza (Fig. 5E, average MLBS = 71.5%, average LPP = 0.96).

Discussion

Ferns have long been recognized as a plant lineage with frequent genome duplications and hybridization events (Otto and Whitton 2000; Wood et al. 2009; Huang et al. 2020), but a lack of low copy nuclear sequence loci and other genomic resources has impeded phylogenetic reconstruction within many fern groups (Sessa et al. 2014). Here we demonstrate the power of TC methods and new bioinformatic tools, specifically the GoFlag 408 TC probes and the SORTER bioinformatic pipeline, for inferring evolutionary relationships in rapidly diverging fern lineages with polyploidy and reticulation complexes. Focusing on Polypodium s.s., we recovered not only topologies largely consistent with previous hypotheses about species relationships within the major clades of the P. vulgare complex (Haufler et al. 1995a, 1995b; Sigel et al. 2014a), but also novel relationships among these major clades and new insights into the relationships among Mesoamerican species of Polypodium (Figs. 3, 4). In addition, we recovered strong support for previous hypotheses about the parentage of five allopolyploid species belonging to the P. vulgare complex (Figs. 1, 5; Haufler et al. 1995b). Below we discuss implications of our phylogenomic analyses for our understanding of evolutionary relationships within the Polypodium s.s. and the broader Polypodiaceae, as well as highlight the performance of the SORTER pipeline for processing of TC data.

Increased Resolution for Polypodium s.s. and Polypodiaceae—The results of this study are largely consistent with the findings of Sigel et al. (2014a); they both provide strong support for the monophyly of the Polypodium s.s. and its placement as sister to Pleurosoriopsis makinoi, the monotypic Asian taxon with twice-pinnate leaves. Within Polypodium s.s. we recovered two large clades, the temperate Polypodium vulgare "V" reticulation complex and the Mesoamerican M clade (Figs. 3, 4; Haufler et al. 1995a, 1995b; Schneider et al. 2004; Otto et al. 2009; Sigel et al. 2014a). As in Sigel et al. (2014a), we found complete support for four major clades within the P. vulgare complex (the G, C, S, and A clades) each comprising two or more diploid species (Figs. 1, 3, 4). However, Sigel et al. (2014a) failed to recover strong support for

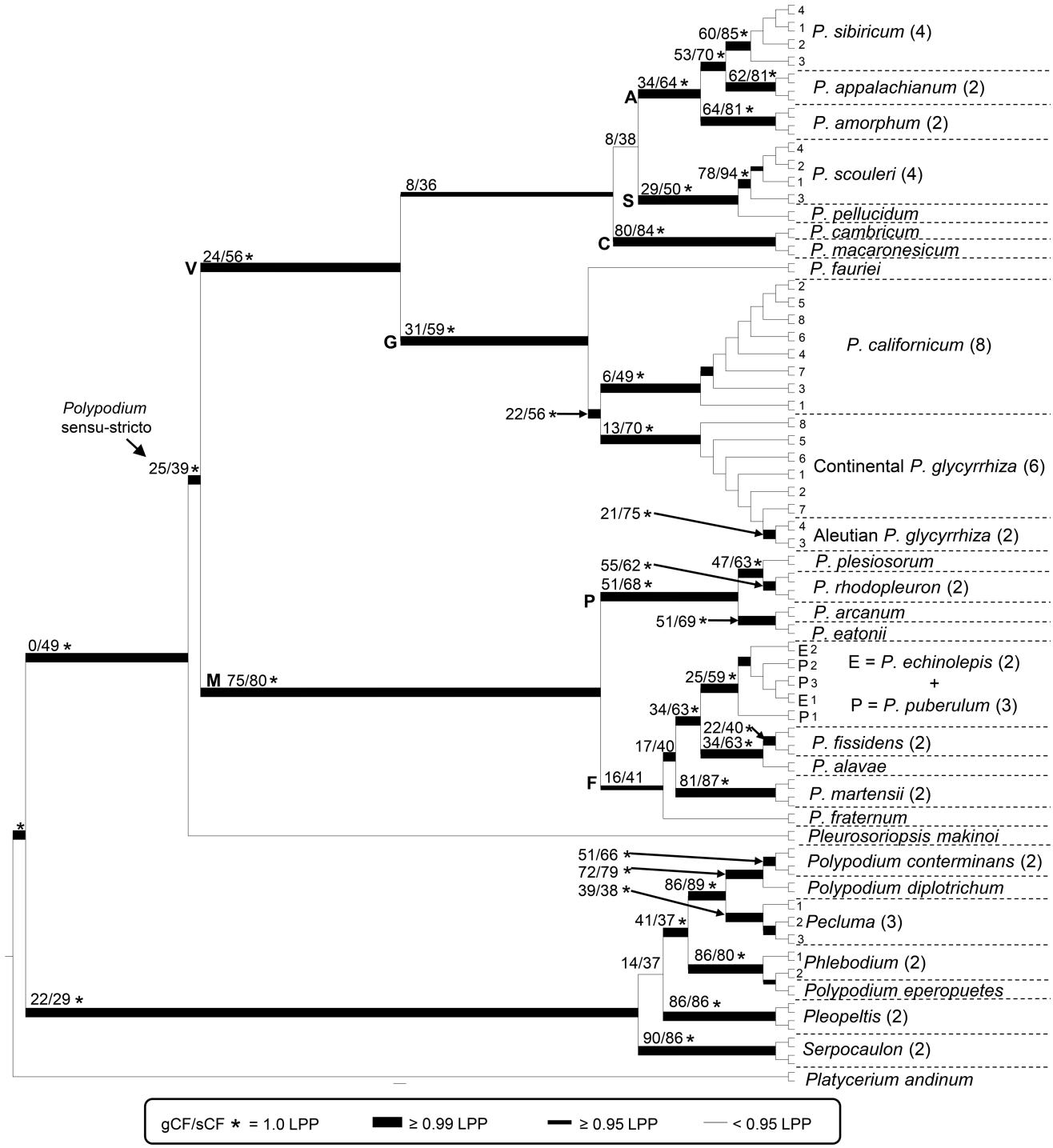

FIG. 4. ASTRAL multispecies coalescent (MSC) tree based on phased dataset C (Table 1). Branch widths indicate local posterior probability (LPP) node support values, with asterisks representing 100% LPP support. Gene concordance factor (gCF) and site concordance factor (sCF) values are presented for key nodes. Bolded letters V, A, G, C, S, M, P, and F indicate the major clades within Polypodium s.s. as depicted in Fig. 3. Numbers in parentheses after taxa names represent the number of individuals sampled, unique sample identifier numbers are given at tips for Polypodium species with more than two samples (Appendix 1).

relationships among the four major clades, as well as resolve relationships among diploid species of the A clade. Below we focus our discussion primarily on the novel relationships among the four major clades of the P. vulgare complex, interspecific relationships within the A clade, and previously unexplored relationships within the Mesoamerican M clade.

