In the Belly of the Beast: a cophylogenetic study of the pitcher plant flies (*Fletcherimyia*) and their carnivorous hosts (*Sarracenia*)

by

Peter T. Kann

July, 2021

Director of Thesis: Dr. Trip Lamb

Major Department: Biology

SARCOPHAGIDAE, commonly known as the flesh flies, comprises one of the more behaviorally diverse families among the insects. In addition to carrion feeding, sarcophagids have evolved life history strategies that include predation, parasitism, and kleptoparasitism. Most kleptoparasitic species specialize on solitary hymenopterans, but one genus, *Fletcherimyia*, has developed a relationship with unlikely hosts—the North American pitcher plants (*Sarracenia*, Sarraceniaceae). Well known for their carnivory, *Sarracenia* paradoxically supports an ecologically distinct arthropod community, several members of which are obligate associates. For example, *Fletcherimyia* flies undergo larval development exclusively within pitchers and feed on captured insect prey. Eight species are currently recognized within *Fletcherimyia*, all morphologically-defined constructs based largely on genital morphology; the species have yet to be confirmed genetically and have never been placed in any phylogenetic context. Previous studies have characterized *Fletcherimyia-Sarracenia* interactions: five of the eight fly species appear to be restricted to a single host species, whereas the remaining three fly species affiliate with multiple pitcher species. However, fly-pitcher affiliations are largely based on limited observation with narrow geographic scope. The evolutionary history of the *Fletcherimyia-Sarracenia* system as a whole has yet to be addressed.

We conducted the most comprehensive ecological sampling of *Fletcherimyia* to date to 1) examine the status of species constructs; 2) present the first phylogeny for the genus; and 3) conduct a cophylogenetic analysis of the flies and their pitcher hosts. To do so, we generated two molecular datasets (mitochondrial *cox1*, 2bRADseq) for all eight fly species across their respective geographic ranges and hosts. We provide strong molecular support for each species and present the first phylogeny for the genus based on our 2bRAD data, providing evolutionary insight and context to original species descriptions. Our results demonstrate the efficacy of 2bRAD— a recent modification of RADseq protocol—for phylogenetic analysis of recently diverged taxa.

To reevaluate host plant usage, we defined *Fletcherimyia-Sarracenia* interactions by larval presence, as larvae are bound by pitcher deposition whereas adults potentially visit multiple pitcher species. In the absence of diagnostic larval morphology, we typed larval specimens genetically, using *cox1* as a genetic marker for species identification. For cophylogenetic analysis of the fly-pitcher system, we compared a recent phylogeny of *Sarracenia* to our 2bRAD phylogeny. We found evidence for early cospeciation and subsequent host-switching between eastern and western pitcher species, indicating a protracted coevolutionary history between *Fletcherimyia* and *Sarracenia* lineages.

In the Belly of the Beast: a cophylogenetic study of the pitcher plant flies (*Fletcherimyia*) and their carnivorous hosts (*Sarracenia*)

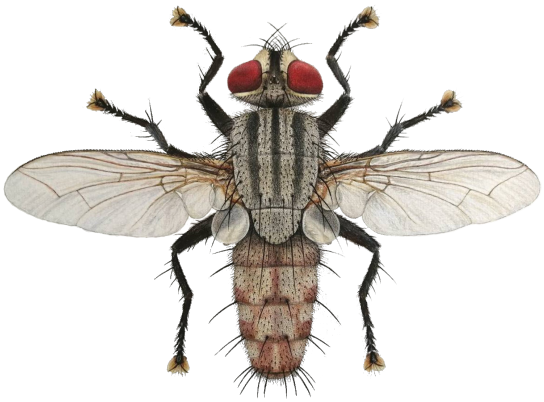
A Thesis

Presented to the Faculty of the Department of Biology

East Carolina University

In Partial Fulfillment of the Requirements for the Degree

Master of Science in Biology



by

Peter Kann

July, 2021

In the Belly of the Beast: a cophylogenetic study of the pitcher plant flies (*Fletcherimyia*) and their carnivorous hosts (*Sarracenia*)

By

Peter T. Kann

APPROVED BY:

Director of Thesis _____

Trip Lamb, PhD

Committee Member _____

Michael S. Brewer, PhD

Committee Member _____

Carol Goodwillie, PhD

Chair of the Department of Biology _____

David Chalcraft, PhD

Dean of the Graduate School _____

Paul Gemperline, PhD

Acknowledgements

I would like to acknowledge the following institutions and people who enabled my research, enriched my mind, and broadened my appreciation for the natural world. Thank you to the Smithsonian National Museum of Natural History's Williston Diptera Fund, the American Museum of Natural History's Theodore Roosevelt Memorial Fund, the North Carolina Native Plants Society's Shinn Grant, and the Garden Club of America's Summer Scholarship in Field Botany for funding this research. Without your generosity, the scope of our field work and the depth of our genetic datasets would not have been possible. Thank you to my committee members Dr. Michael S. Brewer and Dr. Carol Goodwillie for your guidance, your technical expertise, and your teaching. Thank you to my family for your love and support, and occasionally field assistance, while I've been away chasing flies. Thank you to all members of the Brewer Lab: Emily Scott, Dr. Chris Cohen, and Dr. Jeff Cole for keeping me sane and keeping my computer in one piece. Thank you to all members of the M. McCoy Lab: Ellen Titus, Ellie Baker, and Lu Kernstine for relentlessly boosting my morale and assisting in the field. Thank you to Dr. Claudia Jolls, who infected me with a new love and understanding of the botanical world. Thank you to my advisor, mentor, collaborator, and friend Dr. Trip Lamb, who has taught me too many lessons to recount here, but most importantly to appreciate and share the profound beauty that defines nature. Finally, thank you to Sarah Goodnight who, among countless acts of happiness, kindness, and support, drove west in a van full of herpetologists and made these past two and a half years possible.

I would be remiss not to also acknowledge the 811 flies that were collected for this study. Though they were not people, they were unique living members of an extraordinary ecosystem. The knowledge gained from each individual has greatly expanded our understanding of their evolution and natural history, and for that I am forever grateful.

Table of Contents

List of Tables

List of Figures

Introduction

Flesh flies (Sarcophagidae) form a large, cosmopolitan family of ~3,000 described species best known for their namesake sarcophagous larvae and central role as decomposers (Buenaventura et al. 2019). Indeed, the majority of Sarcophagidae feed on carrion or necrotic tissue, accounting for substantive research on medically relevant, economically impactful taxa (Yadong Guo et al. 2014; Meiklejohn, Wallman, and Dowton 2011; Wells, Pape, and Sperling 2001; C. Zhang et al. 2015). Beyond sarcosaprophagy (the consumption of dead or decaying animal tissue), sarcophagids present a wide array of larval feeding habits that include coprophagy, predation, parasitism, parasitoidism, and kleptoparasitism (Buenaventura 2021). These alternative life histories generally entail close ecological if not truly symbiotic relationships with associate taxa. Most often sarcophagid relationships involve other arthropods: orthopterans, mantids, beetles, cicadas, and scorpions are routinely targeted by parasitoids representing some eight genera across Sarcophagidae (Pape 1994; Stucky 2015). To a lesser extent, terrestrial gastropods (Reeves, Pape, and Adler 2000) and even some vertebrates (Mulieri et al. 2018) serve as hosts.

