

Nuclear-plastid phylogenomic discordance in *Isoetes* (Lycopodiophyta) from North America



Peter Schafran^{1,2}, GoFlag Consortium³, Elizabeth Zimmer¹, Carl Taylor¹, and Lytton Musselman²

¹National Museum of Natural History, Washington, DC

²Old Dominion University, Norfolk, VA

³University of Florida, Field Museum, and University of Arizona. Project personnel include (at UF): J. Gordon Burleigh, Emily Sessa, Stuart McDaniel, Christine Davis, Pavlo Antonenko, Sarah Carey, Lorena Endara, Weston Testo; (at Field): Matt von Konrat, Eve Gaus; (at UA): Hong Cui.

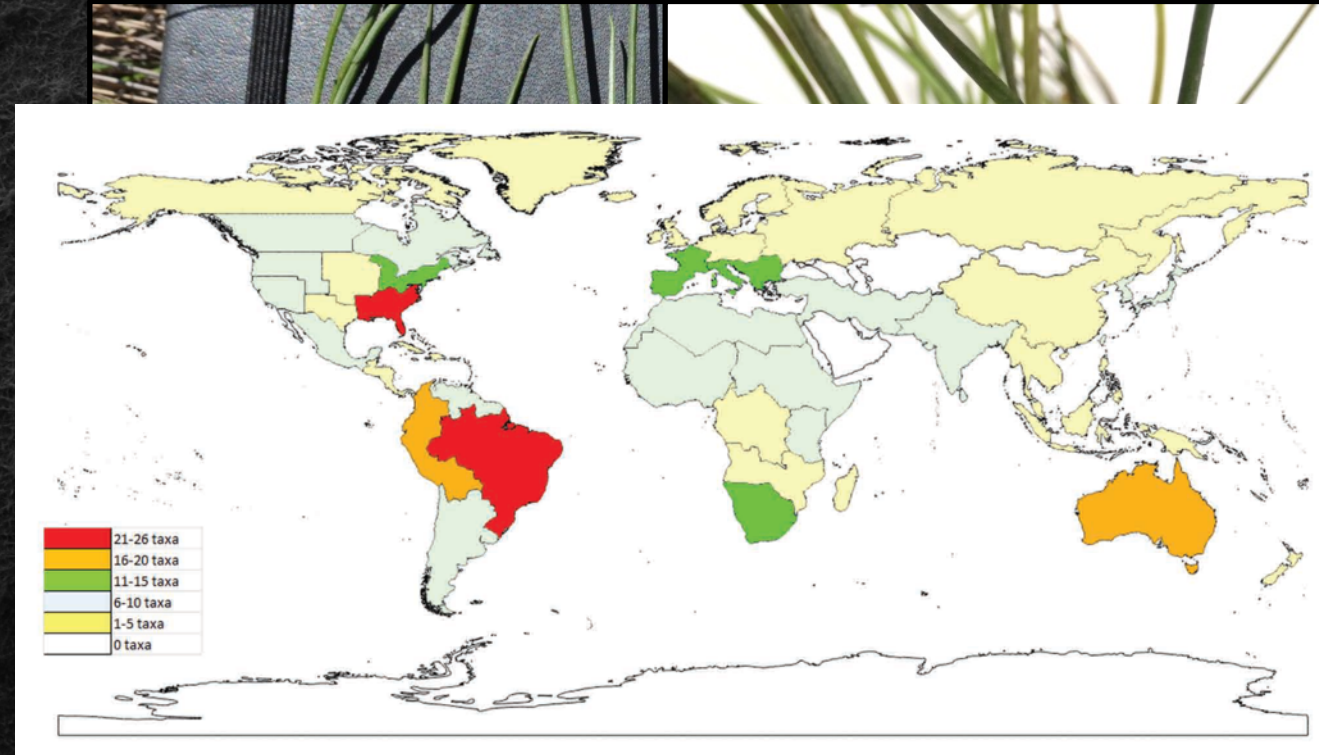
Isoetes (Quillworts, Merlin's Grass)

