Revealing the evolutionary history of a reticulate polyploid complex in the genus *Isoëtes*

^{1,2,4}Jacob S. Suissa*, ³Sylvia P. Kinoshian, ^{4,5}Peter W. Schafran, ⁶Jay F. Bolin, ⁴W. Carl Taylor,
⁴Elizabeth A. Zimmer

¹The Department of Organismic and Evolutionary Biology, Harvard University, Cambridge, MA; ²The Arnold Arboretum of Harvard University, Boston, MA; ³Department of Biology & Ecology, Utah State University, Logan, UT 84322, USA; ⁴Department of Botany, National Museum of Natural History, Smithsonian Institution, Washington, DC, USA; ⁵Boyce Thompson Institute, Ithaca, NY, USA; ⁶Department of Biology, Catawba College, Salisbury, NC, USA

*Corresponding author: jsuissa@g.harvard.edu

Summary

- Polyploidy and hybridization are important processes in the evolution of spore-dispersed plants. Few studies, however, focus these dynamics in heterosporous lycophytes, such as *Isoetes*, where polyploid hybrids are common and thought to be important in the generation of their extant diversity. We investigate reticulate evolution in a complex of western North American quillworts (*Isoetes*) and provide insights into the evolutionary history of hybrids, and the role of polyploidy in maintaining novel diversity.
- We utilize low copy nuclear markers, whole plastomes, restriction site-associated DNA sequencing, cytology, and reproductive status (fertile or sterile) to investigate the reticulate evolutionary history of western North American *Isoetes*.
- We reconstruct the reticulate evolutionary history and directionality of hybridization events in this complex. The presence of high level polyploids, plus frequent homoploid and interploid hybridization suggests that there are low prezygotic reproductive barriers in this complex, hybridization is common and bidirectional between similar—but not divergent—cytotypes, and that allopolyploidization is important to restore fertility in some hybrid taxa.
- Our data provide five lines of evidence suggesting that hybridization and polyploidy can occur with frequency in the genus, and these evolutionary processes may be important in shaping extant *Isoetes* diversity.

Key words: evolution, hybrid, lycophytes, polyploid, speciation, quillwort, RADseq, reticulation

Introduction

Processes such as hybridization and genome duplication (polyploidization) can lead to evolutionary histories that are better explained through reticulating, rather than bifurcating, branches on a phylogenetic tree (Soltis & Soltis, 2000; Linder & Rieseberg, 2004; Nakhleh *et al.*, 2005; Degnan & Rosenberg, 2009; Edelman *et al.*, 2019). These evolutionary processes are important because they lead to the generation and maintenance of novel lineages. These lineages form the basis for reticulate species complexes, which are webs of related taxa involving multiple parental species, sterile (or partially sterile) interspecific hybrids (Dobzhansky, 1982), and fertile auto- or allopolyploids (i.e., products of genome duplication with one or more prenatal genomes, respectively; Stebbins, 1969; Grant, 1971; Rieseberg, 1997). Sterile hybrids within these complexes are sometimes considered evolutionary dead ends, but through allopolyploidization, fertility can be restored because each chromosome has a match in its newly duplicated genome (Stebbins, 1947; Wagner & Wagner, 1980; Soltis & Soltis, 2000). In addition, allopolyploidy can lead to fixed genomic variation (heterozygosity) allowing for lower inbreeding depression following self-fertilization (Soltis & Soltis, 2000; Husband *et al.*, 2008; Soltis *et al.*, 2014; Haufler *et al.*, 2016). For these reasons, allopolyploidy is suggested to be evolutionarily advantageous within some lineages—however the importance of this process as a broader driver of diversification is debated (Mayrose *et al.*, 2011, 2015; Soltis *et al.*, 2014; Landis *et al.*, 2018).

Reticulate species complexes are common in plants (Burnier *et al.*, 2009; Nauheimer *et al.*, 2019; Sandstedt *et al.*, 2020), especially in spore-dispersed vascular plants (ferns and lycophytes) (Barrington *et al.*, 1989; Wood *et al.*, 2009; Sigel, 2016). Among spore dispersed plants, a majority of this research focuses on homosporous ferns, with less work on the lycophytes, and especially the genus *Isoetes* L (Isoetaceae) where these processes are thought to be rampant (Taylor & Luebke, 1988; Hickey *et al.*, 1989; Britton & Brunton, 1989; Taylor & Hickey, 1992; Brunton & Britton, 2000; Small & Hickey, 2001; Pereira *et al.*, 2018). *Isoetes* is a cosmopolitan genus comprising 250-350 species, ~60% of which are polyploid (Troia *et al.*, 2016). This genus has been described as a “living fossil” as it is the last remaining lineage of the ancient Isoetalean lycopsids—a clade that dates back to the middle Devonian (~400mya; Pigg, 1992). However, most of the extant diversity is relatively young (<60mya; Wood *et al.*, 2020) and a product of recent diversification where hybridization and polyploidization are thought to

play an important role in species generation (Taylor & Hickey, 1992; Hoot *et al.*, 2004; Dai *et al.*, 2020). *Isoetes* is an ideal study system for investigating the evolutionary implications of hybridization and polyploidization within vascular plants because of its unique evolutionary dynamics (i.e., an old clade with recent diversification), high species richness, pervasiveness of polyploids, and complex biogeographical patterns.

Although there is great potential for investigating reticulate evolution within *Isoetes*, discerning species relationships in the genus has historically proven difficult due to the conserved body plan with limited morphological characters and character states (Pfeiffer, 1922; Hickey *et al.*, 1989). In fact, early Isoetologist W. N. Clute (1905) lamented, “*the marks by which the species [of Isoetes] are distinguished are so obscure as to be puzzling to all but the select few, and in consequence the species have been largely taken upon by faith.*” In addition, recent work demonstrates that low molecular divergence further compounds the difficulties of circumscribing species within this genus (Pereira *et al.*, 2017; Wood *et al.*, 2020).

One species complex of *Isoetes* that has remained particularly challenging is known from western North America. This complex includes two diploid progenitors (*I. echinospora* Durieu and *I. bolanderi* Engelm.), a sterile diploid hybrid (*I. × herb-wagneri* W. C. Taylor), a fertile allotetraploid (*I. maritima* Underwood), fertile hexaploid (*I. occidentalis* L. F. Hend.), and a handful of sterile interploid hybrids (*I. × pseudotruncata* D. M. Britton & D. F. Brunt. and *I. × truncata* Clute; Fig. 1, 2a). The evolutionary relationships among taxa in this complex have been suggested before using morphology and cytology, but a consensus has not been reached (Fig. 1; Britton & Brunton, 1993, 1996; Britton *et al.*, 1999; Taylor, 2002).

Here, we revisit this complex with new data and use phylogenomic comparisons to investigate the evolutionary implications of polyploidy and hybridization. We leverage five distinct data sets to accomplish this: low copy nuclear markers, whole chloroplast genomes, restriction site-associated DNA sequencing (RADseq), cytology (genome size estimates and chromosome counts), and reproductive status (fertile or sterile) from spore morphology. With these data, we are able to reconstruct the reticulate evolutionary history and directionality of hybridization events, as well as ploidy and fertility of hybrid taxa. This is the first phylogenomic study investigating the evolutionary history of a polyploid complex in *Isoetes*. These data provide robust evidence that hybridization and polyploidy occur with frequency in the genus and

that allopolyploidy—and potentially autopolyploidy—may be important evolutionary processes shaping extant *Isoetes* diversity.

Materials and Methods

We collected five datasets to investigate the reticulate relationships in a complex of western North American *Isoetes*. We performed flow cytometry to estimate genome size; utilized Illumina short-read ddRADseq to obtain genomic data for population structure analysis and phylogenetic network analysis; identified fertile vs. aborted spores; sequenced whole chloroplasts with Illumina; and sequenced all alleles of the *LFY* nuclear marker with the Pacific Biosciences long-read platform to perform phylogenetic analyses.

Taxon sampling

We assembled five distinct datasets, with collections from various sources. For RADseq and cytology, we collected fresh and silica dried leaf tissue from localities with historical information for each taxon across western North America (Fig. 2a). These included high-elevation lakes in Montana, California, and British Columbia. To determine specimen fertility, we examined spores from these collections. Specimen voucher information can be found in (Supplementary Table S1). For chloroplast genome sequencing and *LFY* allele sequencing we sampled from silica dried specimens of vouchered specimens (Table S2).

