BioOne COMPLETE

A Whole Chloroplast Genome Phylogeny of Diploid Species of Isoëtes (Isoëtaceae, Lycopodiophyta) in the Southeastern United States

Authors: Schafran, Peter W., Zimmer, Elizabeth A., Taylor, W. Carl, and Musselman, Lytton J.

Source: Castanea, 83(2) : 224-235

Published By: Southern Appalachian Botanical Society

URL: https://doi.org/10.2179/17-132

BioOne Complete (complete.BioOne.org) is a full-text database of 200 subscribed and open-access titles in the biological, ecological, and environmental sciences published by nonprofit societies, associations, museums, institutions, and presses.

Your use of this PDF, the BioOne Complete website, and all posted and associated content indicates your acceptance of BioOne's Terms of Use, available at www.bioone.org/terms-of-use.

Usage of BioOne Complete content is strictly limited to personal, educational, and non - commercial use. Commercial inquiries or rights and permissions requests should be directed to the individual publisher as copyright holder.

BioOne sees sustainable scholarly publishing as an inherently collaborative enterprise connecting authors, nonprofit publishers, academic institutions, research libraries, and research funders in the common goal of maximizing access to critical research.

Downloaded From: https://bioone.org/journals/Castanea on 29 Mar 2024
Terms of Use: https://bioone.org/terms-of-use

CASTANEA 83(2): 224–235. JULY
Copyright 2018 Southern Appalachian Botanical Society

A Whole Chloroplast Genome Phylogeny of Diploid Species of *Isoëtes* (Isoëtaceae, Lycopodiophyta) in the Southeastern United States

Peter W. Schafran,[1,2*] Elizabeth A. Zimmer,[2] W. Carl Taylor,[2] and Lytton J. Musselman[1]
[1]Department of Biological Sciences, Old Dominion University, Norfolk, Virginia 23529
[2]Department of Botany, National Museum of Natural History, Smithsonian Institution, Washington, D.C. 20013

ABSTRACT The phylogeny of closely related species of *Isoëtes* has been difficult to infer using morphological or molecular data. The southeastern United States exhibits among the highest diversity of *Isoëtes* species, which likely are the result of a relatively recent radiation. We used whole chloroplast genome (plastome) sequences to infer a phylogeny of diploid taxa in the Southeast. The phylogeny of the entire plastome was well supported across various models of evolution. Separate phylogenetic analyses of coding regions and introns mostly returned the same tree topology as the whole plastome, although the placement of four species was discordant. The plastome phylogeny highlights taxonomic issues, because two pairs of infrageneric taxa are resolved as polyphyletic. Inference of ancestral character states suggests a common ancestor of the southeastern clade with white, tuberculate megaspores and papillose microspores maturing in the spring. This study is the first to resolve molecular-phylogenetic relationships in diploid *Isoëtes* from the southeastern USA and shows the utility of plastome sequences for unraveling relationships between closely related taxa.

Key words: *Isoëtes*, lycophyte, phylogenomics, plastome, quillwort.

INTRODUCTION

Isoëtes (Isoëtaceae, Lycopodiophyta) is a cosmopolitan genus of about 200 described species (Troia et al. 2016). Their common names, "quillwort" and "Merlin's grass," originate from their morphology, generally consisting of many linear, subterete sporophylls (leaves) borne on a subterranean rootstock. All species of *Isoëtes* reproduce through production of mega- and microspores in sporangia on separate sporophylls (occasionally separate plants, e.g., *I. butleri*).

The evolutionary history of *Isoëtes* is notoriously difficult to infer. Early attempts were hampered by the dearth of variable character states between species and the phenotypic plasticity that occurs within some characters (Hickey 1986, Taylor and Hickey 1992). Some subgeneric classifications were proposed such as subgenus *Euphyllum*, characterized by alate leaves, and subgenus *Isoëtes*, comprised of species with nonalate leaves (Hickey 1990). Within subgenus *Isoëtes*, sections *Coromandelina* and *Isoëtes* were proposed (Taylor and Hickey 1992). Despite fairly strong morphological phylogenetic hypotheses, none of these subgeneric classifications were supported by molecular data (Hoot and Taylor 2001, Rydin and Wikström 2002, Hoot et al. 2006, Schuettpelz and Hoot 2006, Larsén and Rydin 2016). This suggests numerous changes and reversions in character states and habitat preference, hindering any phylogenetic inference based on nonmolecular data (Taylor and Hickey 1992).

Among molecular phylogenies of species in the genus, the clade containing taxa of the southeastern United States is consistently diffi-

*email addresses: pscha005@odu.edu, schafranp@si.edu
Received March 9, 2017; Accepted May 17, 2018.
Published: August 30, 2018.
DOI: 10.2179/17-132

Downloaded From: https://bioone.org/journals/Castanea on 29 Mar 2024
Terms of Use: https://bioone.org/terms-of-use

cult to resolve (Hoot and Taylor 2001, Hoot et al. 2006, Larsén and Rydin 2016). Currently, about 25 taxa are recognized in the region comprising Delaware, Maryland, Virginia, West Virginia, Kentucky, Tennessee, North Carolina, South Carolina, Georgia, Florida, Alabama, Mississippi, Arkansas, Louisiana, Oklahoma, and Texas (Brunton 2015). Of these, 15 are thought to be basic diploids, hypothesized to have evolved through vicariance. The remaining taxa are hybrids and allopolyploids ranging in ploidy level from diploid to octoploid (Brunton 2015). One of the shortcomings of previous phylogenies is the use of few markers with low phylogenetic signal sufficient to discriminate among the southeastern US species. Common markers, plastid *rbcL* and the *atpB-rbcL* spacer, as well as the nuclear ribosomal internal transcribed spacer (ITS), cannot resolve relationships within the southeastern American clade (Rydin and Wikström 2002, Hoot et al. 2006, Larsén and Rydin 2016).

The plastid genome (plastome) has long been a source of phylogenetic markers for plant systematics. Studies have employed coding and noncoding regions of the plastome to evaluate the relationships from the infraspecific level to the backbone of all multicellular plants (Shaw et al. 2005, 2007, 2014). With a high selective pressure on the photometabolic genes encoded by the plastome, its nucleotide sequence and structure are relatively slowly evolving (Wicke and Schneeweiss 2015). Protein coding regions show greater conservation than noncoding introns and intergenic spacers, but rates of mutation vary across these categories and across lineages (Wicke and Schneeweiss 2015). Therefore, typical studies that use few ($\leq$6) chloroplast markers often cannot resolve low-level phylogenies (Shaw et al. 2014). With the increasing ease of whole plastome assembly from next generation sequencing data, however, phylogenomic studies can generate robust phylogenies at all taxonomic levels (Wicke and Schneeweiss 2015).