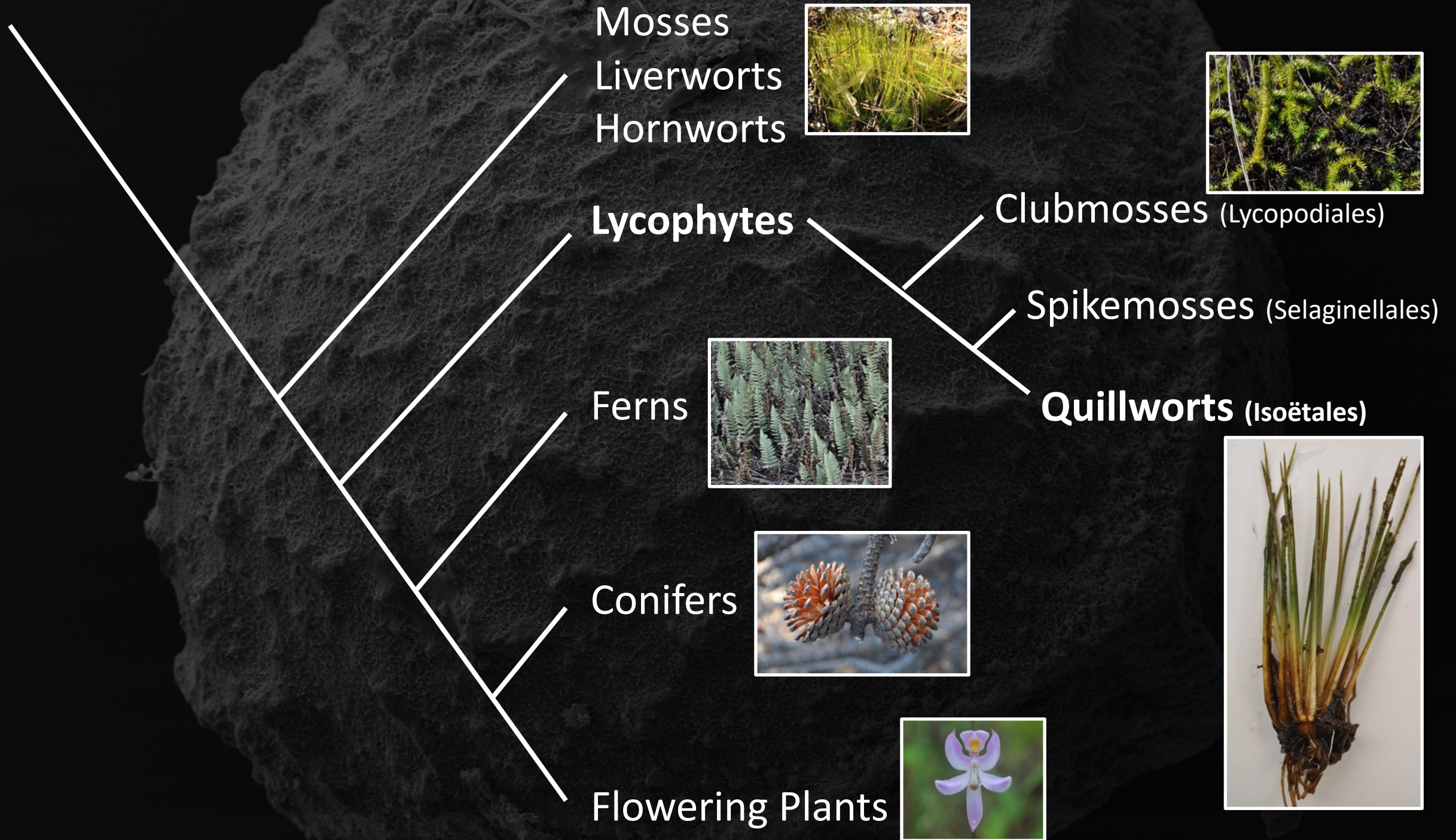
Cosmopolitan except Antarctica

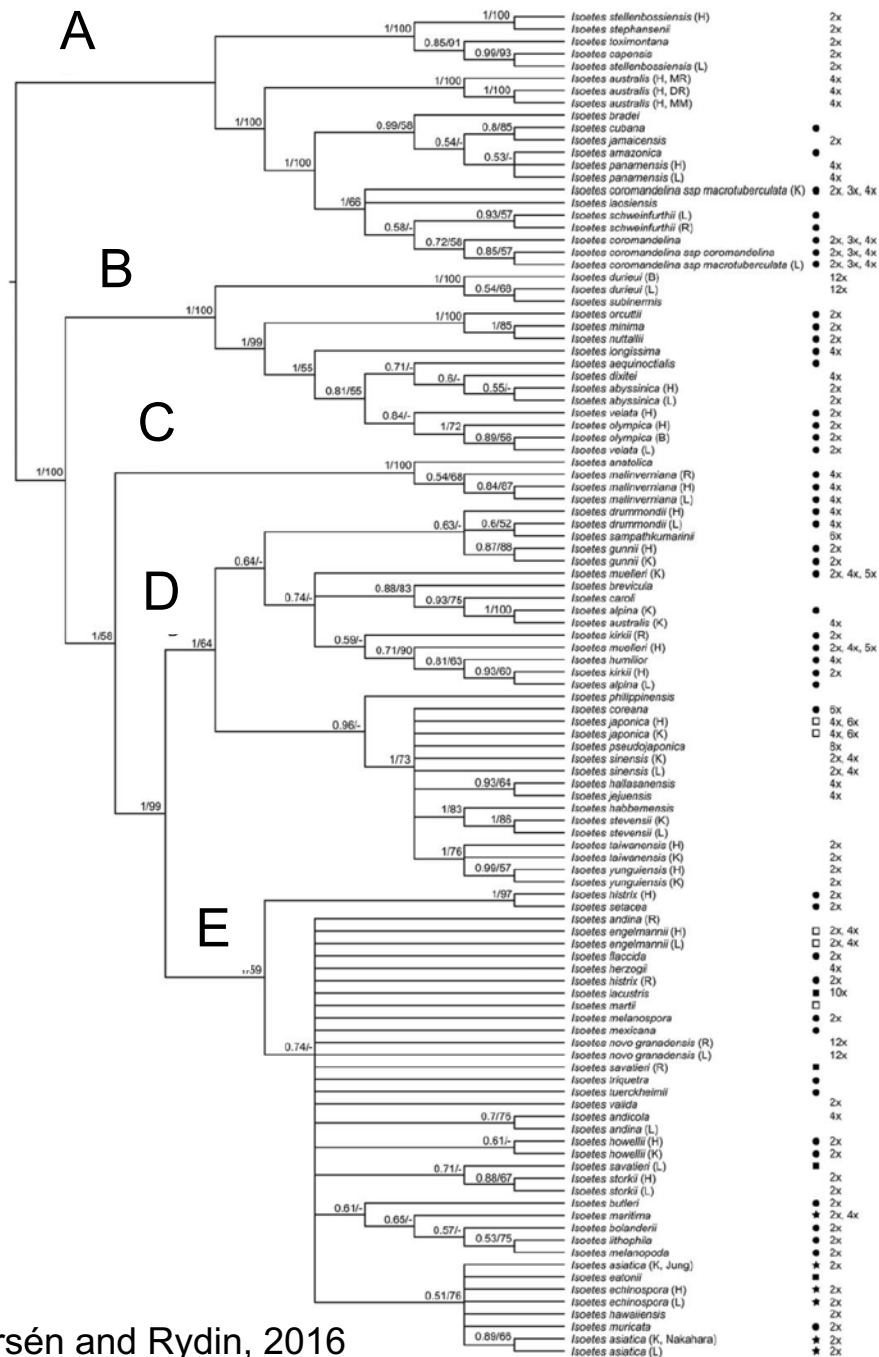
Estimated 150-300 spp.

Occur in nearly any freshwater habitat

Extremely conserved morphology:
“Cone on a corm”







Africa, Australasia, Neotropics

Mediterranean, California, Africa, India

Mediterranean

Australasia

Americas

Basic
diploids (2x)



(3x)



(6x)



Sterile hybrids (2x)



Fertile polyploids (4x)



Genome doubling
Instant speciation

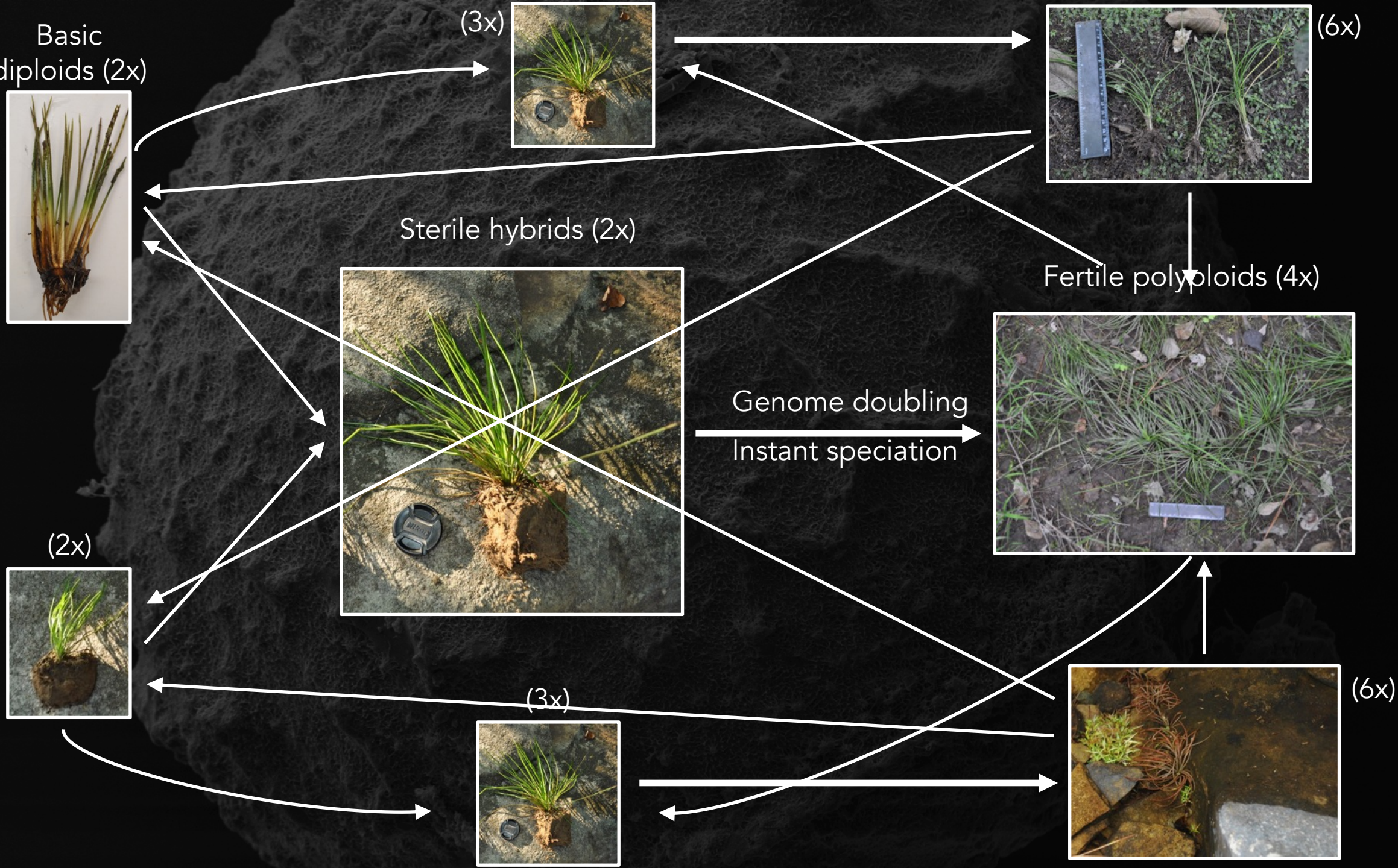
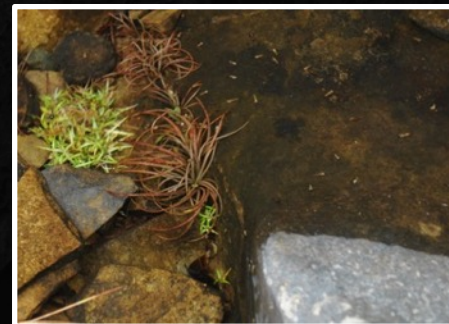
(2x)