Cytology

Fresh leaf material was used to measure C-values for individuals of each taxon in order to estimate their ploidy. C-values were estimated using microphylls from living plants. Leaves were prepared for DNA flow cytometry as described in Bolin *et al.*, (2017) using propidium iodide stain (P121493; Molecular Probes FluoroPure, ThermoFisher Scientific, Waltham, Massachusetts, USA). Newly expanded leaves of the *Glycine max* ‘Polanka’ (Doležel *et al.*, 1994) were used as the standard for genome size estimation. Samples were analyzed with a BD Accuri C6 Flow Cytometer and associated software (Becton Dickinson and Company, Franklin Lakes, New Jersey, USA), using gating and peak estimation parameters described in Bolin *et al.*, (2017). C-values were calculated for each individual following the methods of Doležel & Bartos,

(2005). By measuring the C-value for taxa with well documented chromosome counts we were able to infer ploidy levels for individual plants using C-value estimates.

RADseq DNA extraction, library construction, and sequencing

Silica-dried plant tissue samples were supplied to the University of Wisconsin-Madison Biotechnology Center for DNA extractions, library preparation, and sequencing. DNA was extracted using the QIAGEN DNeasy mericon 96 QIAcube HT Kit and quantified using the Quant-iT™ PicoGreen dsDNA kit (Life Technologies, Grand Island, NY).

Libraries were prepared following Elshire *et al.*, (2011) with minimal modification. 100 ng of DNA were digested using PstI and MspI (New England Biolabs, Ipswich, MA) after which barcoded adapters amenable to Illumina sequencing were added by ligation with T4 ligase (New England Biolabs, Ipswich, MA). The 96 adapter-ligated samples were pooled and amplified to provide library quantities amenable for sequencing, and adapter dimers were removed by SPRI bead purification. Quality and quantity of the finished libraries were assessed using the Agilent Bioanalyzer High Sensitivity Chip (Agilent Technologies, Inc., Santa Clara, CA) and Qubit dsDNA HS Assay Kit (Life Technologies, Grand Island, NY), respectively. Size selection was performed to obtain 300 - 450 BP fragments. Sequencing was done on Illumina NovaSeq 6000 2x150 S2. Data were then analyzed using the standard Illumina Pipeline, version 1.8.2.

RADseq data processing

All RADseq data processing was done on the Pearse Lab (Utah State University) high-performance workstation; downstream analyses were performed on the University of Utah Center for Higher Performance Computing. Raw data were demultiplexed using *stacks* v. 2.4 *process_radtags* (Catchen *et al.*, 2011, 2013) allowing for a maximum of one mismatch per barcode. Demultiplexed FASTQ files were paired and merged using *ipyrad* v. 0.9.52 (Eaton & Overcast, 2020). Low quality bases, adapters, and primers were removed from each read. Filtered reads were clustered at 90% similarity, and we required a sequencing depth of 6 or greater per base and a minimum of 35 samples per locus to be included in the final assembly. The *ipyrad* pipeline defines a locus as a short sequence present across samples. From each locus retained in the final assembly, *ipyrad* identifies single nucleotide polymorphisms (SNPs); these SNPs are the variation used in our downstream analyses. Although many of the species of *Isoetes*

included in this study are polyploid, *stacks* and *ipyrad* assume diploidy, so we processed all samples as such.

Population structure analysis

The program STRUCTURE v. 2.3.4 (Pritchard *et al.*, 2000) estimates admixture (hybridization) between populations of closely related taxa. The program operates with the assumption that each individual's genome is a mosaic from K source or ancestral populations. In our analysis, we ran STRUCTURE for $K = 2 - 5$ with 50 chains for each K . We then used CLUMPAK (Kopelman *et al.*, 2015) to process the STRUCTURE output and estimate the best K values (Pritchard *et al.*, 2000; Evanno *et al.*, 2005).

Split network analysis

Split networks are implicit representations of reticulate evolution. They incorporate uncertainty by depicting multiple phylogenetic hypotheses: internal nodes do not represent ancestral taxa, but rather incongruencies between potential evolutionary histories. We utilized the NeighborNet (Bryant & Moulton, 2002) split network algorithm in SplitsTree v. 4.16.1 (Huson & Bryant, 2006). NeighborNet is similar to a neighbor joining tree in that it pairs samples based on similarity, sorting taxa into larger and larger groups (Bryant & Moulton, 2002). NeighborNet creates a network where parallel lines indicate splits of taxa, and boxes created by these lines indicate conflicting signals (Bryant & Moulton, 2002).

Identification of fertility

Spores from sequenced individuals were photographed under a Zeiss Discovery v12 instrument (Carl Zeiss Microscopy, LLC, Thornwood, NY, USA). Some individuals were too immature to have viable spores and were not included. Individual spores were visualized so as to detect any abnormalities that would indicate sterility, such as flattening of proximal hemisphere, polymorphism, irregularity in size and texture, and connections of meiotic tetrads (Wagner *et al.*, 1986).

DNA extraction for Illumina and PacBio sequencing

For whole plastome and nuclear gene sequencing, total genomic DNA was isolated from silica dried leaf tissue using the DNeasy Plant Mini Kit (Qiagen Inc., Valencia, CA, USA) at the Smithsonian NMNH Laboratory of Analytical Biology (LAB). DNA quantity and quality were measured using a Qubit 2.0 fluorometer (ThermoFisher Scientific, Waltham, MA, USA) and an Epoch spectrophotometer (BioTek Instruments Inc., Winooski, VT, USA).

Chloroplast library preparation & sequencing

Whole genomic DNA was quantified using a Qubit 2.0 fluorometer (ThermoFisher Scientific, Waltham, MA, USA) and quality was measured by 260nm: 280nm absorption ratio using an Epoch spectrophotometer (BioTek Instruments Inc., Winooski, VT, USA). 150-1000ng of whole genomic DNA was diluted to 60µL and sheared to ~500BP fragments using the Q800R2 sonicator (Qsonica LLC, Newtown, CT, USA). Fragmented DNA was then prepared for sequencing on a MiSeq Illumina sequencer (Illumina Inc., San Diego, CA, USA) using the NEBNext Ultra II DNA Library Prep Kit for Illumina (New England Biolabs Inc., Ipswich, MA, USA). Manufacturers' instructions were followed for end repair, adaptor ligation, indexing, and PCR enrichment for each sample. Libraries were quantified using a ViiA™ 7 Real-Time PCR System (Applied Biosystems Corp., Foster City, CA, USA). Libraries were then pooled and diluted to 4 nM and submitted for sequencing.

Chloroplast genome assembly and phylogenetic inference

Illumina's BaseSpace database was used to download index separated raw reads. Cutadapt (Martin, 2011) was used to trim NEB indices and filter out low quality bases. Pairing, assembling and referencing the chloroplast reads were done according to Schafran *et al.*, (2018). Adapters and low quality bases were trimmed using Trimmomatic 0.39 (Bolger et al. 2004). Paired reads were then mapped to the reference plastome of *Isoetes flaccida* (Karol *et al.*, 2010) using Bowtie2 (Langmead & Salzberg, 2012), to extract chloroplast reads. The putative chloroplast reads were then *de novo* assembled using SPAdes 3.10.1 with k-mer lengths 21, 33, 55, 66, 99, and 127bp (Bankevich *et al.*, 2012). Reference-based assemblies were constructed in Geneious Prime 2020.2.3 (<https://www.geneious.com>) by mapping putative chloroplast reads to the *I. flaccida* plastome and extracting the majority consensus sequence. For each sample, *de novo* assembled contigs were aligned to the reference-based assembly and any discrepancies

were manually corrected using mapped reads. Whole plastomes were aligned with previously sequenced plastomes (Schafran *et al.*, 2018) using MAFFT in Geneious (Kearse *et al.*, 2012) and trees were generated using RAxML 7.3.0 (Stamatakis, 2006) and MrBayes 3.2.6 (Ronquist & Huelsenbeck, 2003). The GTR+G model, with 4000 bootstrap iterations was implemented in RAxML. All MrBayes analyses were run for 50 million generations, sampling every 1,000 generations, with one cold chain and three heated chains. MCMC output was visualized in Tracer 1.7 to check for convergence among the chains (Rambaut *et al.*, 2018). These computations were run on the FASRC Odyssey cluster supported by the FAS Division of Science Research Computing Group at Harvard University.