To infer the phylogeny of *Isoëtes* in the southeastern USA, we employ a diploids-first approach, wherein a phylogeny of the basic diploid taxa provides a framework to infer the parentage of hybrids and allopolyploids (Beck et al. 2010, Burgess et al. 2015). Here we present a phylogeny of southeastern diploid *Isoëtes* based on plastome data.

METHODS

Sample Collection

Plants were collected from type localities whenever possible (Table 1). If a taxon was no longer extant at its type locality, another representative population was selected. For two species, *I. texana* and *I. mattaponica*, material for DNA extraction could not be obtained. For most species, approximately five sporophylls from a single plant were stored in silica gel for DNA extraction. *Isoëtes tegetiformans* and *I. melanospora* were too small to acquire enough material from a single plant, so sporophylls from several plants in the same population were pooled. *Isoëtes nuttallii* was selected as the best outgroup from material available for DNA extraction due to its placement well outside the "American" clade in previous molecular phylogenies (Hoot et al. 2006, Larsén and Rydin 2016), despite its actual occurrence in western North America. Voucher specimens were deposited in the Old Dominion University (ODU) and US National (US) herbaria.

DNA Extraction and Sequencing

Total genomic DNA (gDNA) was extracted from approximately 200 mg of silica-dried leaf tissue using either the DNeasy Plant Mini Kit (Qiagen Inc., Valencia, California) or the Gene Prep instrument (Autogen Inc., Holliston, Massachusetts) using the manufacturers' instructions. Quantification of gDNA was done with a Qubit 2.0 fluorometer (ThermoFisher Scientific, Waltham, Massachusetts) and quality assessed by measuring the 260 nm:280 nm absorption ratio with an Epoch spectrophotometer (BioTek Instruments Inc., Winooski, Vermont). For each sample, an aliquot of 300 ng of gDNA was sheared into approximately 500-base pair (bp) fragments with a Q800R2 sonicator (Qsonica LLC, Newtown, Connecticut). Fragmentation to this size range was confirmed by separating sonicated DNAs on a 2% agarose gel by electrophoresis. Fragmented gDNA was cleaned using Agencourt AMPure XP beads (Beckman Coulter Inc., Brea, California) to size-select for approximately 500-bp fragments.

Fragmented gDNA was prepared for sequencing on a MiSeq sequencer (Illumina Inc., San Diego, California) using the NEBNext Ultra DNA Library Prep Kit and Multiplex Oligos for Illumina (New England Biolabs Inc., Ipswich,

Downloaded From: https://bioone.org/journals/Castanea on 29 Mar 2024
Terms of Use: https://bioone.org/terms-of-use

Table 1. List of taxa and specimens included in study. ODU = Old Dominion University herbarium, US = US National Herbarium.

Taxon	Voucher (Herbarium)	Locality	Date	GenBank Accession
I. butleri Engelm.	Schafran 47 (ODU)	Ft. Worth Nature Center, Texas	18 April 2015	MG668891
I. echinospora Durieu	Schafran 32 (ODU)	Cleveland Lake, Lewis Co., New York	16 July 2014	MG668903
I. engelmannii Braun	Schafran 46 (ODU)	Hiawassee River, Reliance, Tennessee	5 April 2015	MG668892
I. flaccida Shuttlew. var. *flaccida*	Taylor 6770 (US)	St. Marks River, Newport, Florida	26 January 2013	MG668893
I. flaccida var. *chapmannii* Engelm.	Bolin JB_FL_01 (ODU)	Chipola River, Marianna, Florida	13 September 2009	MG599108
I. lithophila Pfeiff.	Schafran 61 (ODU)	Enchanted Rock, Llano Co., Texas	20 April 2015	MG668894
I. melanopoda Gay and Durieu ssp. *melanopoda*	Taylor 6940 (US)	Giant City State Park, Jackson Co., Illinois	26 April 2015	MG668895
I. melanopoda ssp. *silvatica* Brunton and Britton	Schafran NC-05 (ODU)	Mecklenburg Co., North Carolina	10 March 2013	MG668896
I. melanospora Engelm.	Schafran 12 (ODU)	Summit of Stone Mountain, Georgia	15 May 2014	MG668897
I. mississippiensis Leonard, Taylor, Musselman, and Bray	Schafran MS-08 (US)	Lotts Creek, Picayune, Mississippi	18 June 2013	MG668898
I. nuttallii A. Braun ex Engelm.	Taylor 6734 (US)	Vernal Fall, Mariposa Co., California	14 June 2012	MG668899
I. piedmontana (Pfeiff.) C.F. Reed	Schafran 18 (ODU)	Heggie's Rock, Appling, Georgia	17 May 2014	MG668900
I. tegetiformans Rury	Schafran 19 (ODU)	Heggie's Rock, Appling, Georgia	17 May 2014	MG668901
I. valida (Engelm.) Clute	Schafran 37 (ODU)	Michaux State Forest, Cumberland Co., Pennsylvania	1 November 2014	MG668902

Massachusetts). End repair, adaptor ligation, indexing for multiplexing, and PCR enrichment were done according to the manufacturers' instructions (Supplementary Methods). Appropriate length and quantity of the libraries was confirmed with an Agilent 2200 Tapestation (Agilent Technologies, Santa Clara, California). Libraries were diluted or concentrated to 4 nM and 5 μL of each was added into one pool. A BluePippin instrument (Sage Science Inc., Beverly, Massachusetts) was used to size-select for 400–550 bp fragments. The size-selected 4 nM library pool then was submitted to the Smithsonian Laboratories of Analytical Biology sequencing facility.

Data Processing and Chloroplast Genome Assembly

Sequencing reads were downloaded from the Illumina BaseSpace database, having already been separated by primer indices into individual samples. A custom Python wrapper script was used to remove adaptor contamination and low-quality bases with Trimmomatic 0.33 (Bolger et al. 2014), and to combine paired-end reads with PEAR 0.9.6 (Zhang et al. 2014). Putative chloroplast reads were extracted by comparing all reads to a reference plastome of *I. flaccida* (Karol et al. 2010) using Bowtie2 (Langmead and Salzberg 2012). The putative chloroplast reads were reference-assembled to the *I. flaccida* plastome with the reference assembler in Geneious R10 (Biomatters Ltd., Auckland, New Zealand). Reads in the reference-based assembly were manually corrected around repeat regions. In addition, the putative chloroplast reads were *de novo* assembled using SPAdes 3.6.0 with k-mer lengths of 21, 33, 55, 77, 99, and 127 bp (Bankevich et al. 2012). Scaffolds from the *de novo* assembly were aligned to the *I. flaccida* plastome, and the consensus sequence was compared to the reference-based assembly. Any disagreements between assemblies, generally in repeat regions, were either manually corrected based on the raw data, or excluded from phylogenetic analysis. Plastome assemblies were annotated by comparison with the annotated *I. flaccida* plastome using Geneious with a cutoff of 90% similarity.