Kleptoparasitism ("stealing" the food of other organisms) is similarly widespread, particularly in the subfamily Miltogramminae, where most species exploit solitary wasps or solitary bees. Flies either target food-bearing hosts in flight or enter brood chambers to deposit larvae on paralyzed prey. Miltogrammines depart notably from the typical kleptoparasitic interaction in which parasite and host—sharing food preferences—are closely related, i.e., congeneric or confamilial (Iyengar 2008). Perhaps the most remarkable (certainly the most phylogenetically distant) case of sarcophagid kleptoparasitism involves a rare instance in the subfamily Sarcophaginae: a genus that associates not with another animal host, but with a genus of carnivorous plants.

The pitcher plant genus Sarracenia *and arthropod associates*

The North American pitcher plants (*Sarracenia*; Sarraceniaceae) are a recently evolved assemblage of 11 carnivorous species distributed across boggy habitats of the eastern United States and Canada (Stephens et al. 2015). Obtaining nutrients via carnivory, *Sarracenia* species possess modified tubular leaves (pitchers) that serve as passive traps lined with slick wax and downward facing hairs to capture insect prey. Once trapped, insects are digested by a microbial community inhabiting the pitcher and by the plant's endogenous enzymes, though the latter have not been documented in all species of *Sarracenia* (Bradshaw and Creelman 1984) and are demonstrably less complex than those found in the convergently evolved pitchers of the Old World tropical genus *Nepenthes* (Adlassnig, Peroutka, and Lendl 2011). Antithetical to the pitcher's function, several arthropod lineages have evolved to exploit the plants' interior cavities and insect prey (Jones 1904), and recent studies suggest they form ecologically distinct communities (Miller, Bradshaw, and Holzapfel 2018; Satler and Carstens 2019). Though levels of ecological dependency vary, many of these inquilines (organisms that utilize the living spaces of others) are obligate pitcher associates, requiring the plant to complete their life cycles. Obligate affiliates include various moths (*Exyra* and *Papaipema*, Noctuidae), aphids (*Macrosiphum*, Aphidae), mites (*Sarraceniopus*, Histiostomatidae; *Macroseius*, Phytoseiidae), and an exceptional diversity of flies (*Wyeomyia*, Culicidae; *Metriocnemus*, Chironomidae; *Fletcherimyia* and *Sarcophaga*, Sarcophagidae; *Bradysia*, Sciaridae; *Aphanotrigonum* and *Tricimba*, Chloropidae) (Lamb and Kalies 2020; D. R. Folkerts 1999; Mlynarek and Wheeler 2018). Other groups of moths and flies are casually affiliated with *Sarracenia*, as are various wasps, bees, and spiders. This list is merely an overview, not an exhaustive description of the biodiversity of the inquiline community.

The sarcophagid genus Fletcherimyia

For most dipteran genera, just one (or two) species are pitcher associates; only in *Fletcherimyia* is the entire genus obligately affiliated. *Fletcherimyia* currently contains eight species, described primarily on the basis of genital morphology (Dahlem and Naczi 2006). Like other sarcophagids they are ovolarviparous (Meier, Kotrba, and Ferrar 1999). Females retain hatched larvae through their first instar then deposit them directly into *Sarracenia* pitchers, where larvae live and feed on prey captured by the plant (Rango 1999). While several larvae may be deposited into a single pitcher, commonly only one survives to pupariation except in cases of particularly abundant prey (Dahlem and Naczi 2006; D. R. Folkerts 1999), likely due to competitive cannibalism. Given their dependency on prey captured by *Sarracenia*, these flies have been classified as kleptoparasites (Buenaventura 2021); despite these feeding habits, larvae do not appear to rob their hosts of nutrients and may actually increase nutrient availability (Underwood 2009). As such, the *Fletcherimyia-Sarracenia* relationship may be better described as kleptobiotic rather than kleptoparasitic (Vollrath 1984).

Mature larvae either crawl out of the pitcher or chew a distinctive hole near its base before pupariating in the surrounding soil. Adult flies emerge and though seemingly no longer bound ecologically to the pitchers, tend to remain in close proximity nonetheless. Adults routinely perch on pitcher leaf sides and opercula (Wray and Brimley 1943) (Figure 1) and have been anecdotally recorded roosting within flowers (Krawchuk and Taylor 1999). On numerous occasions we have observed different species of *Fletcherimyia in copula* on the sides of pitchers, which may be emblematic of the adults' reproductive impetus to stay near their host plants. The flies' dependence on *Sarracenia* carries with it a conservation concern: roughly 97% of the original range of *Sarracenia* has been lost to anthropogenic activity (G. W. Folkerts 1982); six species are on the IUCN Red List, two of which are federally listed as critically endangered. Their obligate insect associates are almost certainly likewise imperiled.

The most recent taxonomic treatment of the genus *Fletcherimyia* (Dahlem and Naczi 2006) recognizes eight species: *F. fletcheri* Aldrich, *F. rileyi* Aldrich, *F. jonesi* Aldrich, *F. celarata* Townsend, *F. abdita* Pape, *F. oreophilae* Dahlem and Naczi, *F. papei* Dahlem and Naczi, and *F. folkertsi* Dahlem and Naczi. Dahlem and Naczi's (2006) account provides a detailed review of the taxonomic history of *Fletcherimyia*, which we summarize here. The foundational entomologist Charles Valentine Riley described the first species of pitcher plant fly, *Sarcophaga sarraceniae*, in his 1874 publication "Descriptions and natural history of two insects which brave the dangers of *Sarracenia variolaris*" based on a number of specimens. Upon revisiting Riley's type material, Aldrich (1916) realized the type series contained three species: *S. sarraceniae* as well as two new species (*S. jonesi* and *S. rileyi*) which Aldrich subsequently described from Riley's specimens and new field collections. Aldrich also described *S. fletcheri* and *S. celarata*, which soon thereafter were assigned to two new monotypic genera, *Fletcherimyia* and *Peltopygia*, respectively (Townsend 1917). Roback (1954) redescribed *Fletcherimyia* to include the four species *F. fletcheri*, *F. rileyi*, *F. jonesi*, and *F. celarata*, though his incarnation of the taxon was subsumed as a subgenus within *Blaesoxipha* Loew by Downes (1965). *Blaesoxipha* is commonly recovered as being closely related to *Fletcherimyia* in molecular phylogenetic analysis (Stamper et al. 2013; Buenaventura 2021), so Downes' subsumption was not without merit. However, it also met with disagreement, notably by Lopes (1971), who had revisited *Fletcherimyia* prior to Roback (1954). Shewell (1987). Pape (1990) also argued that *Fletcherimyia* should be restored to full generic status, and based on paratypes of *F. rileyi*, described the fifth species, *Fletcherimyia abdita*. Finally, Dahlem and Naczi's (2006) revision provided a thorough taxonomic literature review and detailed various historical misidentifications. Most helpfully, they also included a species key, with diagnostic illustrations of male genitalia and female genital sternites. Their revision includes descriptions of three new species: *F. oreophilae*, *F. papei*, and *F. folkertsi*.