Relationships Among and Within the Major Clades of the P. vulgare Complex—Our MSC phylogeny provides novel support for the placement of the G clade as the first diverging lineage within P. vulgare complex, and for the monophyletic grouping of the C, S, and A clades. These results are surprising considering that the chloroplast phylogeny (Sigel et al. 2014a) recovered the G clade sister to the A clade

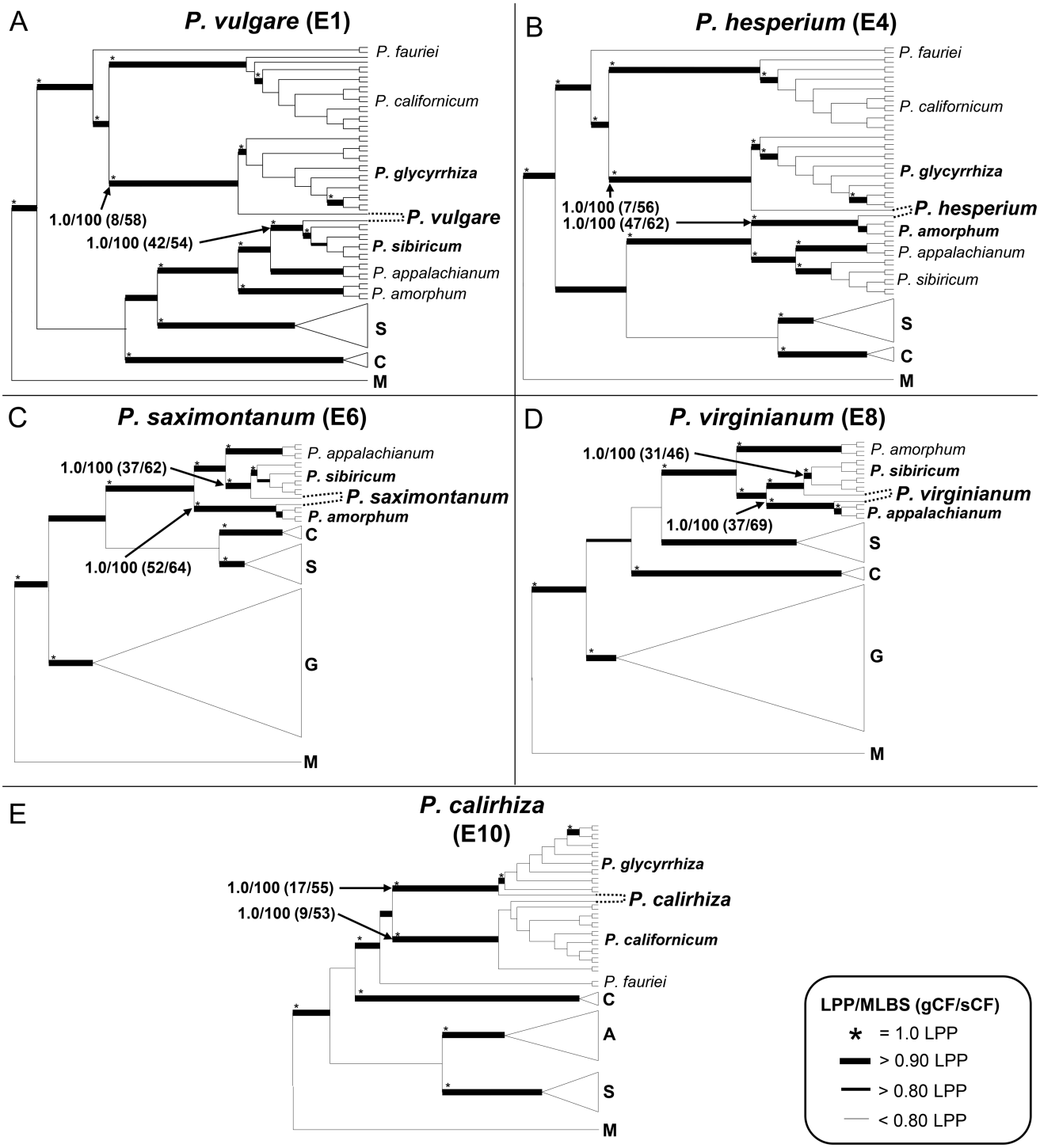

FIG. 5. ASTRAL multispecies coalescent (MSC) trees from analyses of five datasets, each including putative homeolog sequences of one allotetraploid sample. A. Dataset E2, Polypodium vulgare. B. Dataset E3, Polypodium hesperium. C. Dataset E6, Polypodium saximontanum. D. Dataset E7, Polypodium virginianum. E. Dataset E9, Polypodium calirhiza. See Table 1 for dataset descriptions and Supplemental C (Mendez-Reneau et al. 2022) for detailed phylogenies. Dot-ted brackets link tips representing homeologs of each allopolyploid sample. Bold text indicates each allotetraploid taxon and its putative diploid progenitors. Clades not related to each allotetraploid are collapsed for simplicity. Line widths indicate local posterior probability (LPP) support values, with asterisks representing 1.0 LPP. For comparison, specific LPP support values, as well as maximum likelihood bootstrap (MLBS) support values from analyses of corresponding concatenated datasets (Table 1, datasets D2, D3, D6, D7, and D9), are reported for key nodes. Concordance factor values (gCF/sCF) for the MSC tree are reported in parentheses.

and the C clade sister to the S clade, albeit with low support. The novel placement of the G clade suggests that the P. vulgare complex could have originated in Asia. The Japanese taxon P. fauriei is the earliest diverging taxon in the G clade, with P. glycyrrhiza and P. californicum successively extending along the Pacific Rim of Asia and North America from the Kamchatka Peninsula to the Baja California (Haufler et al. 1993). One hypothesis to explain their current distribution is that P. glycyrrhiza may have diverged from P. fauriei via allopatry on the Kamchatka Peninsula or Aleutian Islands, and

P. glycyrrhiza subsequently colonized the west coast of North America. Then, P. glycyrrhiza and P. californicum diverged parapatrically in central California. An Asian origin for the P. vulgare complex is also supported by the placement of the Japanese endemic taxon Pleurosoriopsis makinoi as the sister to Polypodium s.s.