Phylogenetic Analyses

A multiple-sequence alignment program, MAFFT 7 (Katoh and Standley 2013) was used to align all the consensus plastome assemblies. An optimal

Downloaded From: https://bioone.org/journals/Castanea on 29 Mar 2024
Terms of Use: https://bioone.org/terms-of-use

evolutionary model for each alignment was selected with PAUP* v.4 (Swofford 2002) based on corrected Akaike information criterion (AICc) values. Less optimal models and rate variations were also evaluated to determine the effect on tree topology. MrBayes 3.2.6 (Huelsenbeck and Ronquist 2001) and RAxML 7.3.0 (Stamatakis 2006) were used for phylogenetic analysis. Gaps in the alignment were treated as missing data. Parsimony analysis was performed in PAUP*. As was done for total plastome alignments, gene introns and protein coding sequences were extracted separately using Geneious and those alignments were analyzed with MrBayes.

Using the automated model selection implemented in PAUP*, the generalized time reversible model (GTR; Tavaré 1986) of evolution was selected for whole plastome, coding sequence, and intron alignments including *I. nuttallii* (alignments without the outgroup were not analyzed). For the whole plastome, rate variation was best modeled using gamma-distributed rate variation and a proportion of invariable sites (GTR+I+G), whereas the optimal rate variation for the coding sequence and intron alignments included only the proportion of invariable sites parameter (GTR+I). In addition, other iterations of two common evolutionary models, Jukes-Cantor (JC; Jukes and Cantor 1969) and Hasegawa-Kishino-Yano (HKY; Hasegawa et al. 1985) and rate variations equal, proportion of invariable sites alone (+I), and gamma-distributed sites (+G) were applied to all alignments to test the sensitivity of tree topology to model selection. Coding sequence and intron alignments were analyzed both as concatenated and partitioned matrices.

All MrBayes analyses were run for 50 million generations, sampling every 1,000 generations, with one cold chain and three heated chains. Runs were assumed to be converged when split standard deviations were less than 0.001, and potential scale reduction factors equaled 1.000 +/− 0.001. MCMC output was visualized in Tracer 1.6 to check for problems with run convergence (Rambaut et al. 2014). All trees were rooted with *I. nuttallii*.

Ancestral character states were inferred using parsimony in Mesquite 3.31 (Maddison and Maddison 2017). All characters were treated as categorical and unordered.

RESULTS

Assembly and Alignment

Following filtering of low quality reads and merging of paired-end reads, the total number of reads per sample ranged from 310,119 to 2,058,639 (mean=1,345,289). Of the putative chloroplast reads that passed reference-based filtering with Bowtie2, there was a range from 7,479 to 89,260 reads (mean=43,527). The efficiency of genome skimming to collect chloroplast data is variable, with the percent of putative chloroplast reads varying from 0.94% to 8.65% of total reads, with a mean of 3.45%. This might represent inherent variation in the number of chloroplasts present in the collected plant tissues. This does not consider any variation of potential mitochondrial contamination that can pass through this filtering process. Between 6 and 22 scaffolds were constructed for each sample during *de novo* assembly, with a median of 12.5 scaffolds. N50 values ranged from 14,545 to 91,749 (median=26,025.5) and the sum length of scaffolds varied from 132,080 to 155,618 (median=138,690.5).

The final plastome assemblies displayed very little variation in size. Unaligned plastome length ranged from 144,680 to 145,294 (mean=145,102.8) with a percent standard deviation of 0.1%. The alignment of entire plastome sequences for thirteen southeastern *Isoëtes* diploids plus *I. nuttallii* yielded a matrix of 147,946 characters, of which 10,218 (6.9%) were variable (Table 2). Excluding *I. nuttallii*, the alignment length decreases to 146,006 and the proportion of variable sites decreases to 3,129 (2.0%). Pairwise identity between whole plastome alignments with and without the outgroup were 98.8% and 99.6%, respectively, and the percentages of parsimony informative sites were 0.21% and 0.19%, respectively.

Alignments of all 80 coding regions showed similar lengths and pairwise identities, 69,476 characters and 99.6% including *I. nuttallii*, and 69,372 characters and 99.8% excluding it (Table 2). The number of variable and parsimony informative sites decreased by excluding *I. nuttallii*, from 2.7% to 1.3% variable sites and 0.18% to 0.16% parsimony informative sites.

Noncoding intron alignments were 16,425 (with outgroup) and 15,986 characters (without outgroup) long, consisting of 20 separate regions. The regions showed the highest proportions of variable and parsimony informative

Downloaded From: https://bioone.org/journals/Castanea on 29 Mar 2024
Terms of Use: https://bioone.org/terms-of-use

Table 2. Alignment statistics with (left of dashed line) and without (right of dashed line) the outgroup *I. nuttallii*.

	Total Length (characters)		Variable Sites (%)		Parsimony Informative Sites (%)		Pairwise Identity (%)		Optimal Evolutionary Model
Whole plastome	147,946	146,006	6.9	2.0	0.21	0.19	98.8	99.6	GTR+I+G
Coding sequences	69,476	69,372	2.7	1.3	0.18	0.16	99.6	99.8	GTR+I
Introns	16,425	15,986	7.1	2.3	0.22	0.19	98.8	99.5	GTR+I

sites, 7.1% and 0.22%, respectively, including *I. nuttallii*. Excluding *I. nuttallii*, there were 2.3% variable sites and 0.19% parsimony informative sites. Intron alignments also displayed the lowest pairwise identities, 98.8% with *I. nuttallii*, and 99.5% without.