Fly-pitcher coevolution and cophylogenetic analysis

Five species of *Fletcherimyia* reportedly utilize a single host species of *Sarracenia* whereas others use two, or more, host species (Dahlem and Naczi 2006). Host ranges vary in continuity and sympatry with other pitchers (Stephens et al. 2015), and ranges of *Fletcherimyia* vary accordingly. Table 1 summarizes relationships between flies and pitcher hosts and details fly ranges as they are currently known (Dahlem and Naczi 2006). Given the obligate ecological assoaciation of *Fletcherimyia* with *Sarracenia*, there exists potential for different forms of coevolutionary expression (e.g., host switching, cospeciation; Cruaud and Rasplus, 2016).

Cophylogenetic analysis, the study of phylogenetic congruence among interacting lineages, is a useful tool for inferring patterns of codivergence and determining the evolutionary impact of hosts on associates (Merkle and Middendorf 2005). Most work in this field has concentrated on codivergence in isolated relationships like host plants and their obligate seed-parasites, famously including the mutualism seen between figs and fig wasps (Cruaud and Rasplus 2016; Marussich and Machado 2007). Classic predictors of phylogenetic congruence - vertical transmission and host-specificity - are more easily identified in these two-party systems. Less effort has been directed to resolving patterns of codivergence among members of dependent *communities* that share hosts but may vary in ecological dependency and dispersal. The *Fletcherimyia-Sarracenia* system then provides an opportunity for cophylogenetic analysis in a relatively unstudied context.

Comparisons of phylogenies of associated species allow the testing of cospeciation hypotheses and evaluate the evolutionary impact of a host (i.e., independent) lineage on a parasitic (i.e. dependent) one. These analyses are categorized into two broad groups: event-based and global-fit (distance-based) comparisons (Cruaud and Rasplus 2016). Event-based analyses map dependent (parasite, inquiline) trees onto independent (host) trees, searching for the most parsimonious series of coevolutionary events that led to existing topologies.

The events examined fall into five categories: cospeciation, failure to diverge, host loss/sorting, duplication, and host switch (Keller-Schmidt et al. 2011) (Figure 2). Cospeciation occurs when host and parasite lineages experience simultaneous cladogenetic events (Figure 2A); this is not necessarily due to coevolutionary forces and is often a product of vicariance (Cruaud and Rasplus 2016). When the parasite does not speciate simultaneously with its host and instead remains one large population that affiliates with both daughter lineages of a host speciation event, it is considered a failure to diverge (Figure 2B). Host loss, also referred to as sorting, occurs when the host speciates but the parasite does not and the parasite fails to remain associated with both of the resulting host species (Figure 2C, D). Duplication, a speciation event in the parasite lineage but not the host lineage, can either result in two daughter parasite taxa affiliated with the same host (Figure 2C) or can be associated with a host switch (Figure 2D); during host-switching, one daughter parasite lineage "jumps" to a new host lineage and is isolated from its ancestral population. Analysis software packages assign costs to these events, then search for evolutionary histories with the lowest overall cost (Balbuena, Míguez-Lozano, and Blasco-Costa 2013). Alternatively, global-fit methods quantify the congruence of two existing trees (and sometimes identify particular regions that contribute more to that congruence value (Cruaud and Rasplus 2016)) but do not take individual evolutionary events into account. In other words, global-fit methods do not attempt to recreate evolutionary history.

Given the *Sarracenia-Fletcherimyia* system, event-based analysis seems to be the more appropriate choice. However, increasingly complex relationships between dependent and independent taxa dramatically increase computational intensity and possible outcomes, making results difficult to interpret (Desdevises 2007). The existing complexity of the *Fletcherimyia*/*Sarracenia* relationships (i.e., single fly species associating with multiple pitcher species) could compromise an event-based analysis and, thus, fail to resolve cophylogenetic patterns with sufficient clarity. Alternatively, global-fit analyses, which have been shown to

effectively recover patterns of coevolution (Desdevises 2007), could also be used to supplement potential constraints posed by event-based methods.

A cophylogenetic study of the fly/pitcher system requires two critical pieces (yet to be ascertained) of phylogenetic information for the obligate associate (dependent) lineage: 1) a monophyletic *Fletcherimyia*, with appropriate species delimitation, and 2) a robust, well-supported phylogeny. Both monophyletic confirmation and molecular phylogeny are in place for the host *Sarracenia* (Stephens et al. 2015)—its species the products of a rapid diversification dating to ~3 mya. We presume *Sarracenia* diversification shaped evolution history and cladogenesis in *Fletcherimyia* concomitantly during this brief time frame.

We chose the mitochondrial gene cytochrome oxidase 1 (*cox1*) for a preliminary molecular phylogenetic assessment of species identification and relationships in *Fletcherimyia*. Although a rapidly evolving mitochondrial gene such as *cox1* would be considered appropriate for examining molecular genetic variation in *Fletcherimyia*, phylogenetic signal from a single mitochondrial gene could be compromised by introgression and/or incomplete lineage sorting (ILS); both are factors which complicate the phylogenetic relationships within *Sarracenia*. The recent and rapid nature of the *Sarracenia* radiation likely resulted in higher rates of incomplete lineage sorting and, therefore, fewer phylogenetically informative regions of the genome (Maddison and Knowles 2006). All species within *Sarracenia* can also hybridize with one another (Furches, Small, and Furches 2013) and often do so in sympatry, potentially causing gene/species tree discordance. Problematic features like introgression, which stand to produce errant phylogenies, can reveal interesting complexities when mtDNA data is examined in combination with nuclear genes.

We also chose to examine fly species constructs and their relationships with nuclear sequence capture using 2bRADseq (Wang et al. 2012), which has proven successful in defining species limits and recovering relationships in closely related taxa (Kelly and Thacker 2020; Manzello et al. 2019). 2bRAD was shown to be more effective than *cox1* in supporting species

monophyly within the poriferan genus *Ircinia* (Kelly and Thacker 2020), which exhibits high rates of hybridization. In a 2bRAD seq survey of reef-building corals (*Orbicella*), Manzello et al. (2019) not only delimited species satisfactorily but demonstrated monophyly through technical replicates of individual specimens (i.e., the same specimen sequenced multiple times). Simulated 2bRAD datasets were shown to recover accurate phylogenies of 21 *Drosophila* species representing two subgenera (Seetharam and Stuart 2013) demonstrating the potential of 2bRADseq in dipterans as well as for inferring deeper phylogenetic relationships.

2bRADseq is a recent modification of the original RADseq protocol (Davey and Blaxter 2010). It is a form of Reduced Representation Sequencing (RRS), a subsampling technique that captures variation needed for phylogenetic analysis without sequencing the entire genome, thereby reducing time and cost. Restriction enzymes cleave DNA at recognized sites and sequence is captured along those sites. In the 2bRADseq protocol, a single IIb-type restriction endonuclease is introduced to the template material; unlike most endonucleases, they do not cut at their recognition site, but ~10bp in either direction of the site. This results in a 21-36 bp fragment, depending on the specific enzyme used. As is the case with all RADseq methods, reference genomes are not required (Puritz et al. 2014), making 2bRAD particularly well suited for non-model organisms like *Fletcherimyia*. 2bRAD is uniquely reproducible among RADseq techniques (Wang et al. 2012; Robledo et al. 2018); therefore, any potential future additions to the dataset (e.g., additional outgroup taxa, specimens from new localities) should be directly comparable to the initial samples.