Our MSC phylogeny strongly supports the monophyly of the A clade as comprising the three diploid species P. appalachianum, P. amorphum, and P. sibiricum. The affinity of these three species has long been hypothesized based on their shared presence of the sporangiasters, apomorphic sterile structures within sori that are homologous to sporangia (Martens 1943; Lang 1965, 1969, 1971; Peterson and Kott 1974; Siplivinski 1974), and supported by previous phylogenetic analyses of isozymes and plastid sequences (Haufler et al. 1995b; Sigel et al. 2014a). Novel to this study, we recover unequivocal support for the monophyly of each of the three diploid species (Fig. 4), each of which is characterized by a suite of distinct, although somewhat cryptic, morphological characters, spore size, and geographic ranges (Haufler and Windham 1991; Haufler et al. 1993, 1995a, 1995b). Polypodium appalachianum ranges from the southern Appalachian Mountains to eastern Canada, has acute pinnae apices, abundantly glandular sporangiasters, and spores , 52 mm long. Polypodium sibiricum is arctic/boreal in its distribution with rounded pinnae apices, eglandular sporangiasters, and spores , 52 mm long. Polypodium amorphum is differentiated by having spores . 52 mm long and glandular sporangiasters. It is restricted to the Cascade Mountains in the western United States and British Columbia.

An additional novel finding of this study is strong support for the sister pairing of P. appalachianum and P. sibiricum (Figs. 3, 4). Haufler et al. (2000) proposed that the simple, eglandular sporangiasters of P. sibiricum evolved glandular forms in the species P. appalachianum and P. amorphum. However, according to our inferred topology it is more parsimonious that glandular sporangiasters evolved in the common ancestor of the A clade with a secondary loss of glandularity in the ancestor of P. sibiricum, or that glandularity evolved separately in the ancestors of P. appalachianum and P. amorphum.

While our discussion of intraspecific relationships is largely restricted to the A clade, we would be remiss not to point out an intriguing result pertaining to P. glycyrrhiza, a diploid taxon of the G clade (see Sigel et al. 2014a for a detailed discussion of phylogenetic relationships, biogeography, and previous literature pertaining to the G, C, and S clades that is congruent with the results of this study). With a geographic distribution spanning the Pacific Rim from central California to the Kamchatka Peninsula, P. glycyrrhiza has a complex taxonomic history reflecting subtle morphological variation across its range (Lang 1965, 1971; Haufler and Windham 1991; Shalimov and Shmakov 2016). Shalimov and Shmakov (2016) proposed segregation of specimens from the Aleutian Islands and Sitka, Alaska as the species Polypodium aleuticum A.E. Bobrov based on vegetative and spore morphology. Notably, our sampling of P. glycyrrhiza includes two specimens from separate Aleutian Islands that are united in an entirely supported clade (Fig. 4) and appear as reciprocally monophyletic with other specimens of P. glycyrrhiza in the unphased concatenated ML phylogeny (Fig. 3). Despite a small sample size, these results may lend credence to the segregation of P. aleuticum from P. glycyrrhiza, perhaps representing an additional diploid taxon in North America and warranting further study with increased sampling.

The Mesoamerican M clade—All of our phylogenetic analyses recovered complete support for a Mesoamerican clade (M clade) of taxa distributed across Mexico and central America, a group previously hypothesized to be closely allied with Polypodium plesiosorum (Tejero-Díez and Pacheco 2004; Tejero-Díez 2005; Tejero-Díez et al. 2010; Otto et al. 2009; Luna-Vega et al. 2012; Sigel et al. 2014a; Figs. 3, 4). This study provides strong support for two major subclades, which we refer to as the P. plesiosorum P clade and free-veined F clade. The P clade comprises four taxa, P. plesiosorum, P. rhodopleuron, P. eatonii, and P. arcanum, unified by the presence of areoles and netted venation, with most species having glabrous leaves (Tejero-Díez et al. 2010). As with previous phylogenies (Otto et al. 2009; Sigel et al. 2014a), we recovered P. rhodopleuron sister to P. plesiosorum. These two species share very similar morphologies, being differentiated by gla-brous leaves in P. rhodopleuron and hairs on the abaxial leaf surfaces in P. plesiosorum (Tejero- Díez et al. 2010; Luna-Vega et al. 2012). Additionally, we recovered novel support for the monophyly of P. arcanum and P. eatonii, which were previously hypothesized to be allied based on unique vein anastomosis patterns and the fusion of xylem strands along the petiole-leaf rachis (Mickel and Smith 2004; Tejero-Díez et al. 2010).

The F clade comprises six taxa, P. fraternum, P. fissidens, P. puberulum, P. echinolepis, P. alavae, and P. martensii, most of which have hairs on their stipes and leaves, as well as free venation (Mickel and Smith 2004). Within the F clade we recovered unequivocal support for a clade comprising intermixed specimens of P. puberulum and P. echinolepis (Figs. 3, 4; Supplemental C). Mickel and Smith (2004) noted substantial morphological similarity between these two species that are sympatric in wet montane forests in southern Mexico and central America. However, they are differentiated by pinna fusion; Polypodium puberulum has pinnae adnate to the rachis, whereas pinnae are free in P. echinolepis. The paraphyly of these two species suggests that they may hybridize or represent phenotypically plastic variants of the same species. Increased sampling and careful study of morphological variation between these taxa may clarify their relationships.

Outgroup Relationships—Our MSC phylogeny is generally well aligned with previously recovered relationships among the outgroup genera (Smith et al. 2006; Otto et al. 2009; Sigel et al. 2014a). Interestingly, as in several previous studies (Otto et al. 2009; Assis et al. 2016; Almeida et al. 2017), we recovered the placement of several Polypodium species from central and south America among outgroup genera. We recovered P. conterminans and P. diplotrichum united in a clade with Pecluma despite previous taxonomic treatments suggesting they were closely related to the M clade taxon P. plesiosorum (Tejero-Díez and Pacheco 2004; Tejero-Díez 2005). Interestingly, Mickel and Smith (2004) noted that P. diplotrichum is morphologically similar to Pecluma chiapensis (formerly Polypodium chiapense; Carvajal-Hernandez 2018), both lacking "dark axes and clavate soral paraphyses" and suggested that P. diplotrichum should be reassigned to Pecluma. Similarly, we recovered P. eperopuetes as nested within Phlebodium despite previous observations that it is morphologically similar to the M clade taxon P. echinolepis (Mickel and Beitel 1988). Clearly, Polypodium sensu lato, encompassing approximately 100–125 species

worldwide (Mickel and Smith 2004), would benefit from broader molecular sampling and generic revision.

Confirming the Progenitors of Allotetraploid Taxa within the Polypodium vulgare Complex—Delimiting allopolyploids in the Polypodium vulgare complex and their diploid progenitors has been a multi-generational endeavor, beginning with the observation of multiple cytotypes within P. vulgare by Irene Manton (1947). Hybridization experiments, observations of chromosome pairing, and isozyme electrophoresis uncovered numerous other allopolyploid species closely allied to P. vulgare and were used to develop and test hypotheses about their diploid progenitor species. With 408 GoFlag TC probes and the SORTER pipeline to reconstruct homeologs, we have successfully used multi-locus nuclear sequence data to confirm previous hypotheses about the diploid progenitors of five allotetraploid species, P. vulgare, P. hesperium, P. calirhiza, P. saximontanum, and P. virginianum (Figs. 1, 5).