Whole Plastome Phylogeny

Across all model and rate iterations tested (GTR+I+G, GTR+I, GTR+G, GTR, HKY+I+G, HKY+I, HKY+G, HKY, JC+I+G, JC+I, JC+G, JC) there was no change in tree topology and little change in posterior probabilities (data not shown). Bayesian, maximum likelihood (ML), and parsimony-based trees shared the same topology. Support for several nodes was lower in ML compared with Bayesian analyses, although most were still >90 (Figure 1). The basal node of clade A, although well-supported in the Bayesian tree, is significantly weaker (bootstrap value of 69%) in the ML tree (Figure 1). Henceforth, all references will be to the Bayesian phylogeny inferred under the GTR+I+G model. All nodes were well-supported, with posterior probabilities of 100. Two major clades (Figure 2, clades A and B) are evident at the deepest node of the tree. Clade A consisted of *I. melanospora*, *I. engelmannii*, *I. melanopoda* ssp. *silvatica*, *I. tegetiformans*, and *I. piedmontana*, whereas *I. butleri*, *I. mississippiensis*, *I. flaccida* var. *chapmanii*, *I. melanopoda* ssp. *melanopoda*, *I. lithophila*, *I. flaccida* var. *flaccida*, *I. echinospora*, and *I. valida* comprised clade B. Within clade B were two sister groups,

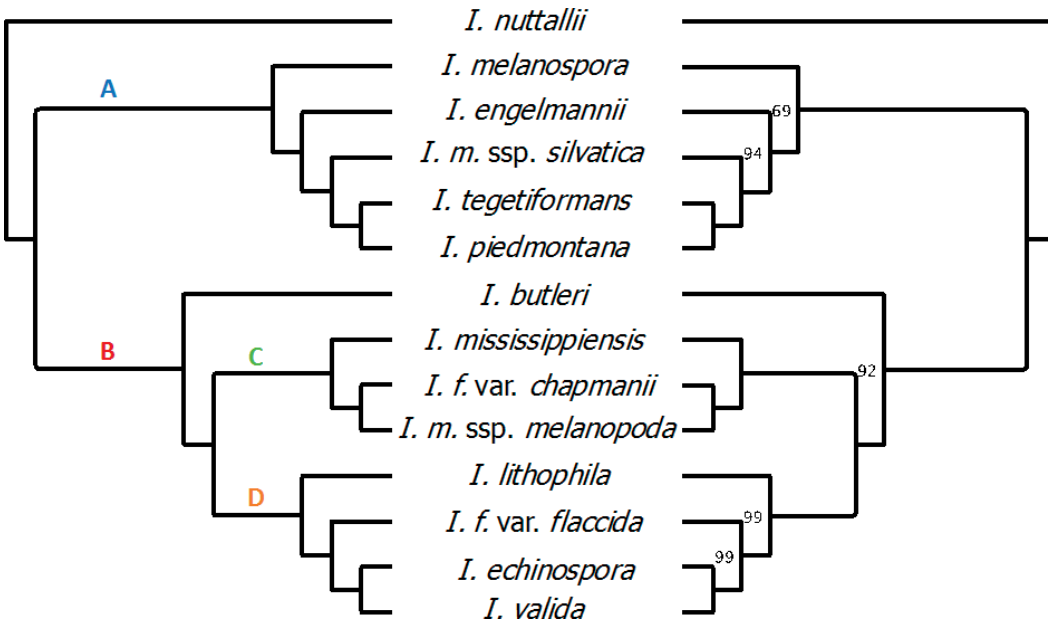


Figure 1. Bayesian (left) and maximum-likelihood (right) phylograms of the whole plastome alignment. All nodes were supported by posterior probabilities or bootstrap values of 100 except where noted. Letters label major clades.

Downloaded From: https://bioone.org/journals/Castanea on 29 Mar 2024
Terms of Use: https://bioone.org/terms-of-use

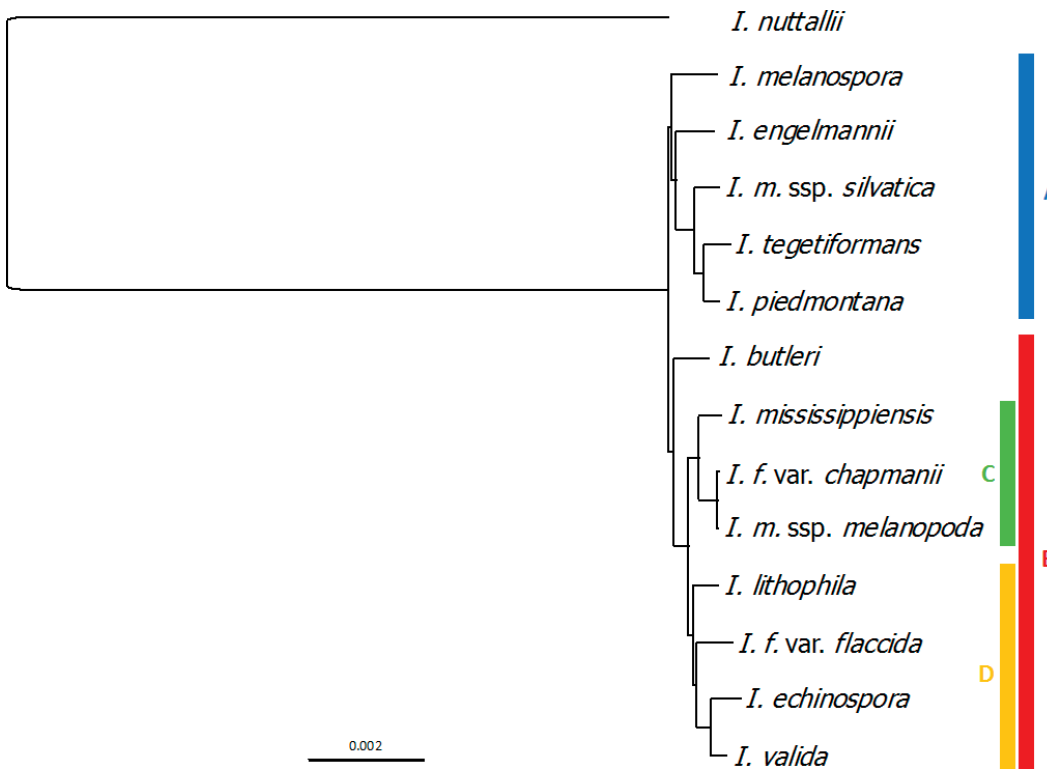


Figure 2. Bayesian phylogram of the whole plastome alignment. All nodes were supported with posterior probabilities of 100. Colored bars/letters represent major clades. Scale bar represents substitutions per site.

clade C (*I. mississippiensis*, *I. flaccida var. chapmanii*, *I. melanopoda* ssp. *melanopoda*) and clade D (*I. lithophila*, *I. flaccida var. flaccida*, *I. echinospora*, *I. valida*). Branch lengths ranged from 3.22e-4 to 5.65e-5 substitutions per site.

Protein Coding Sequence and Intron Phylogenies

In general, the phylogenies inferred from coding sequence and intron alignments separately supported the whole plastome phylogeny, although topologies and support values varied. Most of the relationships within clades A, B, C, and D were supported by both datasets. The placement of *I. butleri* or *I. melanospora* as the most basal branch was weakly supported by the coding sequence and intron phylogenies, respectively (Figure 3). Both of these topologies conflicted with that of the whole plastome, where neither species was sister to the remaining taxa (Figure 2). Except for the placement of *I. melanospora*, the remainder of clade A was supported by both coding sequences and introns. Likewise, clade B was supported by these data, except for the placement of *I. butleri*. Clade C was well supported, with no conflicts in topology among any analysis. Clade D was mostly well supported, but the intron data resolved a polytomy at its backbone, rather than placing *I. lithophila* as the most basal branch of the clade (Figure 3).