Here we report on mitochondrial *cox1* and 2bRAD sequence data generated for *Fletcherimyia*, with broad geographic and host plant coverage of all eight nominal species. We use these molecular data to 1) examine the status of morphological species constructs; 2) present the first phylogeny (molecular or otherwise) for the genus; and 3) conduct a cophylogenetic analysis of the flies and their pitcher hosts. Coupled with updated fly/pitcher relationships based

on larval associations, these results form the basis of a *Fletcherimyia-Sarracenia* cophylogeny which reveals patterns of codiversification and host switching between genera.

Methods

Taxon Sampling

To examine intra- and interspecific genetic variation in *Fletcherimyia,* we pursued a sampling regime to collect the eight nominal species throughout their geographic ranges as well as from all respective pitcher hosts. Adult flies were collected with hand nets whereas larvae were extracted from pitcher fluid/prey mass either by *in situ* pitcher dissection or use of a siphoning device (Rango 1999). Host plant species were recorded for all specimens. This sampling effort yielded 774 identified specimens (263 male, 100 female, 411 larval) representing 91 localities (Figure 4). Adults were euthanized (ethyl acetate inhalation) and their abdomens excised and pinned to reveal terminalia for species identification. At localities where adult flies were abundant, additional adult specimens were collected and pinned whole for supplemental voucher material. Adults were identified using keys and illustrations in Dahlem and Naczi (2006). Specimens (adult and larval) slated for genetic analysis were preserved in RNAlater, left at 4°C for a minimum of 24 hours to allow tissue saturation, then frozen at -80°C until further processing.

DNA Preparation and mtDNA Sequencing

DNA was extracted from head and thoracic tissue of adults and, considering size, the anterior portion to whole specimens of larvae using Qiagen's DNeasy Blood and Tissue Kit. following manufacturer's instructions. A 10µL aliquot of eluted DNA was separated for mtDNA amplification, the remainder being stored at -80°C for subsequent genomic assay. We chose the mitochondrial gene cytochrome oxidase I (*cox1*) as an initial genetic marker for 1) larval species identification, 2) examining population variation and phylogeographic structure, and 3) assessing morphological species constructs. We amplified the entire cox1 gene using the primers 5'-GATTTACAGTCTATTGCCTAAATTTC-3' and 5'-GCTTAAATCCATTGCACTAATCTG-3,' which are modified versions of Simon et al.'s (1994) (Simon et al. 1994) conserved insect primers TY-J-1460 and TL2-N-3014, respectively - altered according to mitogenomic sequences from selected

sarcophaginine genera. We shipped PCR products of 175 samples to Psomagen (Psomagen, Maryland) for bi-directional Sanger sequencing. Raw sequences were trimmed and merged into contigs using *Sequencher* ("Sequencher DNA Sequence Analysis Software from Gene Codes Corporation" n.d.) and exported into FASTA format. Sequences were imported into *Mesquite* and aligned with MAFFT using default parameters.

2bRADseq and Bioinformatics Pipeline

DNA samples for 116 specimens (including several larval samples identified by *cox1* sequencing) were shipped to CD Genomics (CD Genomics, New York) for 2bRADseq. Quality control, library prep, sequencing, and data preprocessing (including quality filtering and demultiplexing) were performed using the established protocol (Wang et al. 2016). Paired-end sequencing was performed using the Illumina HiSeq PE150 platform. Reverse read files were reverse-complemented and merged with forward reads using *PEAR* (J. Zhang et al. 2014) and subsequently treated as single-ended.

Demultiplexed, merged sequence files were processed using the software package *ipyrad* (Eaton and Overcast n.d.), a user-friendly pipeline developed specifically for RADseq data and phylogenetic applications. This software package demultiplexes, clusters, and aligns raw Illumina read data to prepare it for immediate use by most mainstream tree building programs. Assembly parameters were kept at default values with the following exceptions: 1) datatype set to 2bRAD 2) minimum sequence length lowered to 20 to account for short read lengths typifying 2bRAD, and 3) minimum samples per locus; we made three separate alignment matrices with 2, 4, and 28 minimum samples per locus to see the effect on support values and missing data. Phylip, Nexus, and STRUCTURE output formats were generated for tree-building software.

Inferring Phylogenies

Outgroups - We consulted three recent sarcophagid phylogenies [terminalia morphology (Buenaventura and Pape 2018), anchored hybrid enrichment (Buenaventura et al. 2019), and

ultraconserved elements (Buenaventura 2021)] to select appropriate outgroups. These included *Mecynocorpus*, *Titanogrypa*, *Comasarcophaga*, *Blaesoxipha*, and *Spirobolomyia* - genera which, with *Fletcherimyia*, constitute the *Blaesoxipha* clade within Sarcophaginae (Buenaventura and Pape 2018). However, barriers to tissue loans from museum collections (due to COVID-19 restrictions) limited us to our own sarcophagid "bycatch" tissues. We chose two adult male specimens of *Blaesoxipha* (*Acanthodotheca*) and three *Sarcophaga sarraceniae*, a more distant sarcophaginine outgroup.

cox1 - Maximum Likelihood and Bayesian species trees were inferred using *IQTREE* (Nguyen et al. 2015) and *BEAST2* (Bouckaert et al. 2019), respectively, to compare tree topology and support across methods. We ran IQTREE with ModelFinder to determine the best substitution model; TIM2+F+I+G4 was selected according to AIC, AICc, and BIC. IQTREE ran 1000 ultrafast bootstrap replicates to generate nodal support values; nodes with bootstrap support scores (BSS) above 80% are considered to be well supported.

We used *PartitionFinder* v.2.1.1 (Lanfear et al. 2017) to find the optimal partitioning scheme for the Bayesian analysis; a partitioning scheme based on codon position under the GTR+I+G+X model was selected. We ran three independent BEAST runs using the GTR+I+G+X model; each MCMC chain ran for 50 million generations and was sampled every 1000 generations with 25% burn-in. Runs were combined using *LogCombiner* (Drummond and Rambaut 2007) and parameter effective sample sizes were evaluated using *Tracer* v.1.7.2 (Rambaut et al. 2018) to ensure each ESS was above the cutoff of 200. *TreeAnnotator* summarized the combined trees from the three independent runs and inferred a maximum clade credibility tree. We considered nodes with a posterior probability (PP) above 95% to be well supported.

2bRAD - Maximum Likelihood trees were again inferred using *IQTREE* (Nguyen et al. 2015). ModelFinder selected the TVM+F+I+G4 model according to AIC, AICc, and BIC. IQTREE ran 1000 ultrafast bootstrap replicates to generate nodal support values; nodes with BSS above 80%

are considered to be well supported. Of note, most samples (86.7%) failed the χ^2 composition test; this is likely due to the unique structure of 2bRAD reads, which are quite short and all contain the type-IIb restriction enzyme recognition site.

Divergence Dating

Divergence time estimates for *cox1* were calculated using BEAST2.5 (Bouckaert et al. 2019) under the same partitioning scheme based on codon position under the GTR+I+G+X model. Five independent runs were conducted, each with a chain length of 200 million generations, sampling every 1000. Again, runs were combined using LogCombiner and evaluated for parameter ESS >= 200 using Tracer. Divergence estimates were based on a *cox1* substitution rate of 0.01803 sub/s/My/l, averaged from Hawaiian katydids and *Carabus* beetles (Lamb et al. 2018) in light of little fossil evidence or molecular clock estimates for Sarcophagidae (Stevens and Wallman 2006, 1). All rates in the GTR model were estimated based on empirical frequencies under a strict molecular clock. Haploidy was specified and site and clock models were linked across codon positions.