(3x)



(6x)



15 diploid species

I. butleri
I. echinospora
I. engelmannii
I. flaccida s.l.
I. lithophila
I. mattaponica
I. melanopoda s.l.
I. melanospora
I. mississippiensis
I. piedmontana
I. prototypus
I. tegetiformans
I. texana
I. valida
I. viridimontana



RETICULATE
EVOLUTION



14 allopolyploid species

I. appalachiana
I. boomii
I. georgiana
I. graniticola
I. hyemalis
I. junciformis
I. lacustris
I. louisianensis
I. microvela
I. septentrionalis
I. tennesseeensis
I. tuckermanii
I. riparia
I. virginica



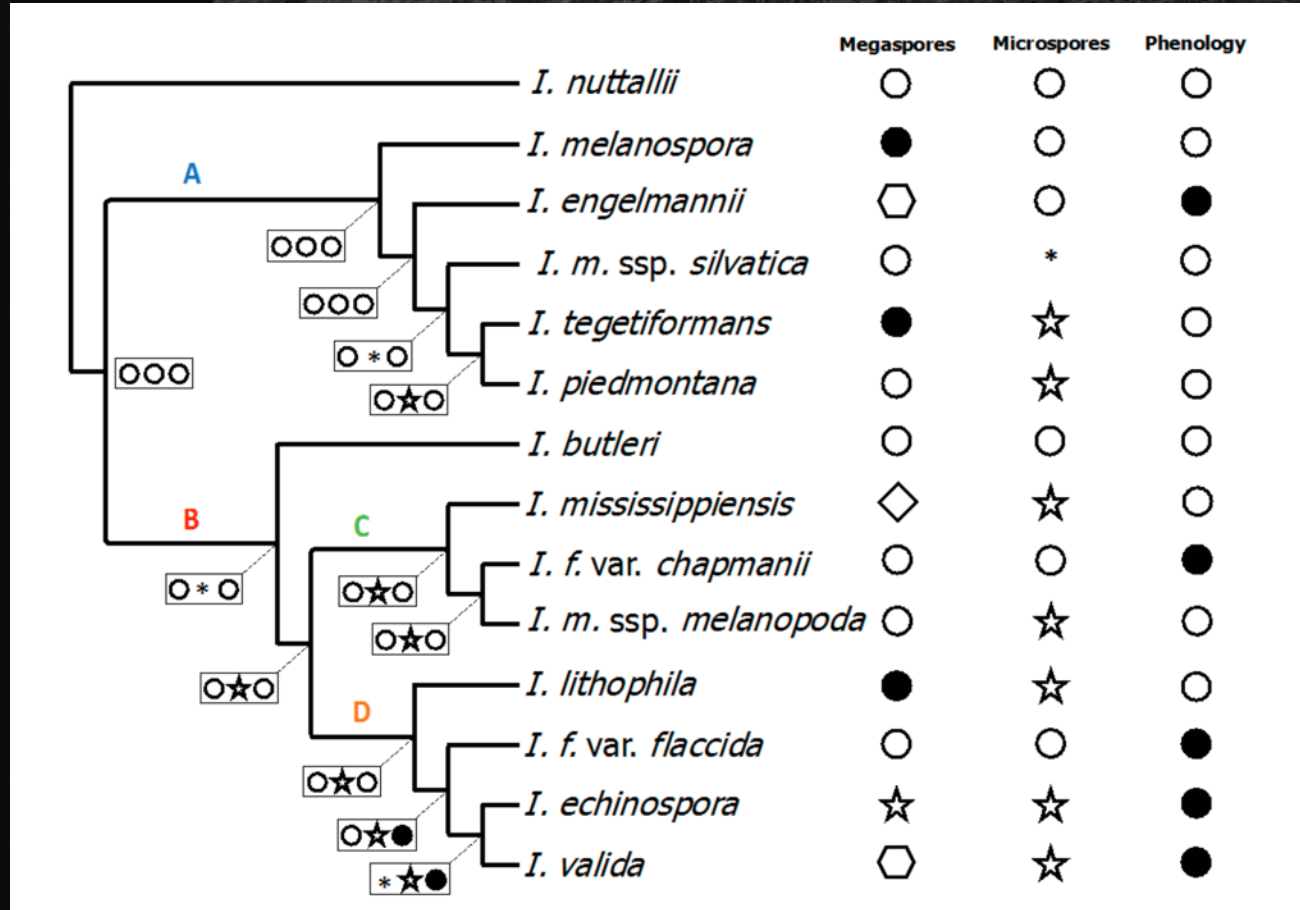
Diploid and hybrid from Hiwassee River, TN

Resolve Diploid Backbone Phylogeny

1. Whole chloroplast genomes
 - Genome skimming on Illumina MiSeq/HiSeq
 - Combined reference-based and *de novo* assembly
2. Target-capture low-copy nuclear markers
 - Generated by GoFlag project
 - Target loci reconstructed by several pipelines: Phyluce, HybPiper, HybPhyloMaker
 - Alignments trimmed by various parameters
 - Trees inferred by RAxML and IQ-TREE