PacBio library preparation

In preparation for PacBio sequencing the second intron of the nuclear gene *LEAFY* (*LFY*) was amplified by the polymerase chain reaction (PCR), using modified 30F (5'-GATCTTTATGAACAATGTGG-3') and 1190R (5'-GAAATACCTGATTTGTAACC3') primers. Amplifications in 25µL reactions were conducted using the following reagents: 12.5µL GoTaq (Taq polymerase master mix), 0.5µL of 0.1mM BSA (Bovine Serum Albumin), 1µL of DMSO (Dimethyl Sulfoxide), 1µL of 10µM Forward and Reverse primer, respectively, 6.5µL of Nuclease free H₂O, and 2.5µL of DNA. PCR protocol began with an initial denaturation temperature of 94°C for 4 minutes, 10 cycles of 94°C for 1 minute, 58°C for 1 minute and 72°C for 1 min, followed by 25 cycles of 94°C for 30 seconds, 53°C for 1 minute, and 72°C for 30 seconds. A final extension cycle of 72°C for 4 minutes followed the reaction. *LFY* amplicons were bead cleaned using Kapa beads following the (manufacturer's protocol). A 1:1 ratio of bead solution to amplicon was used to size select for ~1100bp fragments. The solution was resuspended in 20µL of nuclease free H₂O. Each cleaned sample was then proportionally pooled, based on concentration and ploidy-level into a 1.5ml microcentrifuge tube and sent to the Duke University Sequencing and Genomic Technologies Shared Resource (Durham, NC) for PacBio RSII sequencing using two SMRT cells.

PacBio data processing and phylogenetic inference

Circular consensus sequences (CCS) provided by the Duke University Sequencing and Genomic Technologies Shared Resource were size filtered to the expected size of the amplicon

(1000-1200 bp) and then demultiplexed and clustered with PURC (Rothfels *et al.*, 2017; Dauphin *et al.*, 2018). Demultiplexed reads were also filtered by DADA2 (Callahan *et al.*, 2016) to remove those with any “N” base calls and with more than 5 expected errors. Amplicon sequence variants (ASVs) were identified from the filtered, demultiplexed reads also using DADA2. Mock community data show that DADA2 more reliably recovers the true composition of a mixture of amplicon CCS reads compared to OTU clustering, which tends to oversplit clusters (Nelson *et al.*, 2020a). Therefore, we use only ASVs in downstream analyses. ASVs were aligned to a selection of published *Isoëtes LFY* sequences (Hoot *et al.*, 2004; Taylor *et al.*, 2004; Kim *et al.*, 2010; Rosenthal & Rosenthal, 2014) using the G-INS-i algorithm in MAFFT V7.313 (Kato & Standley, 2013) and phylogenies were inferred using MrBayes (Ronquist & Huelsenbeck, 2003) using the parameters described above for chloroplast phylogenetic inference.

Results

We initially identified all specimens based on morphology (Pfeiffer, 1922; Britton & Brunton, 1993, 1996; Taylor, 2002), although some specimens could not be assigned to a named species. Our morphological identifications were compared to our molecular results (plastid, LFY nuclear marker, and RADseq data; discussed in detail below), and updated accordingly. Most specimens were ultimately identified as named taxa. However, we detected two taxa (with a combination of cytology, molecular data, and examination of spores) that did not match any named species. These are hereafter referred to as *Isoëtes species novo* A (tetraploid) and B (septaploid). Further morphological data are needed for formal taxonomic description.

Cytology

C-values were only obtained from specimens collected with fresh leaf tissue, for this reason genome size was not measured for every individual in this study (Table S1). Ploidy was inferred from C-values based on the relationship between chromosome number and C-value (Bolin *et al.*, 2017). For other individuals, chromosome counts have been made. As expected, C-values of the putative diploids, *I. echinospora* and *I. bolanderi* were the smallest. The values of all samples of *I. bolanderi* were between 2.4–2.6, while the value for *I. echinospora* was 3.9 (Table S1). Specimens of *I. echinospora* from eastern North America are also known to have a larger C-value compared to other *Isoëtes* diploids, and the genome size of diploid *Isoëtes* species

can vary (Bolin *et al.*, 2017). The values of the tetraploids *I. maritima* as well as *I. sp. nov.* A are between 5.7–6.9, and the values of the hybrid between *I. occidentalis* and *I. bolanderi* are between 5.7–6.9 (Table S1). C-values of the hexaploid *I. occidentalis* ranged from 9.4–9.8, and values for *I. sp. nov.* B (7x) had a value between 9.9–10.8 (Table S1). Generally, there is a tight correlation between C-value and ploidy, but within *Isoetes* this relationship starts to break down at higher ploidy levels (Bolin *et al.*, 2017). Because our lower ploidy level taxa have relatively high C-values, the high values of the tetraploids, hexaploids, and septaploids are mathematically consistent. For instance, as *I. maritima* (average C-value: 6.2) is a hybrid between *I. echinospora* (3.9) and *I. bolanderi* (average C-value: 2.5), simple addition of the genomic size of these two taxa corroborates the larger genome size found in our polyploids.

Population genomic analysis

Our raw data obtained from the University of Wisconsin-Madison Biotechnology Center consisted of an average of 2.33×10^6 raw reads per sample. In the final assembly, we retained an average of 2,980 loci per sample, with a standard deviation of 858 loci per sample. Our STRUCTURE and SplitsTree analyses utilized a dataset of 3,343 SNPs, composed of one SNP per locus in the final assembly. Raw, demultiplexed sequences can be accessed from the NCBI GenBank Sequence Read Archive (PRJNA665237).

To identify the optimal number of populations (K) in our population genomic analysis, we used the best K method (Evanno *et al.*, 2005), as well as visually examining $K = 2-5$ as recommended by Pritchard *et al.* (2000). The best K method indicated a value of 3 as the best fit for our data (Fig. S1), which agreed with our prior morphological and cytological hypotheses on the genomic makeup of these taxa (Fig. 2b). This was also the most biologically meaningful number of groups because it separated each named diploid, and elucidated parentage in the allopolyploid hybrids. Decreasing K below 3 leads to a loss in information and increasing it only adds noise and does not add any meaningful population clusters (Fig. S1). Diploids identified as *I. bolanderi* and *I. echinospora* fall into respective single population groups (Fig. 2b). Putative hybrids including *I. × herb-wagneri* and *I. maritima* cluster together with half of their genomic constitution coming from *I. echinospora* and *I. bolanderi*. *Isoetes × herb-wagneri* is a sterile diploid hybrid, while *I. maritima* is a fertile allotetraploid. Based on population structure analyses, the genomic constitution of *I. maritima* has slightly more *I. echinospora* ancestry than

I. bolanderi (Fig. **2b**), but this could be accounted for by sequencing or analysis errors in the polyploid taxon.

All samples of the fertile hexaploid *I. occidentalis* clustered together as a distinct population, without signatures of any other species (Fig. **2b**). We recovered a sterile hybrid cross between *I. occidentalis* and *I. bolanderi* (*I. sp. nov. A*). Based on cytology this taxon is likely to be a tetraploid. There was an additional septaploid hybrid between *I. occidentalis* and *I. bolanderi* (*I. sp. nov. B*) (Fig. **2b**).

Our split network analysis showed very similar groups as our population structure analysis. *Isoëtes echinospora* and *I. occidentalis* fall into two separate, but closely-related groups; *Isoëtes bolanderi* is also placed as a distinct group, but farther away from the latter two species (Fig. **S2**). *Isoëtes* × *herb-wagneri* and *I. maritima* sit very close to one another, further substantiating *I. maritima* as an autopolyploid derivative of *I. × herb-wagneri*. These two species are placed in between their hypothesized progenitors, and the boxes created by the split network indicate some conflicting evolutionary signals (i.e., hybridization). The allopolyploid taxa *I. sp. nov. A* is placed in between *I. bolanderi* and *I. occidentalis*. *Isoëtes sp. nov. B* appears both next to the former hybrid, and with *I. occidentalis*. Due to its similarity to the population structure analysis (Fig. **S2**).

Determination of spore viability

To determine whether a taxon is fertile or sterile, and thus whether it is of hybrid origin, megaspores were visualized. For each mature individual, megaspores were visualized under a dissecting scope. We followed Wagner *et al.* (1986) to identify whether spores are spherical, uniform in size and ornamentation indicating viability and fertility, or flattened, uneven in size and ornamentation indicating non-viable and sterility (Wagner *et al.*, 1986). We determined that most of the spores from *I. sp. nov. A*, *I. sp. nov. B*, and *I. × herb-wagneri* were variable in shape and size, suggesting sterility, while spores from the diploid species (*I. echinospora* and *I. bolanderi*), the tetraploid *I. maritima*, and the hexaploid *I. occidentalis* were uniform in size and shape, suggesting fertility (Fig. **3a-g**).