Cophylogenetic Analysis

Global-fit and Event-based analyses were performed on pruned host (*Sarracenia*) and parasite (*Fletcherimyia*) trees. The host tree was a recent MP-EST species phylogeny of *Sarracenia* (Stephens et al. 2015) generated using target enrichment from multiple accessions; we pruned this tree using the `drop.tip` tool in *ape* (Paradis and Schliep 2019) to collapse subspecies assemblages into single tips unless those subspecies were paraphyletic (e.g. *S. purpurea* spp, *S. rubra* spp). Parasite topologies were inferred from our 2bRAD data (discussed below) and were also pruned to a single tip per species with the exception of *F. papei*; geographically structured genetic divergence congruent across nuclear and mitochondrial datasets rivaled or exceeded that of formally recognized species (Figures 5, 6), therefore these populations were treated separately. Pitcher/fly relationship matrices were based on observed usage of host plants by larvae,

summarized as a tanglegram in Figure 9. Due to low support for various nodes on the *Sarracenia* phylogeny (Figure 10), multiple host topologies were used in this analysis. Nodes with less than 70% support were collapsed into polytomies; alternatively, several fully-resolved trees with varying placements of species with poorly-supported nodes were also used for comparison. Event-based cophylogenetic analysis was performed using *Jane 4.0* (Conow et al. 2010), a tool that assigns costs to coevolutionary events to find optimal explanations of evolutionary history. We used the default cost scheme (cospeciation = 0, duplication = 1, duplication with host switch = 2, loss = 1, failure to diverge = 1) and ran simulations for 2,000 generations with a population size of 4,000 following recommendations from the original publication (population size >= 2*#generations). Though Jane's successor *eMPRess* (Santichaivekin et al. 2020) has been released, it currently lacks compatibility for parasites mapped to multiple host tips, which is necessary for accurate assessments of *Fletcherimyia*/*Sarracenia* given the relationships seen in the literature and in field sampling. Global-fit assessment of tree congruence was performed using *PACo* (Hutchinson et al. 2017), an analysis package in R explicitly capable of quantifying tree congruence when dependent tips share interactions with more than one host tip. *PACo* uses a Procrustes analysis to assess overall tree topology and the degree to which parasite lineages map onto linked host lineages, thereby determining the amount of mirroring between trees. Tanglegrams illustrating pitcher/fly species affiliations for use with *PACo* were generated using the R package *dendextend* (Galili et al. 2019), which allows for visually intuitive labeling of these affiliations (e.g., grouping by fly species). Larger tanglegrams were generated using TreeMap3.

Historical Biogeography

We used the R package *BioGeoBears* (Matzke 2012) under a DEC+J model to infer historical ranges for internal nodes of the *Fletcherimyia* phylogeny, using the same pruned tree referenced above. Species were sorted into three geographic areas: "Gulf" (TX, LA, MS, AL, FL west of the Apalachicola River, GA west of the Fall Line), "East" (FL east of the Apalachicola River, GA east of the Fall Line, piedmont and coastal plain of Carolinas), and "Mountain" (Blue Ridge Mountains

in extreme western Carolinas, northern GA, and northeast AL). “Eastern” and “Western” areas were delimited based on *Fletcherimyia* species presence, genetic breaks present in *F. papei*, and known plant hybrid contact zones (Swenson and Howard 2005). The northern expanse of *F. fletcheri* was omitted from this analysis as it is the only species of its genus to occupy that area, a shared distribution with *S. purpurea*, and is almost certainly a result of post-glacial northward colonization, a phenomenon seen in *Wyeomyia smithii* (Merz et al. 2013). Therefore, the northern presence of *F. fletcheri* was considered autapomorphic and not included.

Results

Molecular Datasets

cox1 - The *cox1* final alignment comprised 1539 bases each for 190 specimens representing 10 *Sarcophaga sarraceniae*, 3 *Blaesoxipha* (downloaded from GenBank), 29 *F. papei*, 27 *F. rileyi*, 32 *F. fletcheri*, 10 *F. folkertsi*, 10 *F. oreophilae*, 20 *F. celarata*, 22 *F. abdita*, and 27 *F. jonesi*. Mitochondrial DNA also proved to be effective in barcode ID of larval specimens, for which there are no known distinguishing morphological structures. With the ability to easily type larvae genetically came a wider range of sampling possibilities. Larvae are more abundant and persistent than adults, which have much shorter phenologies and are affected by weather and time of day; indeed, we recovered larvae exclusively at many of our collecting localities.

2bRAD - ipyrad recovered 382,692 total filtered loci in our highest supported alignment matrix that included the outgroups *Blaesoxipha* and *Sarcophaga*. Of these loci, 90,400 (23.6%) contained at least one parsimony informative site. The final alignment was 10,283,609 bases long (547,587 SNPs) with 86.67% missing sites. Alternate alignment matrices based on different filtering parameters contained 50,713 loci with 55% missing data and 686,368 loci with 91.7% missing data. This relatively high percentage of missing data is the cost of higher locus inclusivity; this trade-off has been shown to increase phylogenetic signal with little detriment to overall reconstructions (Booher et al. 2021).

Phylogenetic Analyses

cox1 - BI and ML analyses both recovered the same overall topology for *cox1* (Figure 6), other than minor within-species shuffling of tips, and supported the monophyly of *Fletcherimyia*. Moreover, all eight morphologically defined species were recovered as being monophyletic, each receiving 100% posterior (BI) and bootstrap (ML) support. Internal nodes of the tree were strongly supported with the exception of the sister group to *F. papei* (PP 69.3%, BSS 56%), where varying the total specimen number per species resulted in different sister group assignments. Thus, *F.*

papei was recovered either as sister to all other *Fletcherimyia* or sister to *F. rileyi* + (*F. fletcheri* + *F. folkertsi*); this discordance likely contributed to low support at this node. BI gave only moderate support (PP=90.2%) to the root node of the clade *F. rileyi* + (*F. fletcheri* + *F. folkertsi*), though ML support was much stronger (BSS=94%).

2bRAD - ML analysis of RADseq data again supported the monophyly of *Fletcherimyia* (100% BSS) and the monophyly of each species therein (100% BSS). Overall topology varied substantially from that of trees inferred from *cox1* data but generally had much higher support (>94% ingroup nodes highly supported; Figure 7). Our topology identified two major clades: the *papei* clade (including *F. papei*, *F. fletcheri*, and *F. folkertsi*) and the *rileyi* clade (including *F. rileyi, F. oreophilae, F. abdita, F. celarata,* and *F. jonesi*) (Figure 7). The only node with low support in alignment matrices with fewer loci was the *papei* clade (72% BSS), though this increased with more included loci (93% BSS). This clade also has the shortest branch length of the entire tree, and it is likely that rapid speciation occurred resulting in ILS and loss of signal. The relationships within *F. papei* are consistent between datasets, revealing two phylogeographic lineages. Overall topology did not differ between matrices comprising different numbers of loci, but support for individual nodes did vary (Figure 7).