Plastome Phylogeny

Diploid species in the SE US

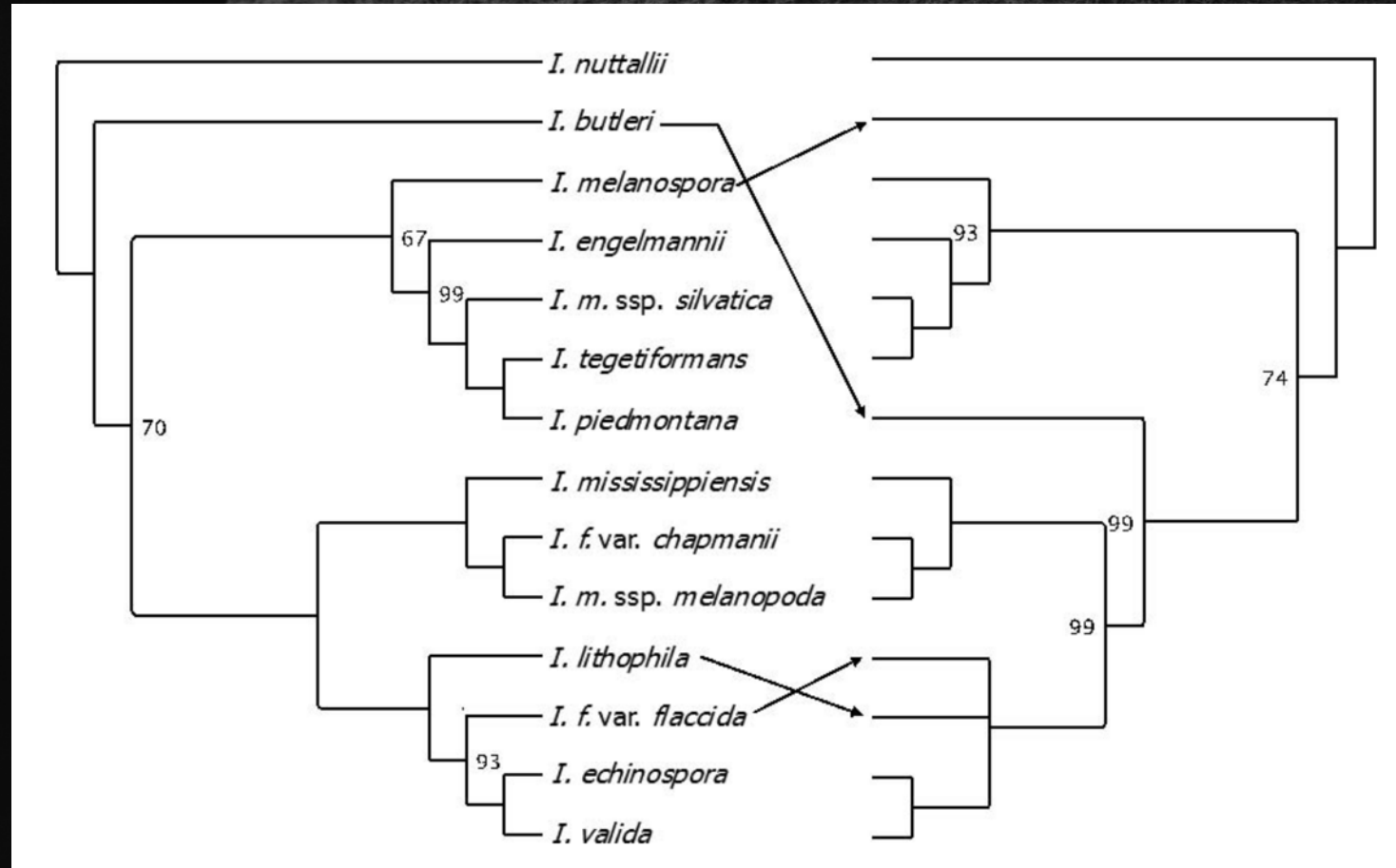


ML analysis of concatenated data

Resolves relationships between species, subspecies, varieties

Plastome Phylogeny

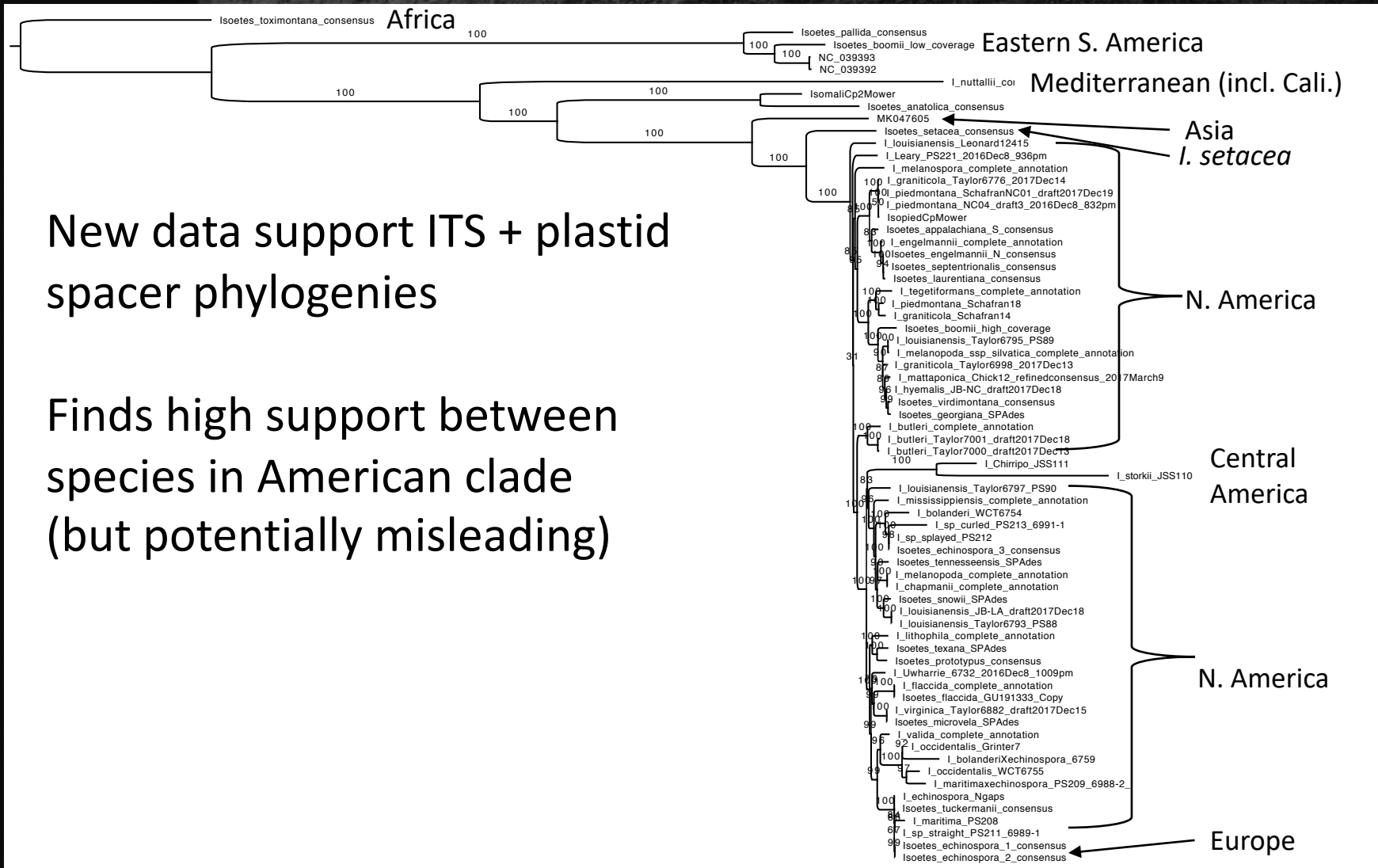
Diploid species in the SE US



Discordance based on plastid subsets

Coding sequences (left) vs. non-coding (right)