Whole chloroplast genomes

In order to gauge the directionality of hybridization events whole plastome phylogenies

were reconstructed (Fig. 4). Bootstrap support as well as posterior probability were high (>95; >0.95) for almost every node in the plastome-derived phylogeny (Fig. 4). Three primary clades of interest were recovered: an *I. bolanderi* clade (blue); an *I. echinospora* clade (brown); and an *I. occidentalis* clade (green). The branch leading to each of these clades has a posterior probability and bootstrap support of 1.00 and 100% respectively. The *I. echinospora* and *I. occidentalis* clades fall out as sister groups to each other, while the *I. bolanderi* clade is sister group to *I. melanopoda* (Fig. 4). Two individuals of *I. maritima* appear in different places in the tree with one as sister to *I. echinospora*, while the other appears most closely related to *I. bolanderi*, suggesting that each diploid species can be the maternal parent. Individuals identified in the field as *I. × pseudotruncata*, as well as the hybrid *I. sp. nov. A* are most closely related to *I. occidentalis*. No diploid progenitor was found in the *I. occidentalis* clade. Similar to *I. maritima* two individuals of *I. × herb-wagneri* appear most closely related to either *I. bolanderi*, or *I. echinospora*, suggesting that either parent can be the maternal progenitor in those hybridization events (Fig. 4).

LFY phylogenetic inference

PacBio sequencing yielded 5684 CCS reads. Despite efforts to achieve equal sequencing depth, the number of reads per sample ranged from 10 to 926 (average 334, standard deviation 228). Following filtering, the number of reads per sample decreased by an average of 43%. DADA2 identified a total of 29 ASVs compared to the 39 OTU clusters created with PURC. Two examples of disagreement in the number of ASVs recovered from five pairs of PCR replicates suggests that PCR or sequencing bias may still affect results, although one of these disagreements may be explained by very low coverage of one replicate (number of filtered reads = 3).

These single gene nuclear data provided the initial framework we used to examine this reticulate complex. Each taxon sampled using PacBio's RSII platform returned a single sequence or multiple sequences depending on the ploidy level for each taxon (i.e. diploids returned single sequences while allopolyploids returned multiple). The PURC pipeline occasionally outputs multiple clusters of the same sequence due to strictness of clustering parameters (Fig. S3). The output from the PURC pipeline shows multiple clusters for *I. × herb-wagneri*, *I. maritima*, multiple unnamed hybrids, and the two replicates of *I. occidentalis* (Fig. S3). The hypothesized

hybrid (*I. × herb-wagneri*) and allotetraploid (*I. maritima*) shows equal clustering of *LFY* alleles from both putative parent species (*I. bolanderi* and *I. echinospora*). The multiple unnamed taxa (denoted as ‘sp.’) are not yet determined taxonomically and will be dealt with in a later paper. Various individuals of *I. echinospora* are found in multiple places in the phylogeny, indicating either misidentification or analytical errors (Fig. S3). The node splitting *I. occidentalis* and *I. echinospora* has very low support (0.56) and most of the tree is not resolved (i.e., soft polytomy) (Fig. S3). However, this is expected when using a single gene to reconstruct a species tree with extremely closely related and reticulating taxa.

Discussion

Reticulate evolution is thought to be common in the genus *Isoetes* because of the frequent occurrence of sterile plants, various co-occurring cytotypes, and the high proportion of polyploid taxa (Taylor & Hickey, 1992; Troia *et al.*, 2016). Previous attempts to explain the patterns of hybridization and polyploidization have relied on few genetic markers (Hoot *et al.*, 2004; Kim *et al.*, 2010; Pereira *et al.*, 2018; Dai *et al.*, 2020), which do not accurately depict species relationships in recently diverged groups with low molecular divergence, introgression, and incomplete lineage sorting (Nelson *et al.*, 2020b). Using whole chloroplast and nuclear genomes, in conjunction with cytological and reproductive data we investigated the importance of hybridization and polyploidization within a complex of western North American *Isoetes*. We detected several hybrids and polyploids, their parental origin, and the directionality of hybridization events. Specifically, we find evidence for homoploid hybridization leading to the formation of sterile diploid taxa that undergo whole genome duplication resulting in fertile polyploid lineages (Fig. 2, 5, S2, S3). These allopolyploids can hybridize with their diploid progenitors, across ploidy levels, to form sterile interploid hybrids (Fig. 3, 5). Interestingly, the directionality of hybridization events is reciprocal between taxa of the same cytotype but in interploid hybrids, the higher ploidy taxon acts solely as the maternal donor (Fig. 4). Finally, we find evidence for a distinct hexaploid lineage without signatures of hybrid origin, raising questions about its origin (Fig. 2b, 4, S2).

Homoploid hybridization and allopolyploid speciation

The diploid species *Isoetes bolanderi* and *I. echinospora* co-occur throughout western

North America. We find that genomic constitution (Fig. **2b**, **S2**, **S3**), ploidy (Table **S1**), sterility (Fig. **3e**), and chloroplast inheritance (Fig. **4**), demonstrate that these species hybridize to form the sterile homoploid *I. × herb-wagneri* (Fig. **5**), initially hypothesized by Taylor (2002). Since *I. × herb-wagneri* is sterile (Fig. **3e**) it cannot reproduce sexually and, in the absence of asexual reproduction through rootstock division, each individual is a product of a new hybridization event. This means that *I. × herb-wagneri* cannot persist as a sexually reproducing independent lineage, a necessary step in homoploid hybrid speciation (Yakimowski & Rieseberg, 2014). Based on these data, it is clear that *Isoetes bolanderi* and *I. echinospora* are too evolutionary divergent to form fertile homoploid hybrids, and allopolyploidy must be a necessary step to restore fertility (Buggs *et al.*, 2009).

Based on overlapping distributions and similar morphology, *I. maritima* was historically circumscribed as a variety of *I. echinospora* (Löve, 1962). Cytological differences later led to the segregation of *I. maritima* as a distinct tetraploid species (Brunton & Britton, 1999). The formation of viable spores (Fig. **3**), tetraploidy (Table **S1**), chloroplast inheritance (Fig. **4**), and genomic constitution (Fig. **2b**, **S3**) demonstrate that *I. maritima* is the allotetraploid derivative of *I. × herb-wagneri* (Fig. **5**). The formation of *I. maritima* may be achieved in multiple ways including: the fusion of unreduced gametes in *I. × herb-wagneri*; the fusion of unreduced gametes from the diploid parents (*I. echinospora* and *I. bolanderi*); or through a triploid bridge (i.e., the fusion of a reduced gamete of one parent with an unreduced gamete of another, followed by a backcross of that triploid hybrid with a reduced gamete from one of the diploid parents) (Ramsey & Schemske, 2002; Husband, 2004). We seldom find abnormal spores in diploid species, but frequently find them in hybrids (Fig. **3**; JSS, WCT, and PWS, pers. obs), suggesting that the first scenario (two unreduced gametes of *I. × herb-wagneri*) may be most likely. The presence of *I. maritima* and the sterility of *I. × herb-wagneri* suggests that allopolyploid speciation is a more common mechanism of hybrid speciation, compared to homoploid hybrid speciation, within this complex.

High-ploidy hybrids

In the 1990s, two sterile *Isoetes* hybrids (*Isoetes × pseudotruncata* and *I. × truncata*) were documented in northwestern North America (Britton & Brunton, 1993, 1996; Britton *et al.*,

1999). *Isoëtes* $\times$ *pseudotruncata* (3x) was described as the hybrid of *I. maritima* $\times$ *I. echinospora*; and *I. truncata* (5x) was describe as the hybrid of *I. maritima* $\times$ *I. occidentalis* (Fig. 1). Our analyses do not recover any genomic signatures of these putative hybridization events described by Britton and Brunton (1993, 1996; Fig. 1, 2b). It is possible that we did not sample them. Alternatively, they may have gone extinct or their progenitors were misidentified. However, we did find signatures of two other hybrid taxa, a sterile tetraploid (*I. sp. nov. A*) and a sterile septaploid (*I. sp. nov. B*) (Fig. 2b; Table S1). Both have genomic signatures of *I. bolanderi* and *I. occidentalis* (Fig. 2b, S2, S3), and they have polymorphic spores, indicating hybrid sterility (Fig. 3f, g). We suspect that the tetraploid *I. sp. nov. A* is a hybrid between a reduced gamete of *I. occidentalis* (3x) and a reduced gamete of *I. bolanderi* (1x) (Fig. 5); while the septaploid *I. sp. nov. B* is either the backcross between an unreduced gamete of *I. sp. nov. A* (4x) and *I. occidentalis* (3x) (Fig. 5), or the hybrid product between an unreduced gamete of *I. occidentalis* (6x) and a reduced gamete of *I. bolanderi* (1x).