Divergence Dating - *cox1* divergence estimates under the strict clock model show that *Fletcherimyia* originated during the mid to late Pliocene (95% HPD 2.23-2.97 mya, median = 2.56 mya), placing the genus within the estimated range for the radiation of *Sarracenia* (1-3 mya, Ellison et al. 2012). Most other dated nodes are discordant with the RADseq-derived topology and are therefore likely uninformative. However, the *folkertsi* + *fletcheri* and east/west *papei* clades, both congruent with RADseq data, appear to have diverged approximately the same time, 1.50 mya (95% HPD 1.18-1.87 mya) and 1.45 mya (95% HPD 1.15-1.77 mya).

Cophylogenetic Analysis

Global-fit analysis in *PACo* found significant congruence between *Fletcherimyia* and *Sarracenia* (p=6e-4) as compared to the null model of completely unrelated phylogenies generated across 2,000 permutations. Interestingly, *S. flava* and *S. minor* each had both the highest and lowest squared residuals (SQres), where lower SQres indicates shorter distance from linked parasite taxa on the PCA plot. This could be due to the three fly species, *F. abdita*, *F. jonesi* and *F. rileyi*, affiliated with *S. flava* and *S. minor*. *Fletcherimyia abdita* and, to a lesser extent, *F. jonesi* and *F. rileyi* are pitcher generalists relative to their congeners and may contribute to a higher SQres. Different host topologies, including those with collapsed nodes, did not have a meaningful impact on results.

Jane recovered twelve isomorphic solutions, all of which were influenced largely by host loss (n = 15) and failures to diverge (n = 11) primarily within the *Sarracenia oreophila* clade and the poorly resolved *S. rubra* complex. However, Jane did recover two cospeciation events in all solutions: one coinciding with the root node of *Fletcherimyia* and the ((*minor* + (*psittacina* + *flava*)) + (*purpurea* ssp. *montana* + (*rosea* + *purpurea purpurea*)) pitcher clade (hereby referred to as MPFPR), and one coinciding with *F. folkertsi* + *F. fletcheri* and *S. rosea* + *S. purpurea* ssp. *purpurea* (Figure 9). The majority of *Fletcherimyia*'s species (excluding *F. abdita*) originate from the node placed at the root of the MPFPR pitcher clade, suggesting most colonization of the *S. oreophila* clade occurred via a series of duplications and host switches rather than by codiversification. One possible exception is *F. papei*; flies collected in the fall line sandhills of west Georgia mark the eastern boundary of the western *F. papei* phylogroup, and the locality's host plants are currently classified as *S. rubra* ssp. *rubra*. However, these plants are unofficially referred to as *S. rubra* ssp. *viatorum*. If the west Georgia host plants are treated as a separate lineage, Jane recovers a third cospeciation event between *F. papei* populations and *S. rubra* + sister subspecies, regardless of the placement of *S. rubra* ssp. *viatorum* in the *Sarracenia*

phylogeny. Further research is needed to investigate the validity of *S. rubra* ssp. *viatorum*; therefore, this relationship was not included in the final analysis.

Historical Biogeography

BioGeoBears placed the ancestral range of *Fletcherimyia* in the "Mountain" and "Eastern" zones, with several migrations to and from the "Gulf" zone (Figure 10). Parameter estimation under the DEC+J model (d=0, e=0, j=3) suggests that jump dispersal was mostly responsible for current range distributions within the genus. The first bifurcation of the tree shows a shift in the *rileyi* clade towards a gulf range, with two subsequent jumps back east in *F. rileyi* and *F. jonesi*, one jump to the mountains in *F. oreophilae*, and two retentions of the ancestral state in *F. celarata* and *F. abdita*. Interestingly the ancestral *F. papei* species was recovered in the gulf, despite one daughter branch residing in the east and its ancestral range including the eastern and mountain zones. Also of note is the pattern that, with the exception of *F. oreophilae* in the mountains, all extant taxa are in sister couplets of eastern and western species, suggesting that range shifts have often accompanied cladogenesis in *Fletcherimyia*.

Discussion

Species Monophyly and Host Relationships

Mitochondrial sequence data corroborated all morphological species constructs in *Fletcherimyia*, which are based on adult specimens. The only morphological description given of *Fletcherimyia* larvae is of the species *F. fletcheri*; it possesses a posterior "cup" surrounding its spiracles (Dahlem and Naczi 2006) which allows the larva to suspend itself at the surface of the pitcher fluid in *S. purpurea*. As such, larval species identification is presently impossible; potential morphological descriptions would require rearing larvae to adulthood. Previous attempts to rear larvae have been successful in other studies (Aldrich 1916; Dahlem and Naczi 2006), but conditions necessary for successful pupariation appear to be fairly specific. However, diagnostic (i.e., species specific) *cox1* sequences serve effectively as molecular markers for objective larval identification and, in turn, reliable assessment of host use. Typing larvae, we observed all fly-host associations reported in Dahlem and Naczi (2006) and documented new fly-host relationships for certain species (summarized in Figure 3).

Fletcherimyia celarata, considered to be a sole affiliate of *S. leucophylla* (Dahlem and Naczi 2006), has been alluded to affiliate with *S. alata* (Satler and Carstens 2016; 2019). We routinely recovered larval *F. celarata* in pitchers of *S. alata* and commonly observed adult flies on *S. alata* as well. In fact, *F. celarata* was more common at many *S. alata* sites than *F. abdita*, the only species of *Fletcherimyia* previously thought to associate with *S. alata*. We also documented larval *F. celarata* in pitchers of *S. rubra* ssp. *gulfensis* from the Florida panhandle. *Sarracenia rubra* ssp. *gulfensis* possesses pitchers similar in size and structure to those of *S. leucophylla* and *S. alata* and thus may not present much of an ecological barrier for host use by *F. celarata*. The subspecies of *Sarracenia rubra* do not form a monophyletic group (Figure 10), and *S. rubra* ssp. *gulfensis* may be a species in its own right. For this reason, we do not view *F. celarata* as an associate of *S. rubra sensu lato*. We also observed *F. papei* in an isolated population

of *S. alata* in central Texas, previously noted by Dr. Rob Naczi (unpublished data). Despite collecting larvae at the locality, they were typed as *F. celarata* and *Sarcophaga sarraceniae*. We did observe *F. papei in copula* on *S. alata* at the site, which lies beyond the western range termini of other *Sarracenia* species. However, without confirming larval usage, these observations of *F. papei* cannot be considered conclusive regarding host use and were therefore not included in this study. (Inclusion of this relationship was added as an exploratory measure but had no consequential outcome on results.)

Most unexpectedly, we found *F. oreophilae* in association with *S. jonesii*, an endangered montane species endemic to the western Carolinas. This species had no previously reported fly associations, though the presence of *F. oreophilae* makes ecological sense as its only known host, *S. oreophila*, is also endemic to mountain bogs further west. Unfortunately, both montane host species are of grave conservation concern, which implies a similar circumstance for *F. oreophilae*. *Fletcherimyia folkertsi* and *F. fletcheri* are then the only species in the genus to strictly associate with a single host species, *S. rosea* and *S. purpurea*, respectively.