Character State Reconstruction

Within clade A, spores that mature in the spring before plants enter their summer desiccation-dormancy is common to several taxa: *I. melanopoda* ssp. *silvatica*, *I. melanospora*, *I. tegetiformans*, and *I. piedmontana* (Figure 4). Only individuals of *I. engelmannii* persist through the summer to produce their spores in the fall. Megaspore ornamentation is fairly consistent among all members of this clade. All taxa have tuberculate megaspores (occasionally pseudoreticulate to cristate in *I. piedmontana*), except *I. engelmannii* which has strongly reticulate megaspores. Gray-black megaspore coloration

Downloaded From: https://bioone.org/journals/Castanea on 29 Mar 2024
Terms of Use: https://bioone.org/terms-of-use

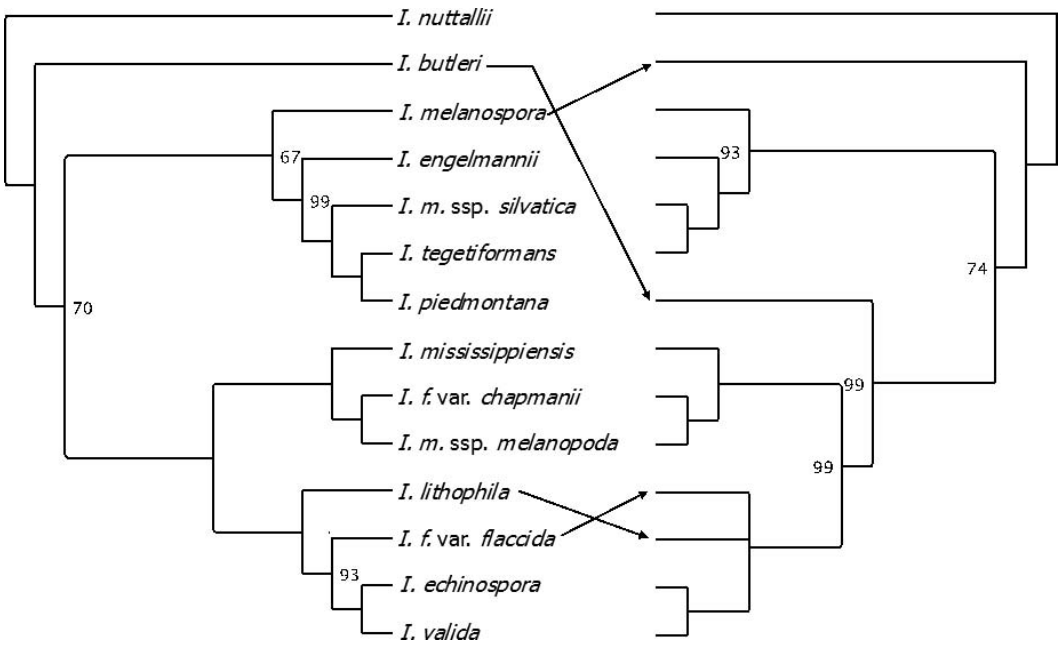


Figure 3. Bayesian phylograms based on coding sequence (left) and intron (right) alignments with arrows indicating changes in topology. All nodes were supported by posterior probabilities of 100 except where noted.

occurs in *I. melanospora* and *I. tegetiformans*, whereas all other have taxa are white megaspores.

In clade B, most taxa (*I. flaccida* var. *chapmanii*, *I. flaccida* var. *flaccida*, *I. echinospora*, and *I. valida*) have spores that mature in late summer or autumn. *Isoëtes butleri*, *I. mississippiensis*, *I. melanopoda* ssp. *melanopoda*, and *I. lithophila* have spores that mature in late spring. Most taxa in this clade have spinulose microspores, except for *I. butleri* and *I. flaccida* s.l., which have papillose microspores. Most taxa in clade B have tuberculate megaspores. Only *I. mississippiensis*, *I. echinospora*, and *I. valida* differ, with laevigate echinate and reticulate-cristate ornamentation, respectively. *Isoëtes lithophila* is the only taxon in this clade with gray-black megaspore coloration.

Ancestral character state reconstruction places white, tuberculate megaspore ornamentation, papillose microspore ornamentation, and springtime maturation at the basal node of the southeastern clade (Figure 4). Of eleven nodes within clades A and B, tuberculate megaspore ornamentation is inferred at 10 (one has more than one most parsimonious state) and white coloration is found at 11 nodes. Microspore ornamentation is inferred to be spinulose at seven nodes, whereas two nodes have papillose ornamentation (two have multiple most parsimonious states). A phenology of spores maturing in spring is the most parsimonious state at nine nodes, with two nodes having autumn-maturing spores.

DISCUSSION Our analyses indicate that plastome DNA sequences are useful for resolving species relationships in closely related groups of *Isoëtes*. Comparison of phylogenies suggests that for the same dataset, evolutionary model selection has little effect on resulting tree topology and support. Algorithm selection (i.e., MrBayes vs. RAxML) appears to have little effect on topology, but can result in different levels of support. Comparison of coding sequence and intron-based phylogenies suggests that different topologies can be inferred depending on marker selection. This should serve as a cautionary note for other phylogenetic studies of *Isoëtes* based on relatively short markers, because selection for a certain type of marker region can influence the resulting phylogeny, especially among closely related taxa.