Abnormal chromosomal behavior is common among interploid hybrids (Sigel, 2016). Chromosomal non-disjunction and uneven chromosomal segmentation can occur during meiosis leading to spores of various ploidy (Sybenga, 1996). Therefore, it is possible that these two interploid hybrids (*I. sp. nov. A* and *I. sp. nov. B*) could produce meiotic products with various cytotypes, resulting in pentaploid and triploid taxa like *I. truncata* and *I. pseudotruncata*, respectively. Given this phenomenon, it is possible that *I. truncata* and *I. pseudotruncata* are not derived from their initially proposed hybridization events (Fig. 1; Britton & Brunton, 1993, 1996), but rather represent multiple cytotypes derived from abnormal chromosomal behavior (Fig. 5). Other lines of evidence may support this, specifically the various spore sizes in these taxa (some large, round and seemingly viable; Fig. 3f, g), and observation that the chloroplast genome of a sterile triploid initially identified as “*I. pseudotruncata*” (Fig. 4; Taylor 6988) is actually derived from *I. occidentalis* (Fig. 4). It would be unlikely for this taxon to inherit its chloroplast from *I. occidentalis* if it were a hybrid between *I. maritima* and *I. echinospora*, as initially hypothesized (Fig. 1; Britton & Brunton, 1993). However, more data including DNA sequences from type species are needed to conclusively determine the identity of *I. pseudotruncata* and *I. truncata*.

These interploid hybrids occur in mixed populations in roughly equal (or lower) densities than their progenitors (JSS, WCT, pers. obs) and also hybridize with them, yet still persist in the population. The minority cytotype exclusion theory postulates that less common cytotypes within a population will be outcompeted by the majority cytotype, unless they have different ecological strategies (Levin, 1975). It is possible that interploid *Isoëtes* fill a disparate niche from their progenitors, or perhaps they are so novel that such evolutionary dynamics are still unfolding, and in the future these hybrids may naturally go extinct.

Directionality of hybridization events

Because *Isoëtes* is heterosporous, both micro- and mega-gametophytes must be present for fertilization to occur. Both male and female gametophytes, however, often co-occur in these populations (Fig. 2a), so one of the main, but not exclusive (Schneller, 1981), prezygotic barriers is sperm size and archegonial neck size (Testo *et al.*, 2015). Chloroplast inheritance shows that different individuals of *I. × herb-wagneri* and *I. maritima* are either more closely related to *I. bolanderi* or *I. echinospora*, meaning either parent can be the maternal donor (Fig. 4). Since *I. echinospora* and *I. bolanderi* are diploids they should have similar sized organs (i.e., sperm, eggs, archegonia), making reciprocal hybridization possible (Haufler *et al.*, 1995; Sigel *et al.*, 2014). In contrast, the interploid hybrid *I. sp. nov. A* is derived from a hexaploid and a diploid (Fig. 2b, 5), with the hexaploid *I. occidentalis* as the sole maternal progenitor (Fig. 4). This is most likely due to sperm size of the hexaploid being larger than the sperm of the diploid (Testo *et al.*, 2015) and may not fit down the narrower archegonial neck of the *I. bolanderi* gametophyte.

The origin of Isoëtes occidentalis and patterns of biogeography

Individuals identified as *Isoëtes occidentalis* in this study comprise a widespread and genomically uniform taxon throughout western North America (Fig. 2a,b). Using preliminary morphological and cytological data, we tested the hypothesis that *I. occidentalis* was an allohexaploid derived from *I. maritima* and *I. echinospora* (Fig. 1). However, we do not find signatures of shared ancestry within the genome of *I. occidentalis* (Fig. 2b). Based on these results, we suggest that *Isoëtes occidentalis* may be either an autopolyploid or allopolyploid derived from extinct or unsampled parents. In our population structure analysis at $K = 2$, we find

that *I. occidentalis* has the same genomic constitution as *I. echinospora* (Fig. S1), and our chloroplast (Fig. 4) and nuclear data (Fig. S3) suggest they are each other's closest sampled relatives. This could mean that *I. occidentalis* is an ancient autopolyploid of *I. echinospora*; alternatively, *I. occidentalis* may be an ancient allopolyploid derived from *I. bolanderi* and *I. echinospora*, but enough time has elapsed to allow for genetic divergence and genomic rearrangement (Sigel, 2016), masking the genomic similarities between these taxa in the population structure analysis (Fig. 2b). Future analyses including more individuals from geographically diverse populations and investigating ancestral genomic signatures of introgression should be conducted to more firmly determine the origin of *I. occidentalis* and other high-ploidy species in *Isoetes*.

Given the distribution of *I. occidentalis* throughout the Sierra Nevada, Rocky Mountains, and Coast Range, glaciation likely played a role in its evolutionary and biogeographic history (Haufler *et al.*, 1995; Burnier *et al.*, 2009; Stein *et al.*, 2010; Sessa *et al.*, 2012a,b; Jorgensen & Barrington, 2020). Progenitors of *I. occidentalis* may have occurred throughout lakes in northwestern North America, and Quaternary glaciation cycles (Dyke & Prest, 1987; Hughes *et al.*, 1989; Mandryk *et al.*, 2001) could have led to extinction of the diploid parents, while the polyploid *I. occidentalis* survived in refugia like the 'ice-free corridor' in western Canada (Schweger, 1989; Dyke, 2005). If true, it is likely that *I. occidentalis* evaded extinction due to a wider distribution and fixed heterozygosity, which buffered populations from inbreeding depression during climatic fluctuations (Brochmann *et al.*, 2004).

Long-distance dispersal could be another aspect of gene flow and biogeographical patterns in these western North American *Isoetes*. Unlike most homosporous ferns, which disperse their spores via air currents (Barrington, 1993), spore dispersal in *Isoetes* is poorly understood because they are aquatic and produce large megaspores (200-1000µm), unable to be dispersed by wind. Anecdotal evidence suggests they can be dispersed by waterways, waterfowl, or even snails and earthworms (Duthie, 1929; Jermy, 1990; Taylor & Hickey, 1992; Larsén & Rydin, 2016), and recent work has documented spores in waterfowl fecal matter (Silva *et al.*, 2020), confirming previous observations. It is not inconceivable that sporangia, spores, or even whole plants are dispersed in the crops of these migratory animals (Les *et al.*, 2003). In this way, separate species that evolved in allopatry, with limited reproductive barriers, could be re-introduced by waterfowl, leading to hybridization. We suspect that long-distance dispersal

through waterfowl and glaciation cycles have also influenced biogeographical patterns of other *Isoëtes* species, and further work should aim to tease apart these processes. Our results demonstrate that hybridization and polyploidization are common and occur with ease in this complex of western North American *Isoëtes*. The occurrence of many polyploids and putative hybrids in the genus (Troia *et al.*, 2016) suggests that these processes are important in shaping extant *Isoëtes* diversity.

Acknowledgements

Thank you to William D. Pearse for thoughtful comments on the manuscript. We also thank Gabriel Johnson and undergraduate research assistance from Catawba College for help with laboratory procedures and C-value generation. We thank the University of Wisconsin Biotechnology Center DNA Sequencing Facility for providing DNA extraction, library prep, and DNA sequencing facilities and services. Parts of this project were funded by The Society of Systematic Biologists Graduate Student Research Award, awarded to JSS. SPK is funded by a National Science Foundation Graduate Research Fellowship.

Author contributions

Taylor and Zimmer initiated the project, Suissa and Taylor carried out fieldwork, Bolin determined C-values, Suissa and Schafran sequenced and analyzed plastid genomes and *LFY* sequences, Suissa and Kinoshita performed RADSeq data processing, Kinoshita performed split network and population genomic analyses, and all authors participated in writing the manuscript.

Data availability

All sequence data can be accessed from the NCBI Genbank Sequence read archive. Demultiplexed sequences from RADSeq: PRJNA665237.

References

Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Prjibelski AD, *et al.* 2012. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *Journal of Computational Biology: a journal of computational molecular cell biology* **19**: 455–477.