Not only did larval genetic typing enable a robust assessment of fly host usage, it also greatly expanded our overall collection efforts. Of the 774 collected specimens, the majority (411) were larvae. Many localities, particularly in northern states, yielded no adults at all. Larval identification via *cox1* barcoding provided a far more comprehensive sampling effort, with wider geographic and ecological scope than would have been possible using adult flies alone.

Fletcherimyia *Phylogeny*

Trees inferred from both molecular datasets were highly supported, particularly the 2bRADseq tree, whose topology gives more meaningful evolutionary context to Dahlem and Naczi's (2006) phenotypic patterns observed among species. They note repeatedly certain morphological similarities among *F. papei*, *F. fletcheri*, and *F. folkertsi*, as well as female genitalic similarities among *F. abdita*, *F. celarata*, *F. jonesi* and *F. rileyi* (*F. oreophilae* was described as being "very

distinctive" from all other species in both male and female terminalia). These two groups form reciprocally monophyletic clades, termed the "*papei* clade" and "*rileyi* clade," in our 2bRAD phylogeny, giving phylogenetic significance to certain morphological similarities. Dahlem and Naczi (2006) also partition species into these same groups in the first couplet of their dichotomous key. Flies of the *papei* clade possess tan or reddish abdomens (Figure 8: F, G, H) and three postsutural dorsocentral setae, whereas flies of the *rileyi* clade possess dark or gray abdomens with a "normal tessellation pattern" (Figure 8: A, B, C, D, E) and four postsutural dorsocentral setae.

The *rileyi clade* - This clade contains more generalist species in terms of host selection, affiliating typically with pitcher species that are similar in overall shape, size, and habit. *Fletcherimyia rileyi* and *F. jonesi* both use *S. minor* and *S. flava* seemingly indiscriminately and often co-occur with one another. Had they been recovered as sister species, it would be difficult to describe an evolutionary pressure that would have resulted in their speciation. This was not the case on our 2bRAD phylogeny; *F. rileyi* and *F. jonesi* are separated by two speciation events. The same applies for *F. abdita* and *F. celarata*, which were recovered as sister species with high support on the *cox1* phylogeny. These two species overlap considerably in range and host preference, though not to the full extent of *F. rileyi* and *F. jonesi*. In addition to *S. leucophylla* and *S. alata*, the common hosts of *F. celarata*, *F. abdita* is also found in *S. rubra* ssp. *wherryi*, *S. flava,* and *S. minor*, though only in the extreme western ranges of the latter two pitcher species. Our 2bRAD data did not recover *F. celarata* and *F. abdita* not as sister species, but instead as sister each to another species with which they share almost no geographic overlap. This pattern suggests that pitcher use has evolved independently among these species, or at least that allopatric speciation has been a driving force within the *rileyi* clade.

The montane range of *F. oreophilae*, and its affiliation with *S. oreophila*, made its shift in topology between datasets the most unexpected. Hypotheses for the diversification of *Sarracenia* posit a montane origin in the southern Appalachians, with multiple migrations outwards to the

east and west (Stephens et al. 2015). This view is supported by primarily water-based (and thus downhill-moving) seed dispersal and by the phylogenetic positions of certain montane taxa; both *S. oreophila* and *S. purpurea* ssp. *montana* are the basal lineages of their respective clades (Figure 11). *Fletcherimyia oreophilae* was recovered as the basal lineage to the clade including *F. abdita*, *F. celarata*, and *F. jonesi* on our *cox1* tree, which would be consistent with a similar origin for *Fletcherimyia* and possibly an earlier pitcher association. In contrast, the 2bRAD tree recovers *F. oreophilae* as a more derived species, sister to a western species. This implies a very different history for *F. oreophilae*, which may have reached its current range by northern dispersal through Alabama to the southern range terminus of *S. oreophila*. This implies that *Fletcherimyia* did not share *Sarracenia*'s montane origin; the two genera became associated only after *Sarracenia* became established in the coastal plain.

The *papei clade* - Members of the *papei* clade are much more host-specific, inhabiting pitchers that tend to possess distinct morphologies among *Sarracenia*. These pitcher species often host other species-specific arthropod associates. *Exyra ridingsii*, an inquiline of *S. flava*, does not affiliate at all with *S. purpurea* even though the plants often occur in sympatry. Instead, a specialized congeneric species, *Exyra fax*, associates with *S. purpurea* and no other pitcher species. The same is true of *F. fletcheri*, which only associates with pitchers of *S. purpurea*. Similarly, *F. folkertsi* associates exclusively with pitchers of *S. rosea*, which are nearly identical to those of *S. purpurea* and are also consistently filled with rainwater. Both of our datasets recovered *F. fletcheri* and *F. folkersti* as sister species, mirroring the sister status of their host plants. These are the only two species in the family Sarcophagidae that are known to have aquatic larvae (Dahlem and Naczi 2006); their larvae possess a posterior spiracle cup that allows them to trap air bubbles and hang from the water's surface. (As an aside, the other species of *Fletcherimyia* have also historically been considered to have aquatic larvae, though the inside of other pitcher species tend to only have a milliliter of fluid at most and are more wet than water-filled.) This

highly specialized behavior is unlikely to have evolved twice; our topologies recover these aquatic adaptations as synapomorphic between the two species.

Fletcherimyia papei, the smallest species in *Fletcherimyia*, affiliates with the smallest pitcher species, *S. rubra*. *Sarracenia rubra* ssp. have the patchiest distributions of the genus; whether this was historically the case or is a result of anthropogenic sources is unknown, though the latter has certainly exacerbated the fragmentation of their habitat and range. This patchiness is most profound in Georgia, where only a few small populations dot their way diagonally across the state (Stephens et al. 2015). Even among these somewhat isolated populations, we observed highly variable health and density of the plants. This may be a driving factor behind the genetic divergence between eastern and western lineages of *F. papei*. This divide coincides with the fall line sandhills, an area well known as a phylogeographic break and contact point between hybridizing species (Swenson and Howard 2005). Whether ancestral *F. papei* originated along the sandhills and spread outward, or whether *F. papei* spread from west to east or vice-versa, is unknown.

nDNA and mtDNA Discordance

Though in agreement regarding species monophyly, the deeper topologies for trees generated from mtDNA and 2bRAD were largely incongruent. Trees from both datasets recovered *F. fletcheri* and *F. folkertsi* as sister species as well as the eastern/western phylogroups within *F. papei*. However, other species relationships differed significanty between datasets, particularly for *F. rileyi*. In the *cox1* tree, *F. rileyi* is sister to *fletcheri* + *folkertsi*, with *papei* in turn sister to that clade. In the 2bRAD tree, *F. rileyi* is placed as sister to *abdita*, *celarata*, *jonesi,* and *oreophilae* and shares a MRCA with *papei*, *fletcheri*, and *folkertsi* only in the root node of the genus. As mentioned above, *F. oreophilae* shifted from the most basal species in its clade to one of the most derived, altering potential interpretations of *Fletcherimyia* evolution.