Downloaded From: https://bioone.org/journals/Castanea on 29 Mar 2024
Terms of Use: https://bioone.org/terms-of-use

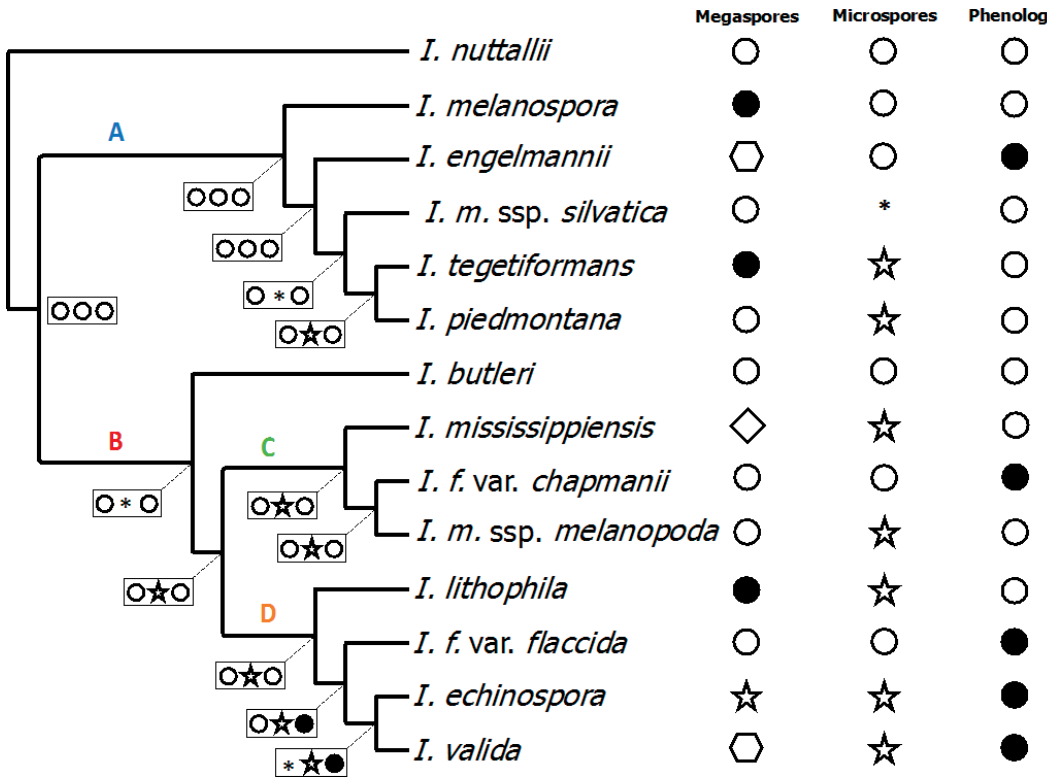


Figure 4. Bayesian whole plastome phylogram with character states. Megaspore key: circle = tuberculate; hexagon = reticulate; star = echinate; diamond = laevigate; open = white spores; filled = gray-black spores. Microspore key: circle = papillose; star = spinulose. Phenology key: open = spores mature in spring; filled = spores mature in summer-autumn. * = character state unknown or more than one most parsimonious state. Node labels follow the order of the table.

The phylogenetic relationships and inference of ancestral character states supports the hypothesis that most of the extant morphological diversity in southeastern *Isoëtes* is relatively recent, and many similar traits are the result of convergence rather than descent (Hickey 1986, Taylor and Hickey 1992). Parsimony indicates that 11 of 14 character-state transitions in the southeastern clade occurred on the terminal branches of the tree. Only three are inferred more deeply in the phylogeny: spinulose microspore ornamentation uniting the clades containing *I. tegetiformans* and *I. piedmontana* (potentially including *I. melanopoda* ssp. *silvatica*) as well as clades C and D (potentially all of clade B), and autumn phenology uniting *I. flaccida* var. *flaccida*, *I. echinospora*, and *I. valida*. These results suggest that species groupings based on these features (e.g., Pfeiffer 1922) do not accurately represent evolution in the genus. Further work should examine these apparently labile characters and what roles they might play in *Isoëtes* speciation, especially where species are adapted to particular environmental conditions (Taylor et al. 1993).

Although a complete analysis of the evolution of the plastome in southeastern *Isoëtes* is not presented here, we note that all genes and transfer RNA and ribosomal RNA coding regions are retained in all taxa. Almost all taxa display numerous autapomorphic sites as inferred from branch lengths, and the number of site differences between pairwise taxa (excluding *I. nuttallii*) is generally several hundred (mean=572; median=607; Figure 5). One exception occurs between *I. flaccida* var. *chapmanii* and *I. melanopoda* ssp. *melanopoda*, which have only 15 site differences—approximately the same variation observed between two individuals of *I. flaccida* var. *flaccida*. This does not appear to be a misidentification, because nuclear DNA sequences from the same individuals of *I.*

Downloaded From: https://bioone.org/journals/Castanea on 29 Mar 2024
Terms of Use: https://bioone.org/terms-of-use