- 585 **Barrington DS. 1993.** Ecological and historical factors in fern biogeography. *Journal of Biogeography*
586 **20:** 275–279.
- 587 **Barrington DS, Haufler CH, Werth CR. 1989.** Hybridization, Reticulation, and Species Concepts in the
588 Ferns. *American Fern Journal* **79:** 55–64.
- 589 **Bolin JF, Hartwig CL, Schafran P, Komarnytsky S. 2017.** Application of DNA Flow Cytometry to
590 Aid Species Delimitation in *Isoetes*. *Castanea* **83:** 38–47.
- 591 **Britton DM, Brunton DF. 1989.** A new *Isoetes* hybrid (*Isoetes echinospora* × *riparia*) for Canada.
592 *Canadian journal of botany. Journal Canadien de Bbotanique* **67:** 2995–3002.
- 593 **Britton DM, Brunton DF. 1993.** *Isoetes* × *truncata*: a newly considered pentaploid hybrid from western
594 North America. *Canadian journal of botany. Journal Canadien de Botanique* **71:** 1016–1025.
- 595 **Britton DM, Brunton DF. 1996.** *Isoetes* × *pseudotruncata*, a new triploid hybrid from western Canada
596 and Alaska. *Canadian journal of botany. Journal Canadien de Botanique* **74:** 51–59.
- 597 **Britton DM, Brunton DF, Talbot SS. 1999.** *Isoetes* in Alaska and the Aleutians. *American Fern Journal*
598 **89:** 133–141.
- 599 **Brochmann C, Brysting AK, Alsos IG, Borgen L, Grundt HH, Scheen A-C, Elven R. 2004.**
600 Polyploidy in arctic plants. *Biological Journal of the Linnean Society* **82:** 521–536.
- 601 **Brunton DF, Britton DM. 1999.** Maritime quillwort, *Isoetes maritima* (Isoetaceae), in the Yukon
602 territory. *Canadian Field-Naturalist* **113:** 641–645.
- 603 **Brunton DF, Britton DM. 2000.** *Isoetes* × *echtuckerii*, hyb. nov., a new triploid quillwort from
604 northeastern North America. *Canadian Journal of Botany* **77:** 1662–1668.
- 605 **Bryant D, Moulton V. 2002.** NeighborNet: An Agglomerative Method for the Construction of Planar
606 Phylogenetic Networks. In: Algorithms in Bioinformatics. Springer Berlin Heidelberg, 375–391.
- 607 **Buggs RJA, Soltis PS, Soltis DE. 2009.** Does hybridization between divergent progenitors drive whole-
608 genome duplication? *Molecular Ecology* **18:** 3334–3339.
- 609 **Burnier J, Buerki S, Arrigo N, Küpfer P, Alvarez N. 2009.** Genetic structure and evolution of Alpine
610 polyploid complexes: *Ranunculus kuepferi* (Ranunculaceae) as a case study. *Molecular Ecology* **18:**
611 3730–3744.
- 612 **Callahan BJ, McMurdie PJ, Rosen MJ, Han AW, Johnson AJA, Holmes SP. 2016.** DADA2: High-
613 resolution sample inference from Illumina amplicon data. *Nature Methods* **13:** 581–583.
- 614 **Catchen JM, Amores A, Hohenlohe P, Cresko W, Postlethwait JH. 2011.** Stacks: building and
615 genotyping Loci *de novo* from short-read sequences. *G3* **1:** 171–182.
- 616 **Catchen J, Hohenlohe PA, Bassham S, Amores A, Cresko WA. 2013.** Stacks: an analysis tool set for
617 population genomics. *Molecular Ecology* **22:** 3124–3140.
- 618 **Clute, N. W. 1905.** What constitutes a species in the genus *Isoëtes*. *Fern Bulletin* **13:** 41–47
- 619 **Dai X, Li X, Huang Y, Liu X. 2020.** The speciation and adaptation of the polyploids: a case study of the
620 Chinese *Isoetes* L. diploid-polyploid complex. *BMC Evolutionary Biology* **20:** 118.

- 621 **Dauphin B, Grant JR, Farrar DR, Rothfels CJ. 2018.** Rapid allopolyploid radiation of moonwort ferns
622 (Botrychium; Ophioglossaceae) revealed by PacBio sequencing of homologous and homeologous nuclear
623 regions. *Molecular Phylogenetics and Evolution* **120**: 342–353.
- 624 **Degnan JH, Rosenberg NA. 2009.** Gene tree discordance, phylogenetic inference and the multispecies
625 coalescent. *Trends in Ecology & Evolution* **24**: 332–340.
- 626 **Dobzhansky T. 1982.** *Genetics and the Origin of Species*. Columbia University Press.
- 627 **Dolezel J, Bartos J. 2005.** Plant DNA flow cytometry and estimation of nuclear genome size. *Annals of*
628 *Botany* **95**: 99–110.
- 629 **Doležel J, Doleželová M, Novák FJ. 1994.** Flow cytometric estimation of nuclear DNA amount in
630 diploid bananas (*Musa acuminata* and *M. balbisiana*). *Biologia Plantarum* **36**: 351.
- 631 **Duthie AV. 1929.** The method of spore dispersal of three South African species of *Isoetes*. *Annals of*
632 *Botany* **43**: 411–412.
- 633 **Dyke A. 2005.** Late Quaternary vegetation history of northern North America based on pollen,
634 macrofossil, and faunal remains. *Géographie Physique et Quaternaire* **59**: 211–262.
- 635 **Dyke AS, Prest VK. 1987.** *Paleogeography of northern North America, 18,000-5,000 years ago*.
636 Geological Survey of Canada.
- 637 **Eaton DAR, Overcast I. 2020.** ipyrad: Interactive assembly and analysis of RADseq datasets.
638 *Bioinformatics* **36**: 2592–2594.
- 639 **Edelman NB, Frandsen PB, Miyagi M, Clavijo B. 2019.** Genomic architecture and introgression shape
640 a butterfly radiation. *Science* **366**: 594–599.
- 641 **Elshire RJ, Glaubitz JC, Sun Q, Poland JA, Kawamoto K, Buckler ES, Mitchell SE. 2011.** A robust,
642 simple genotyping-by-sequencing (GBS) approach for high diversity species. *PloS One* **6**: e19379.
- 643 **Evanno G, Regnaut S, Goudet J. 2005.** Detecting the number of clusters of individuals using the
644 software STRUCTURE: a simulation study. *Molecular Ecology* **14**: 2611–2620.
- 645 **Grant V. 1971.** *Plant Speciation*. Columbia University Press.
- 646 **Haufler CH, Pryer KM, Schuettpelz E, Sessa EB, Farrar DR, Moran R, Schneller JJ, Watkins JE,
647 Windham MD. 2016.** Sex and the Single Gametophyte: Revising the Homosporous Vascular Plant Life
648 Cycle in Light of Contemporary Research. *Bioscience* **66**: 928–937.
- 649 **Haufler CH, Windham MD, Rabe EW. 1995.** Reticulate Evolution in the *Polypodium vulgare*
650 Complex. *Systematic Botany* **20**: 89–109.
- 651 **Hickey RJ, Taylor WC, Luebke NT. 1989.** The species concept in Pteridophyta with special reference
652 to *Isoetes*. *American Fern Journal* **79**: 78–89.
- 653 **Hoot SB, Napier NS, Taylor WC. 2004.** Revealing unknown or extinct lineages within *Isoetes*
654 (Isoetaceae) using DNA sequences from hybrids. *American Journal of Botany* **91**: 899–904.
- 655 **Hughes OL, Rutter NW, Clague JJ, Fulton RJ. 1989.** Yukon Territory (Quaternary stratigraphy and
656 history, Cordilleran Ice Sheet). *Chapter 1*: 58–62.