Notably, these five allopolyploid taxa vary in the relatedness of their progenitors. Polypodium vulgare and P. hesperium each have one progenitor from the A clade and one from the G clade, whereas P. calirhiza has both progenitors from the G clade. Polypodium saximontanum and P. virginianum have both progenitors from the A clade (Figs. 1, 5). Previous estimates by Sigel et al. (2014a) suggest that the A and G clades diverged in the mid-Miocene approximately 13.4 MYA, but the progenitors within the G and A clades diverged from each other in the late Pliocene to early Pleistocene approximately 2.7 MYA and 2.3 MYA, respectively. For the P. vulgare complex, there appears to be a direct relationship between the relatedness of diploid progenitors of an allopolyploid taxon and the number of loci for which two or more distinct homeologs can be reconstructed (Table 1). For example, we were able to recover an average of 183.5 loci for P. vulgare (derived from A clade P. sibiricum and G clade P. glycyrrhiza) but only an average of 54 loci for P. virginianum (derived from the A clade taxa P. amorphum and P. appalachianum). Nonetheless, all datasets produced ample loci to resolve allopolyploid-progenitor relationships within the P. vulgare complex (Table 1; Fig. 5), suggesting that TC markers and the SORTER pipeline have broad utility for resolving reticulation complexes of varying age within the ferns and other lineages.

The Value of TC Data for Resolving the Evolutionary History of Polypodium s.s.—The TC approach enables us, for the first time, to examine the evolutionary history of Polypodium s.s. based on hundreds of presumably independently evolving loci and assess levels of gene tree discordance, or phylogenetic incongruence, among loci. For this study, the ML trees of individual loci had major topological differences within Polypodium s.s. and across the broader Polypodiaceae (Supplemental C). This was reflected in low gCF and sCF values (i.e. , 50%) for major relationships within the P. vulgare complex for both the phased and unphased concatenated ML phylogenies, even when strongly supported by MLBS (Fig. 3). For example, for both the unphased and phased concatenated ML phylogenies, MLBS support for the Polypodium vulgare complex was 100%, but the gene and site concordance factors were 24% and 59%, respectively (Fig. 3). Another notable, clade-level relationship with low concordance factors and moderate to high MLBS support is the sister-pairing of the S and A clades (Fig. 3; unphased concatenated dataset A: MLBS 5 96%, gCF 5 9%; sCF 5 38%; phased concatenated dataset B: MLBS 5 82%, gCF 5 9%, sCF 5 38%), but the MSC analysis recovered marginal support for the sister relationship of the S and A clades (Fig. 4; LPP 5 0.76). Interestingly, we recovered the sister pairing of clades S and C for a subset of our analyses incorporating allopolyploid sequences (Table 1, e.g. datasets D4, D6; Supplemental C) suggesting that the relationships recovered among the four major clades within the P. vulgare complex are highly sensitive to the loci used for analysis. Hence, support for the monophyly of A and S clades reflects a portion of the data, with other loci supporting the monophyly of S and C clades. Taken together, we believe that the divergences of the S, A, and C clades likely occurred rapidly resulting in incomplete lineage sorting (ILS) or ancestral introgression (Degnan and Rosenberg 2009; Liu et al. 2009, 2015; Nakhleh 2013). More broadly, prevalent gene tree discordance and low concordance factors suggest extensive ILS and/or introgression throughout the history of Polypodiaceae s.s., and justifies our adoption of the MSC phylogeny for discerning the evolution of Polypodium s.s. and the broader outgroup sampling of Polypodiaceae (Fig. 4).

The high coverage of the TC data also enables phasing approaches to separate individual allelic or homeologous copies of each locus. Although the value of phasing may be most apparent in inferring the progenitors of the allopolyploids, phasing also increased the number of variable and informative sites in the alignments (Table 1, datasets A for diploids and D1–D10, E1–E10 for allopolyploids) as in Andermann et al. (2019). Phasing also changed the ML bootstrap support for some of the clades within Polypodium s.l. (Fig. 3). However, similar to Kates et al. (2018), phasing had little effect on the phylogenetic relationships. For example, phasing improved support for the monophyly of M clade taxon P. fissidens but reduced support for the sister relationship of the S and A clades within the P. vulgare complex. We interpret the latter example to be caused by conflicting phylogenetic signals for the placement of the S and A clades due to the additional informative variation recovered by phasing.

The SORTER Pipeline is a Flexible Tool for Processing TC Data—The proliferation of TC datasets has led to the development of several bioinformatic pipelines for assembling orthologous sequences for phylogenomic applications, each one adopting a different approach for segregating and capturing paralogous sequences (Faircloth et al. 2020; Johnson et al. 2016; Rothfels et al. 2017; Freyman et al. 2022), allelic variants (Aguiar and Istrail 2012; Andermann et al. 2019; Tiley et al. 2021), or homeologs (Rothfels et al. 2017; Nauheimer et al. 2021; Freyman et al. 2022). The SORTER pipeline is a unified, flexible tool for generating multi-locus nuclear datasets for use in traditional phylogenetic and multi-species coalescent analyses. The three customizable, modular stages of the SORTER pipeline successively cluster putatively paralogous gene copies across samples, phase alleles from diploid taxa, and phase homeologs based on diploid samples for allopolyploid taxa, making it ideal for the phylogenomics of reticulation complexes.

SORTER has numerous advantages. First, SORTER retains sequence variants generated by contig assembly (given user defined filters) through sequential clustering and phasing, resulting in putative orthologous sequences, alleles, and homeologs. This contrasts with other pipelines that initially select a single longest representative contig for each locus for each sample and remove alternative, shorter contigs or collapse slightly overlapping contigs (Faircloth 2016; Johnson et al. 2016). Additionally, SORTER infers and retains putatively paralogous locus copies by clustering sequences by percent identity, effectively generating multiple loci from a