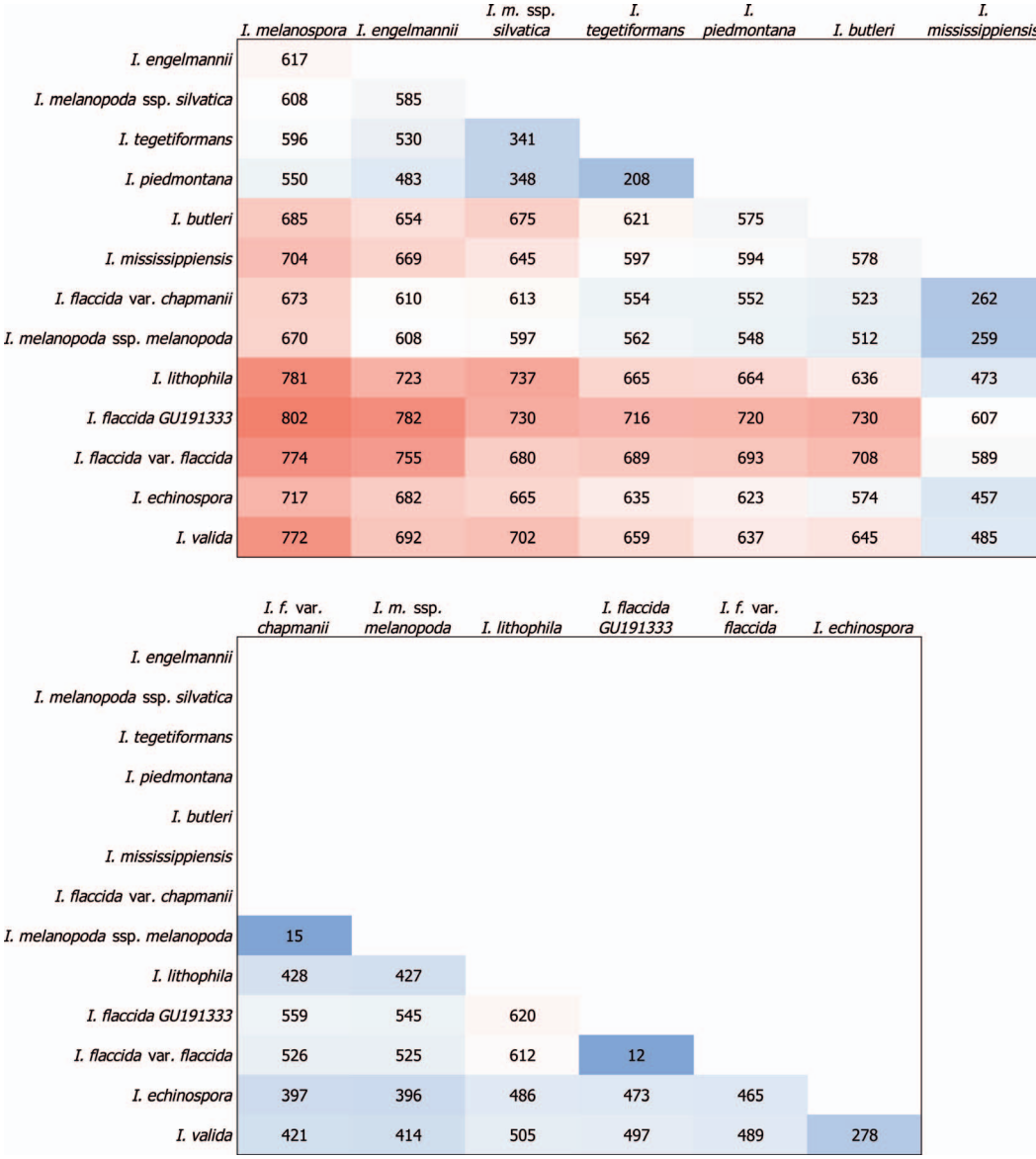


Figure 5. Heat map of number of pairwise differences from the whole plastome alignment, excluding *I. nuttallii*. Includes the reference *I. flaccida* (GU191333) from Karol et al. (2010). Colors scaled from blue (minimum) to white (median) to red (maximum).

flaccida var. *chapmanii* and *I. melanopoda* ssp. *melanopoda* do not indicate a close relationship (unpublished data). Instead, this might represent a hybrid origin of the chloroplast within *I. flaccida* var. *chapmanii* through chloroplast capture. A few structural differences appear in the *I. nuttallii* plastome relative to the southeastern US taxa. Two small inversions are present, one in the *atpB-rbcL* spacer that is 344 bp long, with the other in the *trnK*-UUU-*psbA* spacer that is 24 bp long. In addition, the entire *atpF* intron has been lost in *I. nuttallii*. This does not appear to be a problem of assembly, because sequence reads clearly span the site of the intron.

Downloaded From: https://bioone.org/journals/Castanea on 29 Mar 2024
Terms of Use: https://bioone.org/terms-of-use

Although this study represents a significant advancement in our understanding of the phylogeny of *Isoëtes* in the Southeast, it also highlights many areas needing further research. The plastome represents only one evolutionary lineage and might conflict with nuclear and mitochondrial genomes, causing incongruence between genome phylogenies and species phylogenies. Preliminary nuclear data from *Isoëtes* of the Southeast indicate different species phylogenies are inferred from different genomes (unpublished data). The inheritance and evolution of plastomes, often represented as uniparentally-inherited and nonrecombinational, can be complicated by heteroplasmy, recombination, and intracellular horizontal gene transfer in many plant taxa (Wolfe and Randle 2004, Scarcelli et al. 2016). More complex models of molecular evolution are needed to compensate for intraindividual and intraspecies variation in plastomes (Wicke and Schneeweiss 2015). The level of genetic variation between populations of *Isoëtes* and the vectors by which populations interbreed are very poorly known for most species. In terrestrial habitats with no obvious spore dispersal mechanisms, populations of *Isoëtes* are assumed to be reproductively isolated (Taylor and Hickey 1992). In these cases, small population size and genetic drift might be driving speciation. As more is learned about overlooked morphological characters, molecular phylogenies could gain additional support (Freund 2016, Bray et al. 2018). The addition of unsampled (i.e., *I. texana* and *I. mattaponica*) and potentially unrecognized taxa (e.g., a unique population of *I. melanospora* in South Carolina, Taylor et al. 1993) to this phylogeny could further enhance our understanding of the evolution of diploid *Isoëtes* in the southeastern USA. Finally, this study indicates polyphyly of both *I. flaccida* s.l. and *I. melanopoda* s.l., supporting the raising of *I. flaccida* var. *chapmanii* and *I. melanopoda* ssp. *silvatica* each to the species level, and highlighting the need to update the taxonomy of the genus based on molecular phylogenetic data.

ACKNOWLEDGMENTS The authors would like to thank J. Allison, J. Bolin, R. Brubaker, S. Leonard, R. Matthews, R. O'Kennon, J. Singhurst, K. Sitch, and L. Winslett for their assistance in the field. We thank M. Carlsen, M. Halloran, G. Johnson, and K. Wu for help with DNA sequencing, genome assembly, and phylogenetic analysis. This manuscript was benefited by the comments of two anonymous reviewers. Funding was provided by the Mary Payne Hogan Endowment at ODU and a Graduate Student Fellowship from the Smithsonian Institution. Portions of the laboratory and computer work were conducted in and with the support of the Laboratories of Analytical Biology (LAB) facilities of the National Museum of Natural History.

LITERATURE CITED

Bankevich, A., S. Nurk, D. Antipov, A.A. Gurevich, M. Dvorkin, A.S. Kulikov, V.M. Lesin, S.I. Nikolenko, S. Pham, A.D. Prjibelski, A.V. Pyshkin, A.V. Sirotkin, N. Vyahhi, G. Tesler, M.A. Alekseyev, and P.A. Pevzner. 2012. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. J. Computat. Biol. 19:455–477.

Beck, J.B., M.D. Windham, G. Yatskievych, and K.M. Pryer. 2010. A diploids-first approach to species delimitation and interpreting polyploid evolution in the fern genus *Astrolepis* (Pteridaceae). Syst. Bot. 35:223–234.

Bolger, A.M., M. Lohse, and B. Usadel. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 30:2114–2120.

Bray, R.D, P.W. Schafran, and L.J. Musselman. 2018. Interesting, provocative, and enigmatic: morphological studies on southeastern quillworts (*Isoëtes* Isoëtaceae, Lycophyta). Castanea 83. In press.

Brunton, D.F. 2015. Key to the quillworts (*Isoëtes*: Isoëtaceae) of the southeastern United States. Amer. Fern J. 105:86–100.

Burgess, M.B., K.R. Cushman, E.T. Doucette, C.T. Frye, and C.S. Campbell. 2015. Understanding diploid diversity: a first step in unraveling polyploid, apomictic complexity in *Amelanchier*. Amer. J. Bot. 102:2041–2057.

Freund, F. 2016. Characterizing quantitative variation in the glossopodia of three western North American *Isoëtes* species. Amer. Fern J. 106:87–115.

Hasegawa, M., H. Kishino, and T. Yano. 1985. Dating of the human-ape splitting by a molecular clock of mitochondrial DNA. J. Molec. Evol. 22:160–174.