- 657 **Husband BC. 2004.** The role of triploid hybrids in the evolutionary dynamics of mixed-ploidy
658 populations. *Biological Journal of the Linnean Society* **82**: 537–546.
- 659 **Husband BC, Ozimec B, Martin SL, Pollock L. 2008.** Mating Consequences of Polyploid Evolution in
660 Flowering Plants: Current Trends and Insights from Synthetic Polyploids. *International Journal of Plant*
661 *Sciences* **169**: 195–206.
- 662 **Huson DH, Bryant D. 2006.** Application of phylogenetic networks in evolutionary studies. *Molecular*
663 *Biology and Evolution* **23**: 254–267.
- 664 **Jermey AC. 1990.** Isoetaceae. In: Kramer KU, Green PS, eds. Pteridophytes and Gymnosperms. Berlin,
665 Heidelberg: Springer Berlin Heidelberg, 26–31.
- 666 **Jorgensen SA, Barrington DS. 2020.** Two Beringian Origins for the Allotetraploid Fern *Polystichum*
667 *braunii* (Dryopteridaceae). *Systematic Botany* **42**: 6–16.
- 668 **Karol KG, Arumuganathan K, Boore JL, Duffy AM, Everett KDE, Hall JD, Hansen SK, Kuehl JV,**
669 **Mandoli DF, Mishler BD, et al. 2010.** Complete plastome sequences of *Equisetum arvense* and *Isoetes*
670 *flaccida*: implications for phylogeny and plastid genome evolution of early land plant lineages. *BMC*
671 *Evolutionary Biology* **10**: 321.
- 672 **Katoh K, Standley DM. 2013.** MAFFT multiple sequence alignment software version 7: improvements
673 in performance and usability. *Molecular Biology and Evolution* **30**: 772–780.
- 674 **Kearse M, Moir R, Wilson A, Stones-Havas S, Cheung M, Sturrock S, Buxton S, Cooper A,**
675 **Markowitz S, Duran C, et al. 2012.** Geneious Basic: an integrated and extendable desktop software
676 platform for the organization and analysis of sequence data. *Bioinformatics* **28**: 1647–1649.
- 677 **Kim C, Shin H, Chang Y-T, Choi H-K. 2010.** Speciation pathway of *Isoetes* (Isoëtaceae) in East Asia
678 inferred from molecular phylogenetic relationships. *American Journal of Botany* **97**: 958–969.
- 679 **Kopelman NM, Mayzel J, Jakobsson M, Rosenberg NA, Mayrose I. 2015.** Clumpak: a program for
680 identifying clustering modes and packaging population structure inferences across K. *Molecular Ecology*
681 *Resources* **15**: 1179–1191.
- 682 **Landis JB, Soltis DE, Li Z, Marx HE, Barker MS, Tank DC, Soltis PS. 2018.** Impact of whole-
683 genome duplication events on diversification rates in angiosperms. *American Journal of Botany* **105**:
684 348–363.
- 685 **Langmead B, Salzberg SL. 2012.** Fast gapped-read alignment with Bowtie 2. *Nature methods* **9**: 357–
686 359.
- 687 **Larsén E, Rydin C. 2016.** Disentangling the Phylogeny of *Isoetes* (Isoetales), Using Nuclear and Plastid
688 Data. *International Journal of Plant Sciences* **177**: 157–174.
- 689 **Les DH, Crawford DJ, Kimball RT, Moody ML, Landolt E. 2003.** Biogeography of Discontinuously
690 Distributed Hydrophytes: A Molecular Appraisal of Intercontinental Disjunctions. *International Journal*
691 *of Plant Sciences* **164**: 917–932.
- 692 **Levin DA. 1975.** Minority cytotype exclusion in local plant populations. *Taxon* **24**: 35–43.
- 693 **Linder CR, Rieseberg LH. 2004.** Reconstructing patterns of reticulate evolution in plants. *American*
694 *Journal of Botany* **91**: 1700–1708.

- 695 **Löve Á. 1962.** Cytotaxonomy of the *Isoetes echinospora* Complex. *American Fern Journal* **52**: 113–123.
- 696 **Mandryk CAS, Josenhans H, Fedje DW. 2001.** Late Quaternary paleoenvironments of Northwestern
697 North America: implications for inland versus coastal migration routes. *Quaternary Science Reviews* **20**:
698 301–314.
- 699 **Martin M. 2011.** Cutadapt removes adapter sequences from high-throughput sequencing reads.
700 *EMBnet.journal* **17**: 10–12.
- 701 **Mayrose I, Zhan SH, Rothfels CJ, Arrigo N, Barker MS, Rieseberg LH, Otto SP. 2015.** Methods for
702 studying polyploid diversification and the dead end hypothesis: a reply to Soltis et al. (2014). *The New*
703 *Phytologist* **206**: 27–35.
- 704 **Mayrose I, Zhan SH, Rothfels CJ, Magnuson-Ford K, Barker MS, Rieseberg LH, Otto SP. 2011.**
705 Recently formed polyploid plants diversify at lower rates. *Science* **333**: 1257.
- 706 **Nakhleh L, Warnow T, Linder CR. 2005.** Reconstructing reticulate evolution in species—theory and
707 practice. *Journal of Computational* **12**: 796–811.
- 708 **Nauheimer L, Cui L, Clarke C, Crayn DM, Bourke G, Nargar K. 2019.** Genome skimming provides
709 well resolved plastid and nuclear phylogenies, showing patterns of deep reticulate evolution in the
710 tropical carnivorous plant genus *Nepenthes* (Caryophyllales). *Australian Systematic Botany* **32**: 243–254.
- 711 **Nelson JM, Hauser DA, Li FW. 2020a.** Symbiotic cyanobacteria communities in hornworts across time,
712 space, and host species. *bioRxiv*: 10.1101/2020.06.18.160382
- 713 **Nelson TC, Stathos AM, Vanderpool DD, Finseth FR, Yuan Y-W, Fishman L. 2020b.** Ancient and
714 recent introgression shape the evolutionary history of pollinator adaptation and speciation in a model
715 monkeyflower radiation (*Mimulus* section *Erythranthe*). *bioRxiv*: 10.1101/2020.09.08.287151.
- 716 **Pereira J, Labiak PH, Stützel T. 2017.** Origin and biogeography of the ancient genus *Isoetes* with focus
717 on the Neotropics. *Botanical journal of the Linnean Society. Linnean Society of London* **185**: 253–271.
- 718 **Pereira J, Labiak PH, Stützel T, Schulz C. 2018.** Nuclear multi-locus phylogenetic inferences of
719 polyploid *Isoetes* species (Isoëtaceae) suggest several unknown diploid progenitors and a new polyploid
720 species from South America. *Botanical journal of the Linnean Society. Linnean Society of London* **189**:
721 6–22.
- 722 **Pfeiffer NE. 1922.** Monograph of the Isoetaceae. *Annals of the Missouri Botanical Garden. Missouri*
723 *Botanical Garden* **9**: 79–233.
- 724 **Pigg KB. 1992.** Evolution of Isoetalean lycopsids. *Annals of the Missouri Botanical Garden. Missouri*
725 *Botanical Garden* **79**: 589–612.
- 726 **Pritchard JK, Stephens M, Donnelly P. 2000.** Inference of population structure using multilocus
727 genotype data. *Genetics* **155**: 945–959.
- 728 **Rambaut A, Drummond AJ, Xie D, Baele G, Suchard MA. 2018.** Posterior Summarization in
729 Bayesian Phylogenetics Using Tracer 1.7. *Systematic biology* **67**: 901–904.
- 730 **Ramsey J, Schemske DW. 2002.** Neopolyploidy in Flowering Plants. *Annual Review of Ecology and*
731 *Systematics* **33**: 589–639.

- 732 **Rieseberg LH. 1997.** Hybrid Origins of Plant Species. *Annual Review of Ecology and Systematics* **28**:
733 359–389.
- 734 **Ronquist F, Huelsenbeck JP. 2003.** MrBayes 3: Bayesian phylogenetic inference under mixed models.
735 *Bioinformatics* **19**: 1572–1574.
- 736 **Rosenthal MA, Rosenthal SR. 2014.** *Isoetes viridimontana*: a previously unrecognized quillwort from
737 Vermont, USA. *American Fern Journal* **104**: 7–15.
- 738 **Rothfels CJ, Pryer KM, Li F-W. 2017.** Next-generation polyploid phylogenetics: Rapid resolution of
739 hybrid polyploid complexes using PacBio single-molecule sequencing. *The New Phytologist* **213**: 413–
740 429.
- 741 **Sandstedt GD, Wu CA, Sweigart AL. 2020.** Evolution of multiple postzygotic barriers between species
742 of the *Mimulus tilingii* complex. *BioRxiv*: 2020.08.07.241489.
- 743 **Schafraan PW, Zimmer EA, Taylor WC, Musselman LJ. 2018.** A whole chloroplast genome
744 phylogeny of diploid species of Isoetes (Isoetaceae, Lycopodiophyta) in the southeastern United States.
745 *Castanea* **83**: 224–235.
- 746 **Schneller JJ. 1981.** Evidence for intergeneric incompatibility in ferns. *Plant Systematics and Evolution*
747 **137**: 45–56.
- 748 **Schweger CE. 1989.** Paleoecology of the western Canadian ice-free corridor. *Quaternary geology of the*
749 *Canadian Cordillera; Quaternary geology of Canada and Greenland. Edited by RJ Fulton. Geological*
750 *Survey of Canada, Geology of Canada* **1**: 491–498.
- 751 **Sessa EB, Zimmer EA, Givnish TJ. 2012a.** Unraveling reticulate evolution in North American
752 *Dryopteris* (Dryopteridaceae). *BMC Evolutionary Biology* **12**: 104.
- 753 **Sessa EB, Zimmer EA, Givnish TJ. 2012b.** Reticulate evolution on a global scale: a nuclear phylogeny
754 for New World *Dryopteris* (Dryopteridaceae). *Molecular Phylogenetics and Evolution* **64**: 563–581.
- 755 **Sigel EM. 2016.** Genetic and genomic aspects of hybridization in ferns. *Journal of Systematics and*
756 *Evolution* **54**: 638–655.
- 757 **Sigel EM, Windham MD, Pryer KM. 2014.** Evidence for reciprocal origins in *Polypodium hesperium*
758 (Polypodiaceae): a fern model system for investigating how multiple origins shape allopolyploid
759 genomes. *American Journal of Botany* **101**: 1476–1485.
- 760 **Silva GG, Green AJ, Hoffman P, Weber V. 2020.** Seed dispersal by neotropical waterfowl depends on
761 bird species and seasonality. *Freshwater Biology* **65**: 1–11.
- 762 **Small RL, Hickey RJ. 2001.** Systematics of the Northern Andean *Isoetes karstenii* Complex. *American*
763 *Fern Journal* **91**: 41–69.
- 764 **Soltis DE, Segovia-Salcedo MC, Jordon-Thaden I, Majure L, Miles NM, Mavrodiev EV, Mei W,**
765 **Cortez MB, Soltis PS, Gitzendanner MA. 2014.** Are polyploids really evolutionary dead-ends (again)?
766 A critical reappraisal of Mayrose et al.(2011). *The New Phytologist* **202**: 1105–1117.
- 767 **Soltis PS, Soltis DE. 2000.** The role of genetic and genomic attributes in the success of polyploids.
768 *Proceedings of the National Academy of Sciences of the United States of America* **97**: 7051–7057.