single reference that can be included in the final dataset. Other pipelines detect paralogs post-hoc by identifying loci with disproportionately high levels of SNPs (Nauheimer et al. 2021) or through successive rounds of computationally expensive phylogenetic analysis (Rothfels et al. 2017). By adjusting the percent identity clustering threshold for their dataset in SORTER stage 1, the user can balance the number of loci recovered for phylogenetic analysis with the risk of erroneously clustering paralogous gene copies at lower clustering identities (Supplemental A). Another advantage of SORTER is that it constructs haplotypes of allopolyploid taxa without making any a priori assumptions about ploidy level, the number of homeologous sequences, or progenitor identity (Kamneva et al. 2017), and it does not require reads to overlap across targeted loci (Aguiar and Istrail 2012; Kamneva et al. 2017). Compared to HybPhaser (Nauheimer et al. 2021) that phases hybrid sequences based on the similarity of hybrid reads to putative progenitor references, SORTER compares hybrid and progenitor sequence similarity based on phased allelic sequences derived from consensus-allele assemblies which may improve homeolog inference based on longer sequences with more informative sites than shorter paired end reads. With adequate read depth, we were able to recover up to four homeologous sequences per locus from several allotetraploids. SORTER is designed to be applicable to hybrids of higher ploidy levels, making it amenable to most polyploid reticulation complexes. For this study, we generated allopolyploid datasets to infer the progenitors of one allopolyploid individual per dataset to simplify analyses and presentation of results gauging the effect of progenitor identity on homeolog inference. Although not employed here, we provide an optional script with identical inference methods that allows the inclusion of multiple allopolyploid individuals and taxa in a single dataset.

Acknowledgments

We thank all herbaria (ALA, CITRO, DUKE, KANU, LAF, MEXU, NYBG, RSA, UC, UPS, VT) that lent specimens and allowed destructive sampling of vouchers for DNA extraction. This study would not have been possible without the GoFlag project (NSF # 1541506) that generated the TC probes used, built our Illumina libraries, and sequenced our samples. Support for the development of SORTER was provided, in part, by NSF # 1920858 to E.M.S. We thank Lucy Kerhoulas, Nick Kerhoulas, Michael Sundue, Michael Kessler, and Weston Testo for providing many specimens. We extend additional thanks to Chris Haufler, George Tiley, Nic Kooyers, Carl Rothfels, and three anonymous reviewers for providing comments and suggestions on our manuscript.

Author Contributions

JMR designed the study, wrote the majority of bioinformatic scripts, processed data, performed phylogenetic analyses, and wrote the manuscript. JGB assisted with study design and development of bioinformatic scripts, generated the data, and edited the manuscript. EMS helped with design of the study, generated bioinformatic scripts, and edited the manuscript.

Literature Cited

Aguiar, D. and S. Istrail. 2012. HapCompass: A fast cycle basis algorithm for accurate haplotype assembly of sequence data. Journal of Computational Biology 19: 577–590.

Almeida, T. E., A. Salino, J.-Y. Dubuisson, and S. Hennequin. 2017. Adetogramma (Polypodiaceae), a new monotypic fern genus segregated from Polypodium. PhytoKeys 78: 109–131.

Andermann, T., A. M. Fernandes, U. Olsson, M. Töpel, B. Pfeil, B. Oxelman, A. Aleixo, B. C. Faircloth, and A. Antonelli. 2019. Allele phasing greatly improves the phylogenetic utility of ultraconserved elements. Systematic Biology 68: 32–46.

Andermann, T., M. F Torres Jimenez, P. Matos-Maravı, R. Batista, J. L. Blanco-Pastor, A. L. S. Gustafsson, L. Kistler, I. M. Liberal, B. Oxelman, C. D. Bacon, and A. Antonelli. 2020. A guide to carrying out a phylogenomic target sequence capture project. Frontiers in Genetics 10: 1–14.

Assis, F. C., T. E. Almeida, S. J. Russell, H. Schneider, and A. Salino. 2016. Molecular phylogeny and recircumscription of the fern genus Pecluma (Polypodiaceae–Polypodiopsida). Phytotaxa 247: 235–246.

Beck, J. B., M. L. Markley, M. G. Zielke, J. R. Thomas, H. J. Hale, L. D. Williams, and M. G. Johnson. 2021. Is Palmer's elm leaf goldenrod real? The Angiosperms353 kit provides within-species signal in Solidago ulmifolia s.l. BioRxiv 2021: 1–15.

Bloesch, Z., L. Nauheimer, T. E. Almeida, D. Crayn, and A. R. Field. 2022. HybPhaser identifies hybrid evolution in Australian Thelypteridaceae. Molecular Phylogenetics and Evolution 173: 107526.

Branstetter, M. G., J. T. Longino, P. S. Ward, and B. C. Faircloth. 2017. Enriching the ant tree of life: Enhanced UCE bait set for genome-scale phylogenetics of ants and other Hymenoptera. Methods in Ecology and Evolution 8: 768–776.

Breinholt, J. W., S. B. Carey, G. P. Tiley, E. C. Davis, L. Endara, S. F. McDaniel, L. G. Neves, E. B. Sessa, M. von Konrat, S. Chantanaorrapint, S. Fawcett, S. M. Ickert-Bond, P. H. Labiak, J. Larraın, M. Lehnert, L. R. Lewis, N. S. Nagalingum, N. Patel, S. A. Rensing, W. Testo, A. Vasco, J. C. Villareal, E. Webb Williams, and J. G. Burleigh. 2021. A target enrichment probe set for resolving the flagellate land plant tree of life. Applications in Plant Sciences 9: e11406.

Bryan, F. A. and D. E. Soltis. 1987. Electrophoretic evidence for allopolyploidy in the fern Polypodium virginianum. Systematic Botany 12: 553–561.

Carvajal-Hernandez, C. I. 2018. A new species, new combinations in Pecluma and Pleopeltis, and new records for the state of Veracruz, Mexico. American Fern Journal 108: 139–150.

Cranfill, R. and D. M. Britton. 1983. Typification within the Polypodium virginianum complex (Polypodiaceae). Taxon 32: 557–560.

Degnan, J. H. and N. A. Rosenberg. 2009. Gene tree discordance, phylogenetic inference and the multispecies coalescent. Trends in Ecology & Evolution 24: 332–340.

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.

Faircloth, B. C. 2016. PHYLUCE is a software package for the analysis of conserved genomic loci. Bioinformatics 32: 786–788.

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.

Faircloth, B. C., F. Alda, K. Hoekzema, M. D. Burns, C. Oliveira, J. S. Albert, B. F. Melo, L. Ochoa, F. Roxo, P. Chakrabarty, B. Sidlauskas, and M. Alfaro. 2020. A target enrichment bait set for studying relationships among Ostariophysan fishes. Copeia 108: 47–60.

Fawcett, S., A. R. Smith, M. Sundue, J. G. Burleigh, E. B. Sessa, L.-Y. Kuo, C.-W. Chen, et al. 2021. A global phylogenomic study of the Thelypteridaceae. Systematic Botany 46: 891–915.

Freyman, W. A., M. G. Johnson, and C. J. Rothfels. 2022. homologizer: Phylogenetic phasing of gene copies into polyploid subgenomes. BioRxiv 2022: 1–19.