Hickey, R.J. 1986. *Isoëtes* megaspore surface morphology: nomenclature, variation, and systematic importance. Amer. Fern J. 76:1–16.

Downloaded From: https://bioone.org/journals/Castanea on 29 Mar 2024
Terms of Use: https://bioone.org/terms-of-use

Hickey, R.J. 1990. Studies of neotropical *Isoëtes* L.: I. *Euphyllum*, a new subgenus. Ann. Missouri Bot. Gard. 77:239–245.

Hoot, S.B. and W.C. Taylor. 2001. The utility of nuclear ITS, a *LEAFY* homolog intron, and chloroplast *atpB-rbcL* spacer region data in phylogenetic analyses and species delimitation in *Isoëtes*. Amer. Fern J. 91:166–177.

Hoot, S.B., W.C. Taylor, and N.S. Napier. 2006. Phylogeny and biogeography of *Isoëtes* (Isotaceae) based on nuclear and chloroplast DNA sequence data. Syst. Bot. 31:449–460.

Huelsenbeck, J.P. and F. Ronquist. 2001. MRBAYES: Bayesian inference of phylogeny. Bioinformatics 17:754–755.

Jukes, T.H. and C.R. Cantor. 1969. Evolution of protein molecules. p. 21–132. *In*: Munro H.N. (ed.). Mammalian protein metabolism. Academic Press, New York, New York.

Karol, K., K. Arumuganathan, J.L. Boore, A.M. Duffy, K.D.E. Everett, J.D. Hall, S.K. Hansen, J.V. Kuehl, D.F. Mandoli, B.D. Mishler, R.G. Olmstead, K.S. Renzaglia, and P.G. Wolf. 2010. Complete plastome sequences of *Equisetum arvense* and *Isoëtes flaccida*: implications for phylogeny and plastid genome evolution of early land plant lineages. B. M. C. Evol. Biol. 10:321.

Katoh, K. and D.M. Standley. 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Molec. Biol. Evol. 30:772–780.

Langmead, B. and S. Salzberg. 2012. Fast gapped-read alignment with Bowtie2. Nature Meth. 9: 357–359.

Larsén, E. and C. Rydin. 2016. Disentangling the phylogeny of *Isoëtes* (Isoetales), using nuclear and plastid data. Int. J. Pl. Sci. 177:157–174.

Maddison, W.P. and D.R. Maddison. 2017. Mesquite: A Modular System for Evolutionary Analysis. Version 3.31 (http://mesquiteproject.org).

Pfeiffer, N.E. 1922. Monograph of the Isoetaceae. Ann. Missouri Bot. Gard. 9:79–233.

Rambaut A., A.J. Drummond, D. Xie, G. Baele, and M.A. Suchard. 2018. Posterior summarisation in Bayesian phylogenetics using Tracer 1.7. Systematic Biology syy032.

Rydin, C. and N. Wikström. 2002. Phylogeny of *Isoëtes* (Lycopsida): resolving basal relationships using *rbcL* sequences. Taxon 51:83–89.

Scarcelli, N., C. Mariac, T.L.P. Couvreur, A. Faye, D. Richard, F. Sabot, C. Berthouly-Salazar, and Y. Vigouroux. 2016. Intra-individual polymorphism in chloroplasts from NGS data: where does it come from and how to handle it? Molec. Ecol. Resources 16: 434–445.

Schuettpelz, E. and S.B. Hoot. 2006. Inferring the root of *Isoëtes*: exploring alternatives in the absence of an acceptable outgroup. Syst. Bot. 31:258–270.

Shaw, J., E.B. Lickey, J.T. Beck, S.B. Farmer, W. Liu, J. Miller, K.C. Siripun, C.T. Winder, E.E. Schilling, and R.L. Small. 2005. The tortoise and the hare II: relative utility of 21 noncoding chloroplast DNA sequences for phylogenetic analysis. Amer. J. Bot. 92:142–166.

Shaw, J., E.B. Lickey, E.E. Schilling, and R.L. Small. 2007. Comparison of whole chloroplast genome sequences to choose noncoding regions for phylogenetic studies in angiosperms: the tortoise and the hare III. Amer. J. Bot. 94: 275–288.

Shaw, J., H.L. Shafer, O.R. Leonard, M.J. Kovach, M. Schorr, and A.B. Morris. 2014. Chloroplast DNA sequence utility for the lowest phylogenetic and phylogeographic inferences in angiosperms: the tortoise and the hare IV. Amer. J. Bot. 101:1987–2004.

Stamatakis, A. 2006. RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. Bioinformatics 22:2688–2690.

Swofford, D. L. 2002. PAUP*. Phylogenetic Analysis Using Parsimony (*and Other Methods). Version 4. Sinauer Associates, Sunderland, Massachusetts.

Tavaré, S. 1986. Some probabilistic and statistical problems in the analysis of DNA sequences. Lectures on Mathematics in the Life Sciences 17:57–86.

Taylor, W.C. and R.J. Hickey. 1992. Habitat, evolution, and speciation in *Isoëtes*. Ann. Missouri Bot. Gard. 79:613–622.

Taylor, W.C., N.T. Luebke, D.M. Britton, R.J. Hickey, and D.F. Brunton. 1993. Isoëtaceae. p. 64–75. *In*: FNA Editorial Committee (eds.),

Downloaded From: https://bioone.org/journals/Castanea on 29 Mar 2024
Terms of Use: https://bioone.org/terms-of-use

Flora of North America, North of Mexico, vol. 2, 64–75. Oxford University Press, New York, New York.

Troia, A., J.B. Pereira, C. Kim, and W.C. Taylor. 2016. The genus *Isoetes* (Isoetaceae): a provisional checklist of the accepted and unresolved taxa. Phytotaxa 277:101–145.

Wicke, S. and G.M. Schneeweiss. 2015. Next-generation organellar genomics: potentials and pitfalls of high-throughput technologies for molecular evolutionary studies and plant systematics. p. 1–42. *In*: Hörandl, E. and M.S. Appelhans (eds.). Next-generation sequencing in plant systematics. Koeltz Scientific Books, Oberreifenberg, Germany.

Wolfe, A.D. and C.P. Randle. 2004. Recombination, heteroplasmy, haplotype polymorphism, and paralogy in plastid genes: implications for plant molecular systematics. Syst. Bot. 29: 1011–1020.

Zhang, J., K. Kobert, T. Flouri, and A. Stamatakis. 2014. PEAR: a fast and accurate Illumina paired-end read merger. Bioinformatics 30: 614–620.

Downloaded From: https://bioone.org/journals/Castanea on 29 Mar 2024
Terms of Use: https://bioone.org/terms-of-use