- 769 **Stamatakis A. 2006.** RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with
770 thousands of taxa and mixed models. *Bioinformatics* **22**: 2688–2690.
- 771 **Stebbins GL Jr. 1947.** Types of polyploids; their classification and significance. *Advances in Genetics* **1**:
772 403–429.
- 773 **Stebbins GL. 1969.** The significance of hybridization for plant taxonomy and evolution. *Taxon* **18**: 26–
774 35.
- 775 **Stein DB, Hutton C, Conant DS, Haufler CH, Werth CR. 2010.** Reconstructing *Dryopteris*
776 ‘*semicristata*’ (Dryopteridaceae): Molecular profiles of tetraploids verify their undiscovered diploid
777 ancestor. *American Journal of Botany* **97**: 998–1004.
- 778 **Sybenga J. 1996.** Chromosome pairing affinity and quadrivalent formation in polyploids: do segmental
779 allopolyploids exist? *Genome* **39**: 1176–1184.
- 780 **Taylor WC. 2002.** *Isoetes* × *herb-wagneri*, an interspecific hybrid of *I. bolanderi* × *I. echinospora*
781 (Isoëtaceae). *American Fern Journal* **92**: 161–163.
- 782 **Taylor WC, Hickey RJ. 1992.** Habitat, Evolution, and Speciation in *Isoetes*. *Annals of the Missouri*
783 *Botanical Garden. Missouri Botanical Garden* **79**: 613–622.
- 784 **Taylor WC, Lekschas AR, Wang QF, Liu X, Napier NS, Hoot SB. 2004.** Phylogenetic relationships of
785 *Isoetes* (Isoëtaceae) in China as revealed by nucleotide sequences of the nuclear ribosomal ITS region and
786 the second intron of a LEAFY homolog. *American Fern Journal* **94**: 196–205.
- 787 **Taylor WC, Luebke NT. 1988.** *Isoetes* × *hickeyi*: A naturally occurring hybrid between *I. echinospora*
788 and *I. macrospora*. *American Fern Journal* **78**: 6–13.
- 789 **Testo WL, Watkins JE Jr, Barrington DS. 2015.** Dynamics of asymmetrical hybridization in North
790 American wood ferns: reconciling patterns of inheritance with gametophyte reproductive biology. *The*
791 *New Phytologist* **206**: 785–795.
- 792 **Troia A, Pereira JB, Kim C, Taylor WC. 2016.** The genus *Isoetes* (Isoetaceae): a provisional checklist
793 of the accepted and unresolved taxa. *Phytotaxa* **277**: 101–145.
- 794 **Wagner WH Jr, Wagner FS. 1980.** Polyploidy in pteridophytes. *Basic Life Sciences* **13**: 199–214.
- 795 **Wagner WH, Wagner FS, Taylor WC. 1986.** Detecting Abortive Spores in Herbarium Specimens of
796 Sterile Hybrids. *American Fern Journal* **76**: 129–140.
- 797 **Wood D, Besnard G, Beerling DJ, Osborne CP, Christin P-A. 2020.** Phylogenomics indicates the
798 ‘living fossil’ *Isoetes* diversified in the Cenozoic. *PloS One* **15**: e0227525.
- 799 **Wood TE, Takebayashi N, Barker MS, Mayrose I, Greenspoon PB, Rieseberg LH. 2009.** The
800 frequency of polyploid speciation in vascular plants. *Proceedings of the National Academy of Sciences of*
801 *the United States of America* **106**: 13875–13879.
- 802 **Yakimowski SB, Rieseberg LH. 2014.** The role of homoploid hybridization in evolution: a century of
803 studies synthesizing genetics and ecology. *American Journal of Botany* **101**: 1247–1258.

804

Figure 1. Historically hypothesized relationships of western North American *Isoëtes*, derived from Britton & Brunton, 1993, 1996; Britton *et al.*, 1999; and Taylor, 2002. *Isoëtes bolanderi* and *I. echinospora* hybridize to form the sterile diploid *I. × herb-wagneri*. *Isoëtes maritima* is formed via a whole genome duplication (WGD; indicated by red dot) of *I. × herb-wagneri*. Then, *I. maritima* (2x gamete) hybridizes with *I. echinospora* (1x gamete) to form triploid *I. × pseudotruncata*. The triploid goes through a WGD (red dot) to form *I. occidentalis*. Finally, *I. maritima* (2x gamete) and *I. occidentalis* (3x gamete) hybridize to form the pentaploid *I. truncata*.

Figure 2. Sampling locations and population structure ($K = 3$) of western North American *Isoëtes*; (a) black dots represent lake localities where species were collected, pie charts indicate the ploidy and genomic structure of the different individuals collected at each location. The ploidy was derived from c-values or chromosome counts and the genomic constitution derived from the STRUCTURE analysis; (b) population structure and ploidy level for each sample; the diploid species (*I. bolanderi*, blue; *I. echinospora*, brown) and hexaploid *I. occidentalis* (green) are separated as the three main species with distinct genomic populations.

Figure 3. Spore images for (a) fertile diploid *Isoëtes echinospora*; (b) fertile diploid *I. bolanderi*; (c) fertile hexaploid *I. occidentalis*; (d) fertile tetraploid *I. maritima*; (e) sterile diploid *I. × herb-wagneri*; (f) sterile tetraploid *I. sp. nov. A*; and (g) sterile septaploid *I. sp. nov. B*. (scale bar: 200µm). Fertile spores were round and uniform within the sample (i.e., a, b, c, d). Sterile spores are misshapen (sometimes flattened: e) and vary in size sometimes with the meiotic tetrads visible (i.e., f, g). Central image is of *I. occidentalis* in the field.

Figure 4. Plastome phylogeny of western North American *Isoëtes* including outgroups for reference. All individuals in the group of interest comprise three distinct clades: *I. bolanderi* (blue), *I. echinospora* (brown), *I. occidentalis* (green). Within the blue clade, there are accessions of *I. bolanderi*, *I. × herb-wagneri*, and *I. maritima*. In the green clade, there are accessions of *I. occidentalis* and *I. sp. nov. A*. Finally, in the brown clade there are accessions of *I. echinospora*, *I. × herb-wagneri*, and *I. maritima*. The taxon labeled “*I. pseduotruncata*” Taylor 6988, was initially hypothesized to be *I. pseduotruncata*, however our data suggests this is *I. sp. nov. A*.

Figure 5. Hypothesized relationships of western North American *Isoëtes* as supported by this study. *Isoëtes bolanderi* and *I. echinospora* hybridize to form the sterile diploid *I. × herb-wagneri*. *Isoëtes maritima* is formed via a whole genome duplication (indicated by red dot) of *I. × herb-wagneri*. *Isoëtes bolanderi* (1x gamete) and *I. occidentalis* (3x gamete) hybridize to form the tetraploid *I. sp. nov. A*. Finally, *I. occidentalis* (3x gamete) and *I. sp. nov. A*. (4x gamete) hybridize to form the septaploid *I. sp. nov. B*. The polyploid origin of *I. occidentalis* is still inconclusive.