Haufler, C. H. and T. A. Ranker. 1995. RbcL sequences provide phylogenetic insights among sister species of the fern genus Polypodium. American Fern Journal 85: 361–374.

Haufler, C. H. and M. D. Windham. 1991. New species of North American Cystopteris and Polypodium, with comments on their reticulate relationships. American Fern Journal 81: 7–23.

Haufler, C. H. and W. Zhongren. 1991. Chromosomal analyses and the origin of allopolyploid Polypodium virginianum (Polypodiaceae). American Journal of Botany 78: 624–629.

Haufler, C. H., M. D. Windham, F. A. Lang, and S. A. Whitmore. 1993. Polypodium. Pp. 356–357 in Flora of North America North of Mexico vol. 2, ed. Flora of North America Editorial Committee. New York and Oxford: Oxford University Press.

Haufler, C. H., D. E. Soltis, and P. S. Soltis. 1995a. Phylogeny of the Polypodium vulgare complex: Insights from chloroplast DNA restriction site data. Systematic Botany 20: 110–119.

Haufler, C. H., M. D. Windham, and E. W. Rabe. 1995b. Reticulate evolution in the Polypodium vulgare complex. Systematic Botany 20: 89–109.

Haufler, C. H., E. A. Hooper, and J. P. Therrien. 2000. Modes and mechanisms of speciation in pteridophytes: Implications of contrasting patterns in ferns representing temperate and tropical habitats. Plant Species Biology 15: 223–236.

Huang, C., X. Qi, D. Chen, J. Qi, and H. Ma. 2020. Recurrent genome duplication events likely contributed to both the ancient and recent rise of ferns. Journal of Integrative Plant Biology 62: 433–455.

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: 1–7.

Johnson, M. G., L. Pokorny, S. Dodsworth, L. R. Botigue, R. S. Cowan, A. Devault, W. L. Eiserhardt, N. Epitawalage, F. Forest, J. T. Kim, J. H. Leebens-Mack, I. J. Leitch, O. Maurin, D. E. Soltis, P. S. Soltis, G. K.-S. Wong, W. J. Baker, and N. J. Wickett. 2019. A universal probe set for targeted sequencing of 353 nuclear genes from any flowering plant designed using k-medoids clustering. Systematic Biology 68: 594–606.

Kamneva, O. K., J. Syring, A. Liston, and N. A. Rosenberg. 2017. Evaluating allopolyploid origins in strawberries (Fragaria) using haplotypes generated from target capture sequencing. BMC Evolutionary Biology 17: 1–19.

Kates, H. R., M. G. Johnson, E. M. Gardner, N. J. C. Zerega, and N. J. Wickett. 2018. Allele phasing has minimal impact on phylogenetic reconstruction from targeted nuclear gene sequences in a case study of Artocarpus. American Journal of Botany 105: 404–416.

Kott, L. S. and D. M. Britton. 1982. A comparative study of sporophyte morphology of the three cytotypes of Polypodium virginianum in Ontario. Canadian Journal of Botany 60: 1360–1370.

Lang, F. A. 1965. Cytotaxonomic Study of the Polypodium vulgare Complex in Northwestern North America. Doctoral thesis. Vancouver, British Columbia: University of British Columbia.

Lang, F. A. 1969. A new name for a species of Polypodium from northwestern North America. Madroño 20: 53–60.

Lang, F. A. 1971. The Polypodium vulgare complex in the Pacific Northwest. Madroño 21: 235–254.

Lischer, H. E. L., L. Excoffier, and G. Heckel. 2014. Ignoring heterozygous sites biases phylogenomic estimates of divergence times: implications for the evolutionary history of Microtus voles. Molecular Biology and Evolution 31: 817–831.

Liu, L., L. Yu, L. Kubatko, D. K. Pearl, and S. V. Edwards. 2009. Coalescent methods for estimating phylogenetic trees. Molecular Phylogenetics and Evolution 53: 320–328.

Liu, L., Z. Xi, S. Wu, C. C. Davis, and S. V. Edwards. 2015. Estimating phylogenetic trees from genome-scale data. Annals of the New York Academy of Sciences 1360: 36–53.

Luna-Vega, I., J. D. Tejero-Dıez, R. Contreras-Medina, M. Heads, and G. Rivas. 2012. Biogeographical analysis of two Polypodium species complexes (Polypodiaceae) in Mexico and Central America. Biological Journal of the Linnean Society. Linnean Society of London 106: 940–955.

Mamanova, L., A. J. Coffey, C. E. Scott, I. Kozarewa, E. H. Turner, A. Kumar, E. Howard, J. Shendure, and D. J. Turner. 2010. Target-enrichment strategies for next-generation sequencing. Nature Methods 7: 111–118.

Mandel, J. R., R. B. Dikow, V. A. Funk, R. R. Masalia, S. E. Staton, A. Kozik, R. W. Michelmore, L. H. Rieseberg, and J. M. Burke. 2014. A target enrichment method for gathering phylogenetic information from hundreds of loci: an example from the Compositae. Applications in Plant Sciences 2: 1–6.

Manton, I. 1947. Polyploidy in Polypodium vulgare. Nature 159: 136.

Manton, I. 1950. Problems of Cytology and Evolution in the Pteridophyta. Cambridge: Cambridge University Press.

Manton, I. 1951. Cytology of Polypodium in America. Nature 167: 37.

Manton, I. 1957. The problem of Polypodium virginianum. American Fern Journal 47: 129–134.

Martens, P. 1943. Les organes glanduleux de Polypodium virginianum (P. vulgare var. virginianum) I. Valeur systematique et repartition geographique. Bulletin du Jardin botanique de l'Etat a Bruxelles 17: 1–14.

McCormack, J. E., S. M. Hird, A. J. Zellmer, B. C. Carstens, and R. T. Brumfield. 2013. Applications of next-generation sequencing to phylogeography and phylogenetics. Molecular Phylogenetics and Evolution 66: 526–538.

Mendez-Reneau, J., J. G. Burleigh, and E. M. Sigel. 2022. Data from: Tar-get capture methods offer insight into the evolution of rapidly diverged taxa and resolve allopolyploid homeologs in the fern genus Polypodium s.s. Dryad Digital Repository. https://doi.org/10.5061/dryad.2z34tmpq9.

Mickel, J. T. and J. M. Beitel. 1988. Pteridophyte flora of Oaxaca, Mexico. Memoirs of the New York Botanical Garden 46: 1–568.

Mickel, J. T. and A. R. Smith. 2004. The Pteridophytes of Mexico. New York: The New York Botanical Garden Press.