Plastome Phylogeny

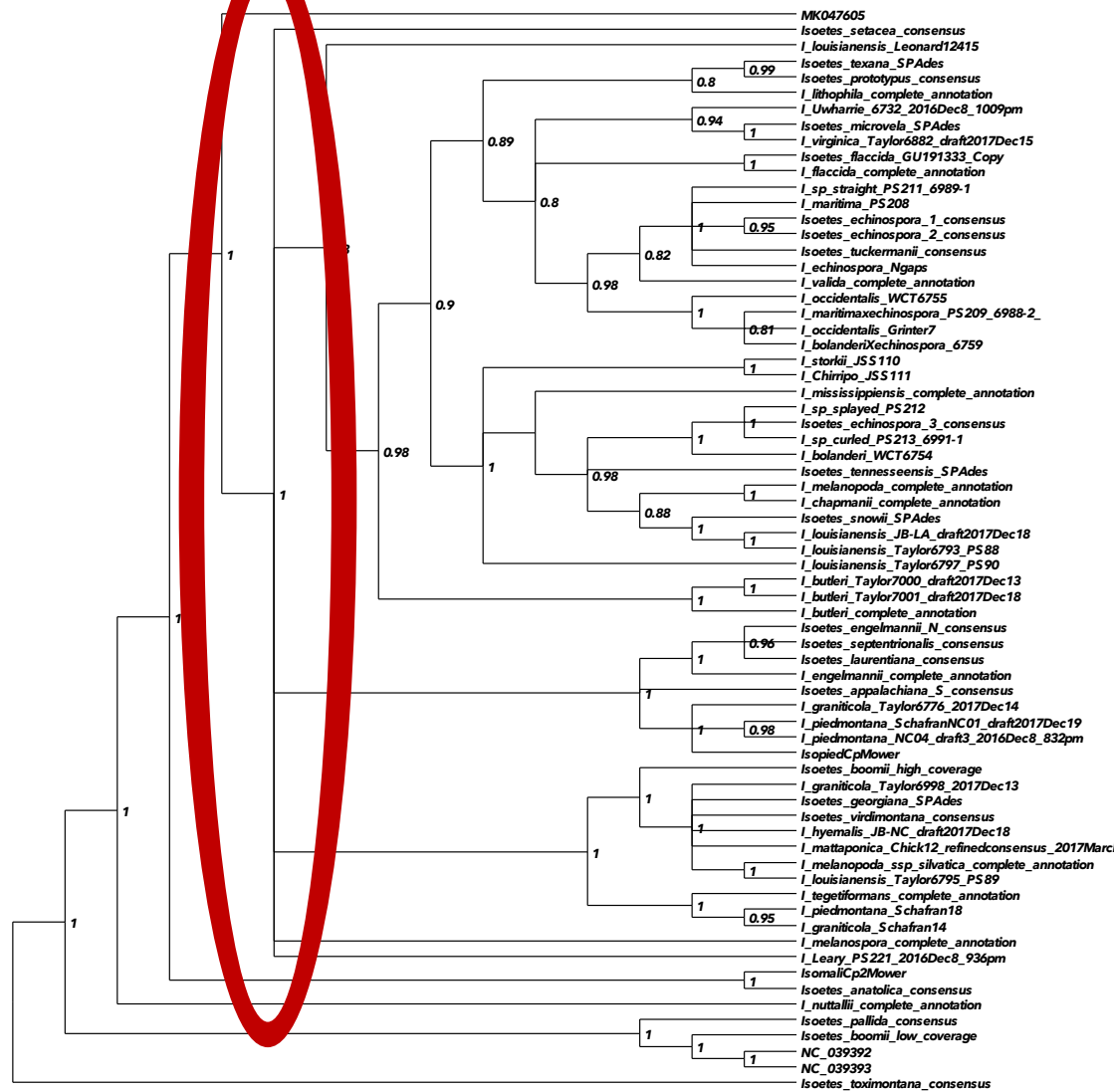


New data support ITS + plastid spacer phylogenies

Finds high support between species in American clade (but potentially misleading)