Miller, M. A., W. Pfeiffer, and T. Schwartz. 2010. Creating the CIPRES Science Gateway for inference of large phylogenetic trees. Pp. 1–8 in Proceedings of the Gateway Computing Environments Workshop (GCE). New Orleans: Gateway Computing.

Minh, B. Q., M. W. Hahn, and R. Lanfear. 2020. New methods to calculate concordance factors for phylogenomic datasets. Molecular Biology and Evolution 37: 2727–2733.

Mirarab, S. 2019. Species Tree Estimation Using ASTRAL: Practical Considerations. arXiv preprint: 1–16.

Nakhleh, L. 2013. Computational approaches to species phylogeny inference and gene tree reconciliation. Trends in Ecology & Evolution 28: 719–728.

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: 1–19.

Nicholls, J. A., R. T. Pennington, E. J. M. Koenen, C. E. Hughes, J. Hearn, L. Bunnefeld, K. G. Dexter, G. N. Stone, and C. A. Kidner. 2015. Using targeted enrichment of nuclear genes to increase phylogenetic resolution in the neotropical rain forest genus Inga (Leguminosae: Mimosoideae). Frontiers in Plant Science 6: 00710.

Otto, E. M., T. Janßen, H.-P. Kreier, and H. Schneider. 2009. New insights into the phylogeny of Pleopeltis and related Neotropical genera (Polypodiaceae, Polypodiopsida). Molecular Phylogenetics and Evolution 53: 190–201.

Otto, S. P. and J. Whitton. 2000. Polyploid incidence and evolution. Annual Review of Genetics 34: 401–437.

Peterson, R. L. and L. S. Kott. 1974. The sorus of Polypodium virginianum: Some aspects of the development and structure of paraphyses and sporangia. Canadian Journal of Botany 52: 2283–2288.

Rothfels, C. J., K. M. Pryer, and F. Li. 2017. Next-generation polyploid phylogenetics: Rapid resolution of hybrid polyploid complexes using PacBio single-molecule sequencing. The New Phytologist 213: 413–429.

Schneider, H., A. R. Smith, R. Cranfill, T. J. Hildebrand, C. H. Haufler, and T. A. Ranker. 2004. Unraveling the phylogeny of polygrammoid ferns (Polypodiaceae and Grammitidaceae): exploring aspects of the diversification of epiphytic plants. Molecular Phylogenetics and Evolution 31: 1041–1063.

Schneider, H.-P. K. H. 2006. Phylogeny and biogeography of the staghorn fern genus Platycerium (Polypodiaceae, Polypodiidae). American Journal of Botany 93: 217–225.

Sessa, E. B., J. A. Banks, M. S. Barker, J. P. Der, A. M. Duffy, S. W. Graham, M. Hasebe, J. Langdale, F.-W. Li, D. B. Marchant, K. M. Pryer, C. J. Rothfels, S. J. Roux, M. L. Salmi, R. M. Sigel, D. E. Soltis, P. S. Soltis, D. W. Stevenson, and P. G. Wolf. 2014. Between Two Fern Genomes. GigaScience 3: 1–17.

Shalimov, A. P. and A. I. Shmakov. 2016. Spore morphology of Polypodium aleuticum (Polypodiaceae) and related species. Biological Bulletin of Bogdan Chmelnitskiy Melitopol State Pedagogical University 6: 28–33.

Shalimov, A. P., A. V. Vaganov, and A. I. Shmakov. 2011. Spore morphology of Polypodium (Polypodiaceae) species from Russia. Turczaninowia 14: 5–13.

Sigel, E. M., M. D. Windham, C. H. Haufler, and K. M. Pryer. 2014a. Phylogeny, divergence time estimates, and phylogeography of the diploid species of the Polypodium vulgare complex (Polypodiaceae). Systematic Botany 39: 1042–1055.

Sigel, E. M., M. D. Windham, and K. M. Pryer. 2014b. Evidence for reciprocal origins in Polypodium hesperium (Polypodiaceae): A fern model system for investigating how multiple origins shape allopolyploid genomes. American Journal of Botany 101: 1476–1485.

Silva, D. N., V. Varzea, O. S. Paulo, and D. Batista. 2018. Population genomic footprints of host adaptation, introgression and recombination in coffee leaf rust. Molecular Plant Pathology 19: 1742–1753.

Siplivinski, V. 1974. Notulae de Flora Baicalensi, II. Novosti Sistematiki Vyshikh Rastenii 11: 327–337.

Smith, A. R., H.-P. Kreier, C. H. Haufler, T. A. Ranker, and H. Schneider. 2006. Serpocaulon (Polypodiaceae), a new genus segregated from Polypodium. Taxon 55: 919–930.

Smith, B. T., M. G. Harvey, B. C. Faircloth, T. C. Glenn, and R. T. Brumfield. 2014. Target capture and massively parallel sequencing of

ultraconserved elements for comparative studies at shallow evolutionary time scales. Systematic Biology 63: 83–95.

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

Tejero-Dıez, J. D. 2005. Revision Taxonomica del Complejo Polypodium plesiosorum (Polypodiaceae, Polypodiophyta). Doctoral dissertation. Unidad Xochimilco: Universidad Autonoma Metropolitana.

Tejero-Dıez, J. D. and L. Pacheco. 2004. Notas taxonomicas y de distribucion en Polypodium plesiosorum y P. rhodopleuron (Polypodiaceae: Pteridophyta). Anales del Instituto de Biologıa. Serie Botanica 75: 11–37.

Tejero-Dıez, D. J., S. Aguilar-Rodrıguez, T. Terrazas, and L. Pacheco. 2010. Arquitectura y anatomıa foliar del complejo Polypodium plesiosorum sensu Moran (Polypodiaceae). Revista de Biologıa Tropical 58: 955–976.

Tiley, G. P., A. A. Crowl, P. S. Manos, E. B. Sessa, C. Solıs-Lemus, A. D. Yoder, and J. G. Burleigh. 2021. Phasing alleles improves network inference with allopolyploids. BioRxiv 2021: 1–41.

Vaidya, G., D. J. Lohman, and R. Meier. 2011. SequenceMatrix: Concatenation software for the fast assembly of multi-gene datasets with character set and codon information. Cladistics 27: 171–180.

Weitemier, K., S. C. K. Straub, R. C. Cronn, M. Fishbein, R. Schmickl, A. McDonnell, and A. Liston. 2014. Hyb-Seq: Combining target enrichment and genome skimming for plant phylogenomics. Applications in Plant Sciences 2: 1–7.

Whitmore, S. A. and A. R. Smith. 1991. Recognition of the tetraploid Polypodium calirhiza (Polypodiaceae) in western North America. Madroño 38: 233–248.

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.

Zhang, C., M. Rabiee, E. Sayyari, and S. Mirarab. 2018. ASTRAL-III: Polynomial time species tree reconstruction from partially resolved gene trees. BMC Bioinformatics 19: 15–30.

APPENDIX 1.— Specimens included in this study. Taxon name (unique specimen identifier if more than one specimen), collection locality, collector and collection number (herbarium).

Pecluma camptophyllaria (Fee) M.G. Price: (1) Puerto Rico, SF 468 (VT); Pecluma dulcis (Poir.) F.C. Assis & Salino: (2) Costa Rica, Testo 761 (VT); Pecluma rhachypterygia (Liebm.) F.C. Assis & Salino: (3) Mexico, Sundue 3622 (VT); Phlebodium araneosum (M. Martens & Galeotti) Mickel & Beitel: (1) Mexico, Coyoacan, Sundue 4193 (VT); Phlebodium decumanum (Willd.) J. Sm.: (2) Cultivated, Ullman s.n. (VT); Platycerium andinum Baker: Cultivated, Farnwerk Gardens, Switzerland, Testo s.n. (VT); Pleopeltis angusta Humb. & Bonpl. ex Willd.: Mexico, Testo 1086 (VT); Pleopeltis monosora (Desv.) A.R. Sm.: Ecuador, Pichincha, Sundue 2596 (VT); Pleurosoriopsis makinoi (Maxim. ex Makino) Fomin: Japan, Fukui-Shi, Ebihara 2972 (DUKE); Polypodium alavae A.R. Sm.: Mexico, Mejia 857 (MEXU); Polypodium amorphum Suksd.: (1) Washington, Mt. Baker, Sigel 2010-75 (DUKE); (2) Oregon, Multonomah County, Sigel 2010-104 (DUKE); Polypodium appalachianum Haufler & Windham: (1) Canada, Nova Scotia, Oldham 38496 (DUKE); (2) North Carolina, Ashe County, Sigel 2010-61 (DUKE); Polypodium arcanum Maxon: Mexico, Testo 1069 (VT); Polypodium californicum Kaulf.: (1) California, San Luis Obispo, Hartman 1418 (KANU); (2) California, San Clemente Island, Dunkle 7248 (KANU); (3) Mexico, Baja, Tenorio 13356 (NYBG); (4) California, Santa Cruz Island, Thorne 52569 (RSA); (5) California, Anacapa Island, Junak 380 (RSA); (6) California, Marin County, Rose 70012 (KANU); (7) California, Santa Ana Mountains, LLK 1111 (KANU); (8) California, San Diego, Smith 2558 (UC); Polypodium calirhiza S.A.Whitmore & A.R.Sm.: (1) Mexico, Jalisco, Dyer 51 (DUKE); (2) California, San Francisco, Mendez-Reneau 5 (LAF); Polypodium cambricum L., Spain, Valencia, Woods 8-61 (RBGE); Polypodium conterminans Liebm.: (1) Mexico, Kromer 3624 (MEXU); (2) Mexico, Castillo-Hernandez 307 (MEXU); Polypodium diplotrichum Mickel & Beitel: Mexico, Mendoza 545 (MEXU); Polypodium eatonii Baker: Mexico, Zacapoaxtla, Schuettpelz 1601 (VT); Polypodium echinolepis Fee: (1) Mexico, Zacapoaxtla, Schuettpelz 1617 (VT); (2) Guatemala, Testo 1279 (VT); Polypodium eperopuetes Mickel & Beitel: Mexico, Chiapas, Carvajal 1178 (CITRO); Polypodium fauriei Christ: Japan, Fukui Prefecture, Ebihara 2973 (DUKE); Polypodium fissidens Maxon: (1) Mexico, Cerro Huitepec, Carvajal 1115 (CITRO); (2) Guatemala, Jimenez 1393 (MEXU); Polypodium fraternum Schltdl. & Cham.: Mexico, Zacapoaxtla, Schuettpelz 1647 (VT); Polypodium glycyrrhiza D.C. Eaton: (1) California, Del Norte County, Kerhoulas G26A4 (LAF); (2) California, Del Norte County, Kerhoulas G23A1 (LAF); (3) Alaska, Amatignak Island, Talbot AMT04-9 (ALAH); (4) Alaska, Attu Island, Parker 9221 (ALAH); (5) Canada, British Columbia, Rothfels 3086 (DUKE); (6) Oregon, Multonomah County, Sigel 2010-109 (DUKE); (7) California, Del Norte County, Morse 11378 (KANU); (8) California, Del Norte County, Kerhoulas G21A5 (LAF); Polypodium hesperium Maxon: (1) Arizona, Pinaleno Mountains, Yatskievych 81-459 (KANU); (2) California, San Bernardino Mountains, Kiefer 575 (KANU); Polypodium macaronesicum A.E. Bobrov: Spain, Tenerife Island, Larsson 47 (UPS); Polypodium martensii Mett.: (1) Mexico, Tamaulipas, LS 418 (CITRO); (2) Mexico, Zacapoaxtla, Schuettpelz 1723 (VT); Polypodium pellucidum Kaulf.: Hawaii, Kauai, Vernon s.n. (DUKE); Polypodium plesiosorum Kunze: Mexico, Testo 829 (VT); Polypodium puberulum Schltdl. & Cham.: (1) Mexico, Oaxaca, Testo 1532 (VT); (2) Mexico, Oaxaca, Testo 1519 (VT); (3) Mexico, Sundue 4070 (VT); Polypodium rhodopleuron Kunze: (1) Mexico, Schuettpelz 1620 (VT); (2) Mexico, CICH 1212 (VT); Polypodium saximontanum Windham: (1) Colorado, Smith 0182 (DUKE); (2) Montana, Sigel 2011-41a (DUKE); Polypodium scouleri Hook. & Grev.: (1) California, Kerhoulas S22A4 (LAF); (2) Canada, Rothfels 4475 (DUKE); (3) California, Huiet 141 (DUKE); (4) California, Halse 7135 (KANU); Polypodium sibiricum Sipliv.: (1) Mongolia, Wu 123 (MO); (2) Japan, Mt. Tiene, Haufler s.n. (KANU); (3) Canada, Saskatchewan, Haufler s.n. (KANU); (4) Russia, Altai Mountains, Argus 13190 (KANU); Polypodium virginianum L.: (1), Canada, Ontario, Rothfels 3943 (DUKE); (2) North Carolina, Rothfels 4031 (DUKE); Polypodium vulgare L.: (1) Switzerland, Bern Canton, Sigel 2010-55 (DUKE); (2) Norway, Lyngen County, Larsson 298 (DUKE); Serpocaulon attenuatum (C. Presl) A.R. Sm.: Costa Rica, Puntarenas, Testo 1241 (VT); Serpocaulon levigatum (Cav.) A.R. Sm.: Colombia, Testo 1157 (VT).