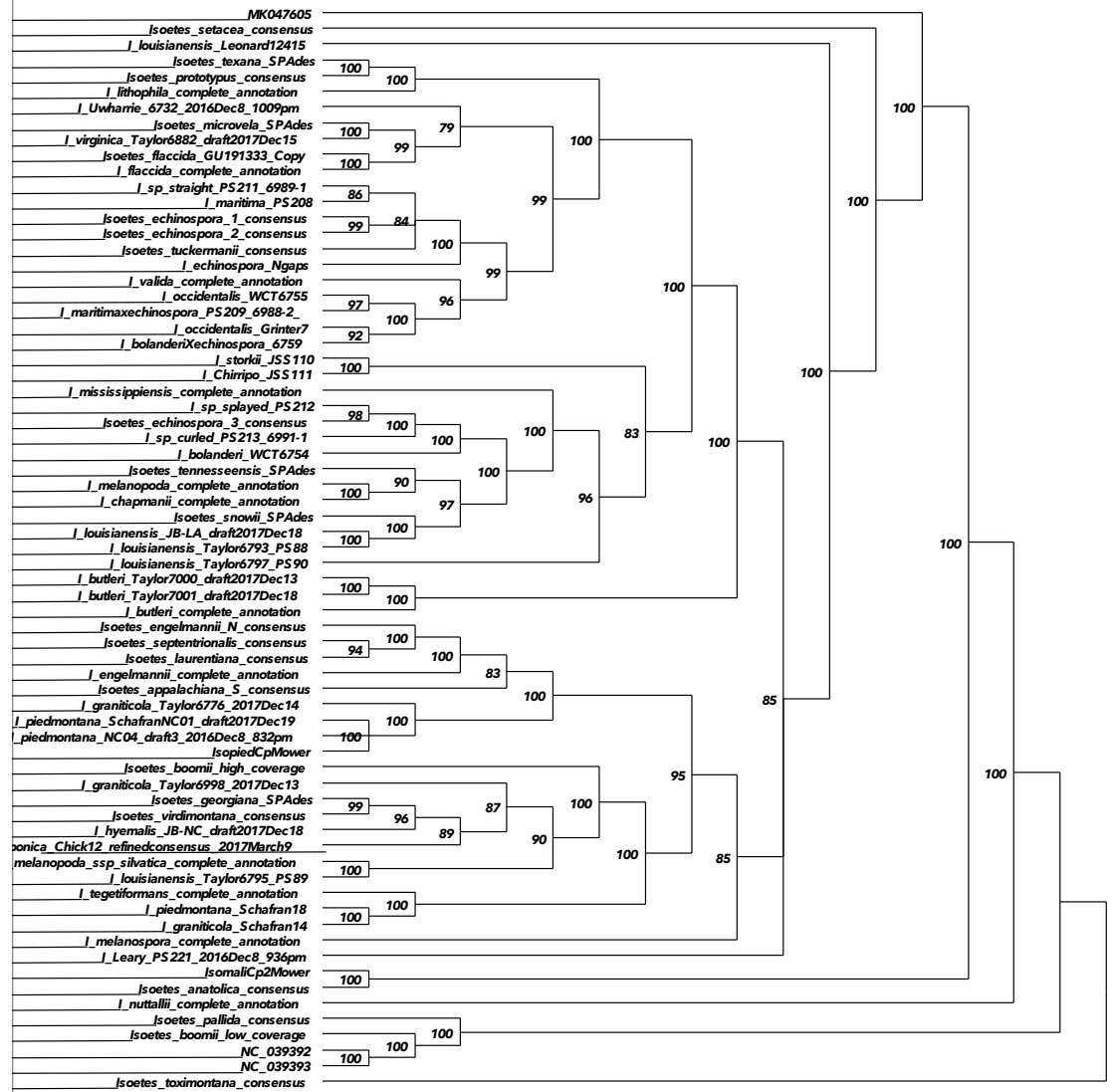
[1]Tree1

ASTRAL, 5kbp segments



[2]Tree1

RxML, concatenated plastome

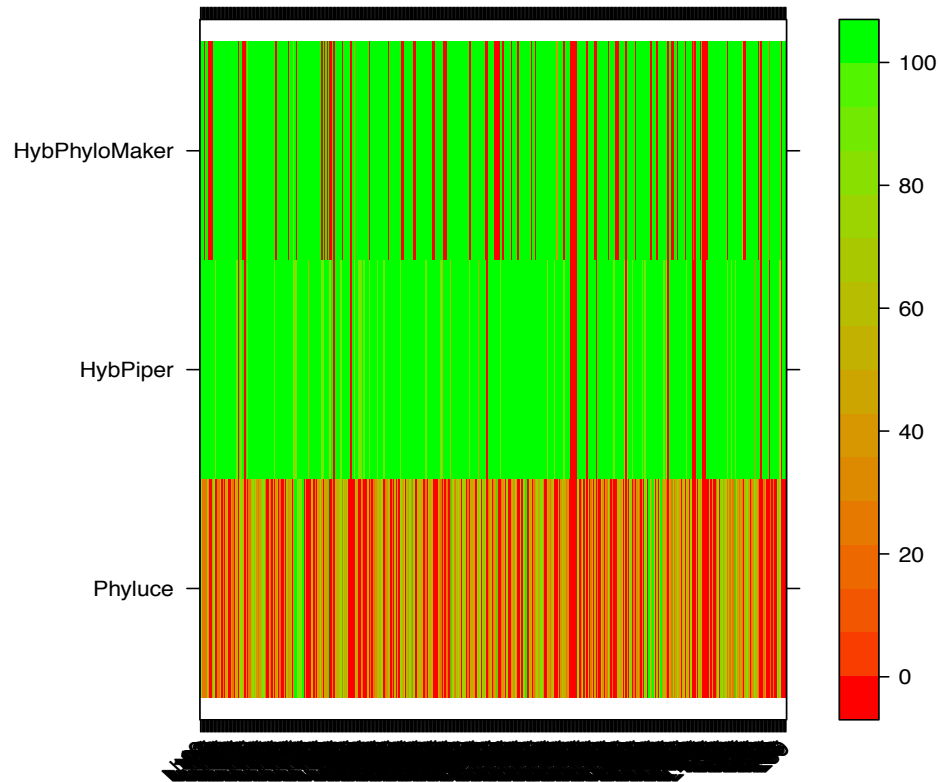


Branches w/ support < 75 bootstrap/0.75 PP collapsed

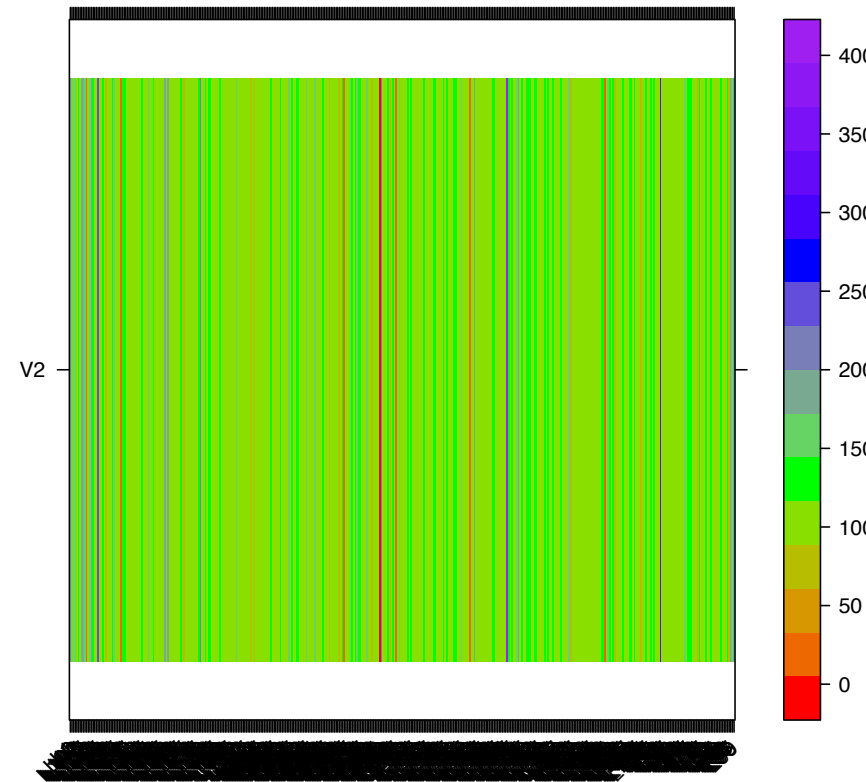
Analysis of Target-Enrichment Data

Sequence data from 11 diploids

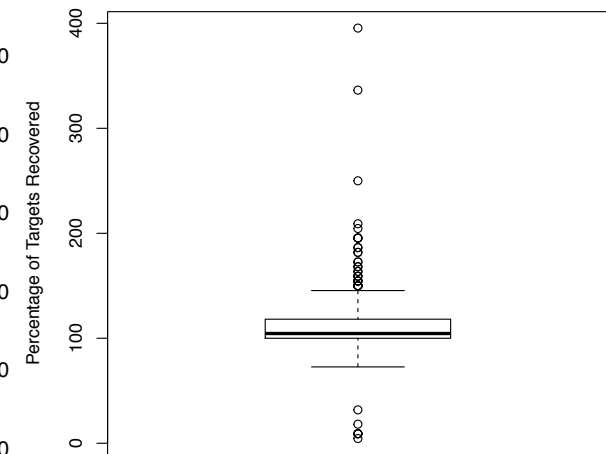
Percent Recovery of Targets



GoFlag Percent Recovery of Targets

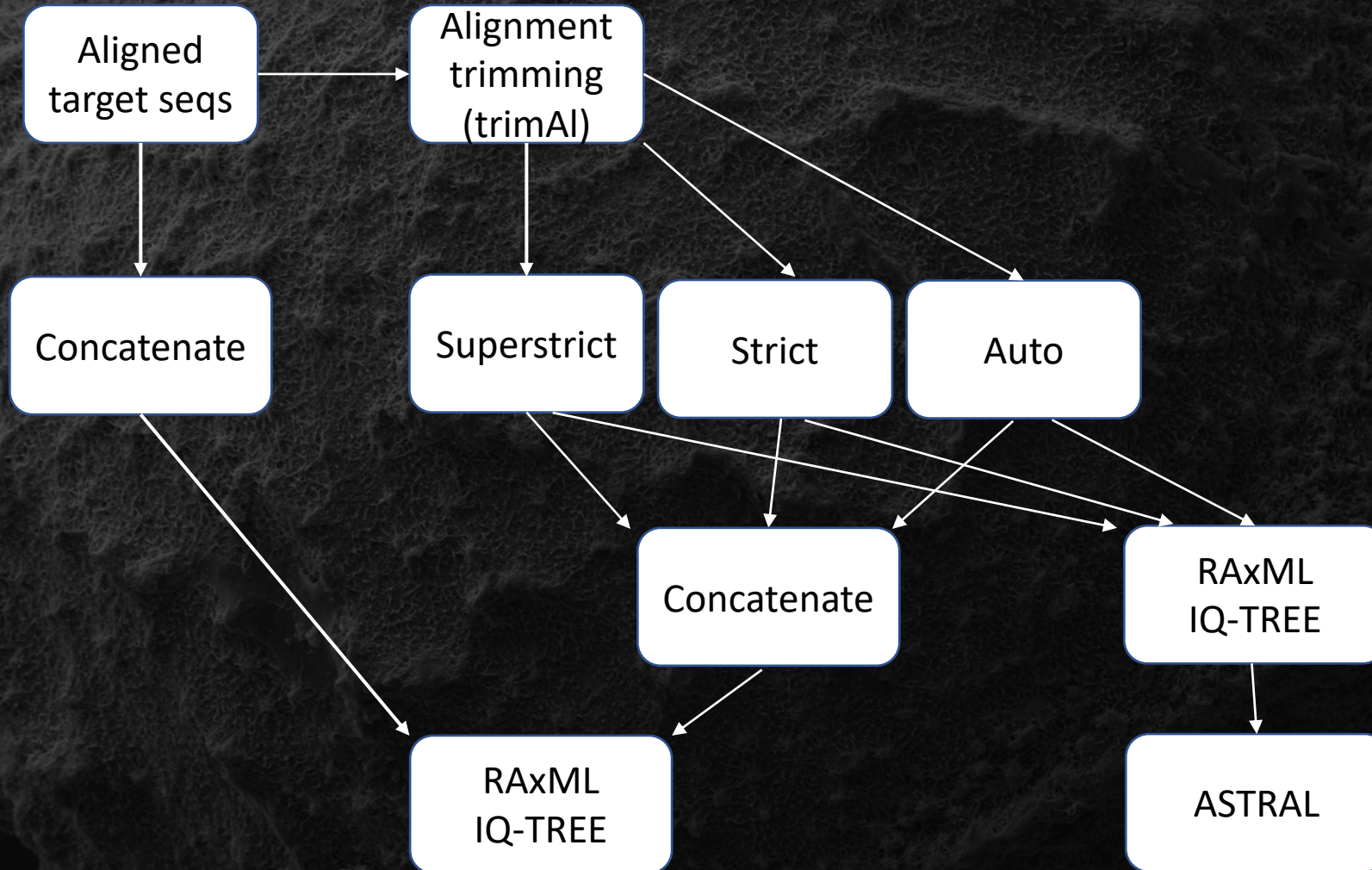


GoFlag Percent Recovery of Targets



GoFlag potentially detecting additional paralogs, but HybPiper reports many fewer

Analysis of Target-Enrichment Data



Tree Comparisons

Pairwise Robinson-Foulds Distance – PAUP*

	tree	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
Complete	1																
	2	9															
	3	6	13														
	4	7	6	11													
	5	6	11	8	9												
	6	5	8	11	6	9											
	7	6	9	12	9	10	3										
	8	9	4	13	4	11	6	7									
Incomplete	9	8	9	12	9	10	9	6	9								
	10	9	10	13	10	11	8	7	10	5							
	11	8	9	12	9	10	7	4	7	2	7						
	12	9	10	11	10	9	8	5	10	5	4	7					
	13	8	9	12	9	10	9	8	9	2	5	4	7				
	14	9	10	11	10	9	8	5	10	5	4	7	0	7			
	15	10	11	14	11	8	9	6	9	6	11	4	9	8	9		
	16	9	10	11	10	9	8	7	10	7	2	9	2	7	2	11	
		COMPLETE								INCOMPLETE							
		AUTO	MAFFT	STRICT	UPERSTRIC	AUTO	MAFFT	STRICT	UPERSTRIC	AUTO	MAFFT	STRICT	UPERSTRIC	AUTO	MAFFT	STRICT	UPERSTRIC
		C	A	C	A	C	A	C	A	C	A	C	A	C	A	C	A

Average pairwise distances:

Complete data vs. complete: 8.88

Complete data vs. incomplete: 9.29

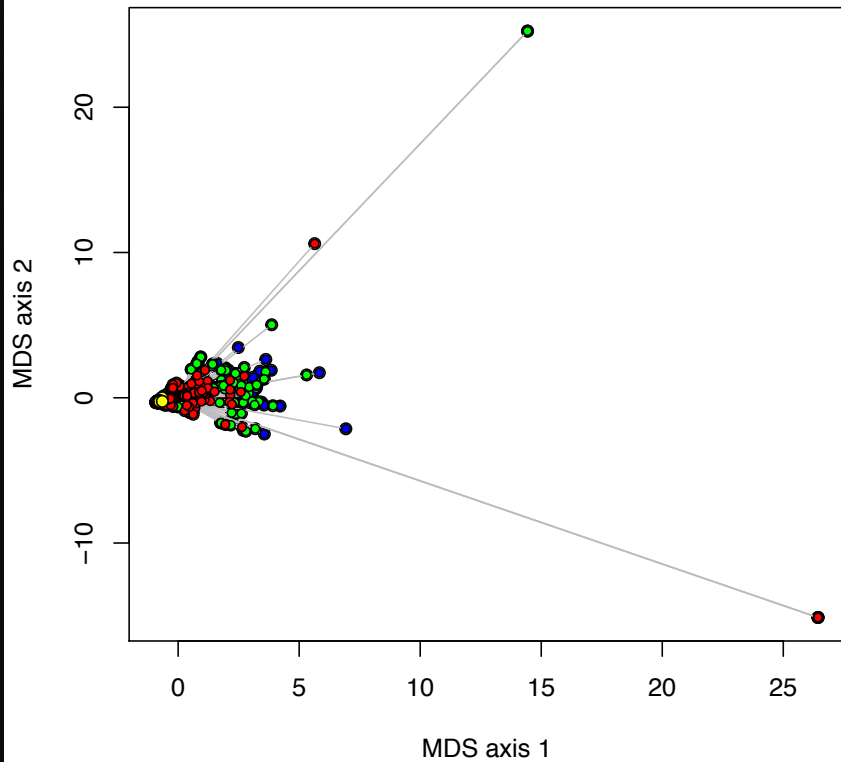
Incomplete data vs. incomplete: 7

Trees from unfiltered, concatenated dataset with greatest RF distances

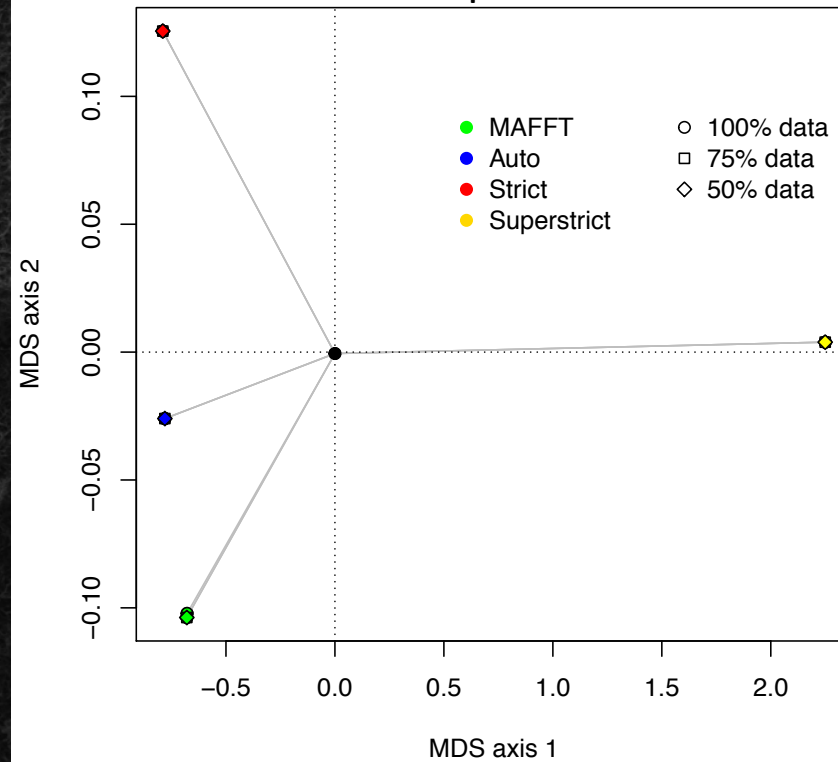
Tree Comparisons

Multidimensional Scaling – R code from Liam Revell

Individual loci by trimming parameters



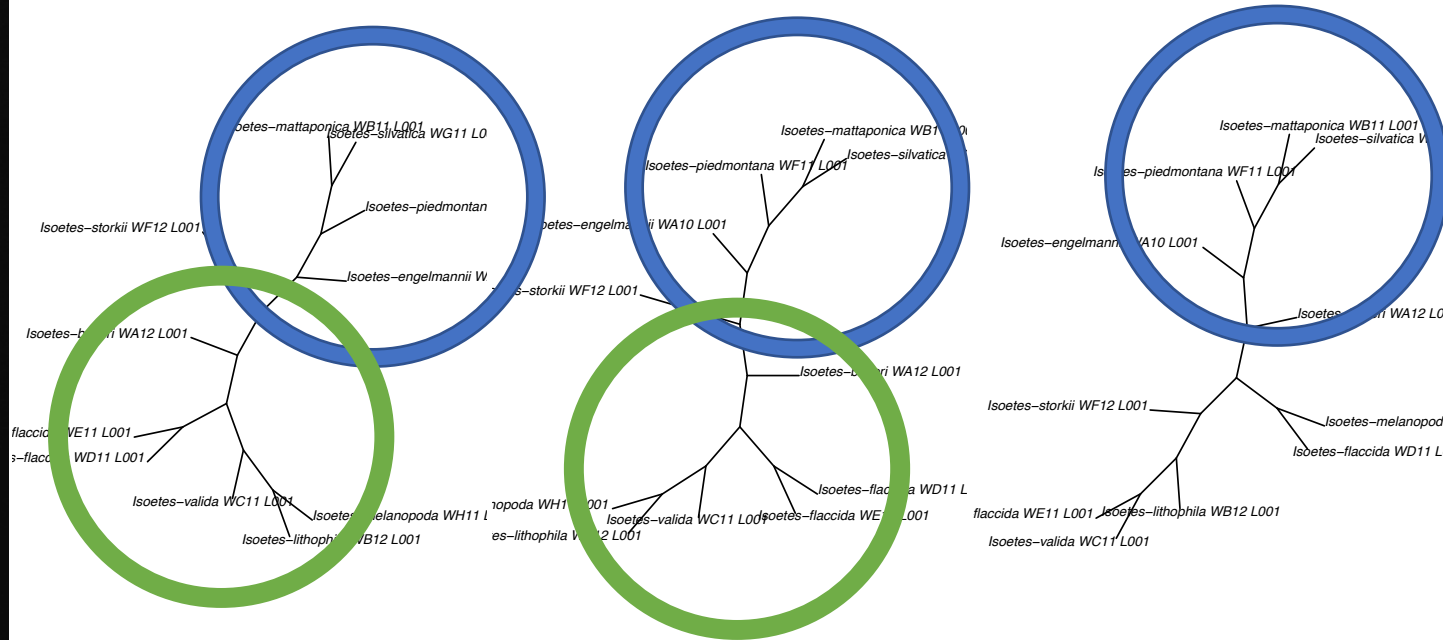
ASTRAL trees by trimming parameters and data completeness



Alignment trimming seems to have strongest effect on tree topology

Seeing the forest for the trees

Superstrict concat vs. majority consensus ASTRAL vs. plastome



4 species consistently resolved together:
I. engelmannii
I. piedmontana
I. silvatica
I. mattaponica } Similar distribution and morphology

Potential chloroplast capture in *I. chapmanii*:
 Plastome sister to *I. melanopoda*, ~99% less distant than other tip pairs

Nuclear position sister to *I. flaccida*, shortest distance in tree (~80% shorter than average)

In progress...allele phasing

Assumption that neotetraploids function like diploids at single-copy loci

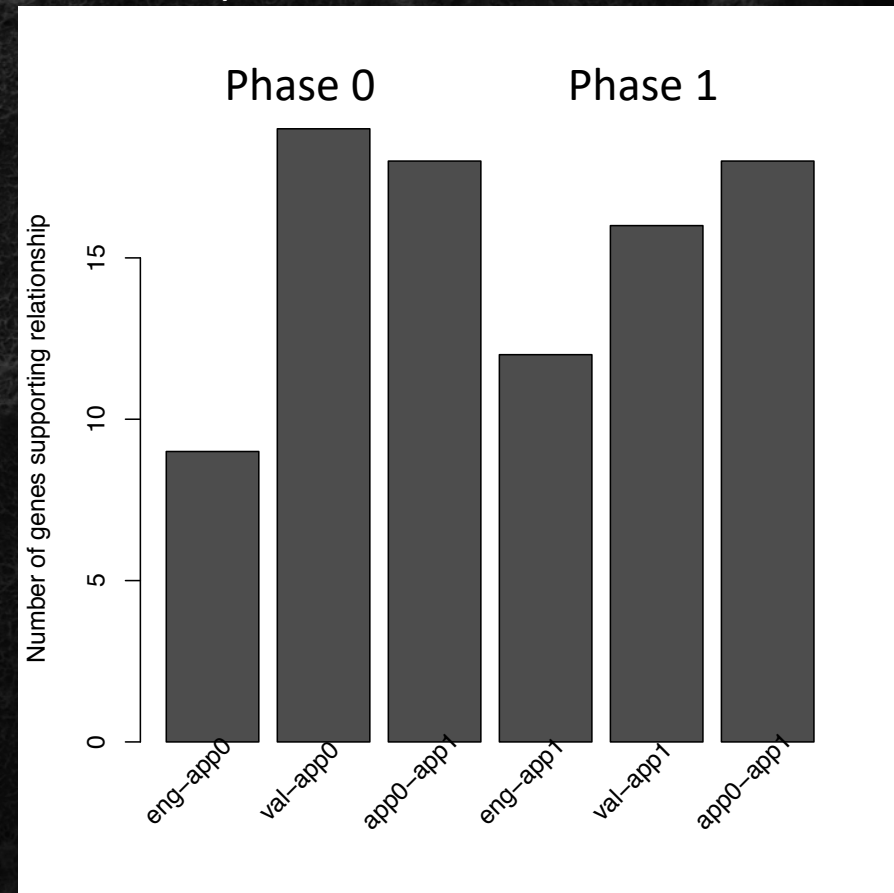
Illumina reads phased with samtools, loci reconstructed with HybPiper

Gene phylogenies constructed, distances from phased copies to putative diploid parents used to test parentage

I. appalachiana (4x) parentage known from Sanger seq of LFY intron 2

I. engelmannii X *I. valida*

Low # of loci supporting (eng,app) conflicts with plastome showing *I. engelmannii* as maternal parent



Acknowledgements

Lab

Gabe Johnson

Maggie Halloran

Monica Carlsen

Carl Rothfels

Fay-Wei Li

Chris Tyrell

Jacob Suissa

Funding/Support

Mary Payne Hogan
Endowment (ODU)

Smithsonian Institution
Fellowships

American Society of Plant
Taxonomists

Washington Field Biologists
Club

NSF DEB 1541506 (GoFlag)

Field

Steve Leonard

Dan Brunton

Alan Cressler

Frankie Snow

Tony Matthews

Mac Alford

Bob O'Kennon

Jason Singhurst

Loran Anderson

Devin Rodgers

Brian Keener

Lesley Starke

Alvin Diamond Jr.

Cameron Byrd

Jacob Fields

Richard Carter

Joey Shaw

Claire Ciafré

and many more...