Discordance between nuclear and organellar genomes is a well known phenomenon and is more prevalent in recently diverged taxa (Stephens et al. 2015; Shaw 2002). This discordance is not unique to organelle genomes, rather it is a commonly encountered issue with phylogenies inferred from single genes (Sanders et al. 2013; Kubatko, Gibbs, and Bloomquist 2011). Genes typically have older coalescence times than their respective species, which can lead to instances of incomplete lineage sorting and discordant gene tree/species tree topologies. Rapid, successive speciation events increase the rate of ILS across the genome, and in recently diverged taxa there may not have been enough time for distinct alleles to go to fixation, further obscuring phylogenetic signal.

Divergence dating suggests *Fletcherimyia* is only ~2.5 million years old (comparable to the estimated age of *Sarracenia* at ~3 million years), which may be responsible for some discordance between datasets. *Fletcherimyia* also experienced several rapid speciation events early in its evolutionary history (Figure 7); this phenomenon also characterizes *Sarracenia* and is likely responsible for its poor resolution and nodal support (Stephens et al. 2015) (Figure 11). The mitochondrial genome has a lower effective population size than typical nuclear genes, due to its maternal pattern of inheritance and haploidy, and therefore has lower rates of ILS. However, ILS is not the only potential source of gene discordance. Phylogenetic signal from mitochondria can be further affected by hybridization and introgression (McGuire et al. 2007; Alves et al. 2008). Although we do not see explicit signs of introgression in *Fletcherimyia*, our mtDNA tree does recover *F. celarata* and *F. abdita*, two species with nearly total range overlap, as sister species. This relationship is not supported by our RADseq tree, in which *F. celarata* and *F. abdita* are sister to *F. jonesi* and *F. oreophilae*, respectively. Introgression and hybridization are common within host plants and have had significant impact on phylogenetic reconstructions of *Sarracenia*, resulting in similar discordance between trees inferred from nuclear and organellar DNA (Stephens et al. 2015). Montane species, distantly related on a MP-EST hybrid enrichment tree,

were recovered as sister taxa on a topology inferred from plastid DNA which was in general much less resolved and supported than trees inferred from nuclear genes.

Whatever the underlying evolutionary cause for gene tree discordance, an effective way to discern the "true" species tree is by incorporating more loci from different portions of the genome (Stephens et al. 2015; Sanders et al. 2013; Kubatko, Gibbs, and Bloomquist 2011; Shaw 2002). RADseq effectively does this by nature of its stochastic capture of sequence associated with restriction sites across the genome, thereby incorporating many gene histories and not relying on any single possible topology.

2bRAD and Impact of Missing Data

2bRAD, a recent modification of the RADseq protocol, has previously been used primarily in studies of aquatic organisms at the population level (Manzello et al. 2019; Kelly and Thacker 2020; Pecoraro et al. 2016; Aslam et al. 2018; Liu et al. 2017) and examples of its efficacy in reconstructing deeper evolutionary relationships, particularly in insects, have been mostly theoretical (Seetharam and Stuart 2013). In this study we provide evidence that 2bRAD is effective not only at infraspecific levels, but also at inferring a phylogeny of a non-model, non-aquatic genus of recently diverged insects. Our 2bRAD phylogeny had high support throughout and recovered a topology that logically progressed from morphological observations. Due to the recent development of 2bRADseq and its historical use with primarily marine organisms, it is not well known how many RAD tags 2bRADseq can recover in other taxa or what coverage those tags will provide across the genome, particularly in non-model organisms (Yu Guo et al. 2014). Here we were able to recover a high number of loci with sufficient genomic coverage to infer a highly supported phylogeny.

As discussed above, mitochondrial DNA can be less informative for older divergence patterns, whereas RADseq has been shown to be effective over a broad range of divergence times. Tripp et. al (2017) provides an excellent summary of relatively old and young taxa whose

phylogenies were effectively inferred using RADseq data; dates range from less than 15,000 years ago in a radiation of African cichlids in Lake Victoria (Wagner et al. 2013) to 80 mya in the deepest divergences of *Paragorgia*, a genus of deep-water corals (Herrera and Shank 2016). Other genetic data seem to have a "sweet spot" for informative phylogenetic signal. For example, *cox1* is most informative in recently diverged taxa, whereas nuclear genes with slower mutation rates maintain signal in much older radiations but have not had time to accrue enough substitutions to be informative for recent taxa. RADseq, perhaps by nature of its random sampling from differently evolving regions across the genome, then seems to be somewhat removed from this issue. It is therefore not surprising that our 2bRAD data agreed with mtDNA in shallow clades of their respective trees (i.e. species) but disagreed with higher support as to the topology of the deeper nodes (88.7% average node support for mtDNA, 95.6% average node support for 2bRAD with 86% missing data, excluding species nodes).

Many recent studies have shown that the number of loci recovered, not the age of the taxa in focus, is most influential on RADseq's utility in a given system (Eaton et al. 2017; Crotti et al. 2019; Tripp et al. 2017; Wagner et al. 2013). Including even the loci shared across as few as two or three samples is seemingly beneficial to overall phylogenetic analysis. However, with an increased number of loci comes a higher proportion of missing data, a hallmark of RADseq studies. In phylogenetics, missing data is usually associated with a decrease in signal and overall support values, but it has been shown that it is not the missing data itself that causes such problems. Rather, when too few characters are used in a phylogeny, taxa with incomplete data have a higher proportion of incomplete characters and therefore contribute to lower overall confidence (Wiens 2003). RADseq data provide many, many characters across loci, therefore the instance of a taxon having few to no complete characters is rare. In fact, missing data tends to correlate positively with overall tree support in RADseq studies when the missing data is a result of less strict filtering of loci. Though we did not see the same level of support improvement when adding more loci at the cost of missing data as other studies (e.g., Crotti et al. 2019), we did

observe increased support for the *papei* clade when we increased the percentage of missing data from 55% (BSS 68%) to 86% (BSS 72%) to 91.7% (BSS 93%). This may be due to the longer reads of traditional RADseq (>100bp on average) compared to that of 2bRAD (25-35bp; 27bp on average in this study). Longer reads are more likely to contain more informative sites, therefore the omission of one locus may be more detrimental. Despite seeing an increase in support for the *papei* clade at the highest percentage of missing data and number of loci, average node support decreased (95.1 vs 95.6 vs 93.4% average node support at 50,713 vs 382,692 vs 686,368 loci, respectively, not including species nodes). This contrasts with other studies and may point to more phylogenetic noise than signal added when including loci shared only by a minimum of two samples. The *papei* clade has the shortest branch length of any node on the 2bRAD tree and therefore likely has the most phylogenetic discordance among different regions of the genome (i.e. ILS). By including more loci, a stronger overall signal may have become apparent; for other regions of the tree which did not have this same issue, the additional loci may have instead obscured the predominant signal.

Fletcherimyia *Phylogenies in the Context of* Sarracenia

The two major clades of *Fletcherimyia* differ consistently in the specificity of their host preferences: the *rileyi* clade is largely generalist whereas the *papei* clade is more host specific. This distinction also correlates with the nature of reconstructed events behind each fly/pitcher affiliation. However, these clades do not show explicit geographic structuring, each containing roughly equal proportions of species in the eastern and western portions of *Sarracenia*'s range. This is in sharp contrast to the most supported phylogeny of *Sarracenia* (Figure 11), which bifurcates into two clades showing obvious geographic structure; each clade is primarily east or west of the Apalachicola region, with only one or two species crossing that break.

Despite the difference in phylogeographic patterns between the fly and pitcher genera, global-fit analysis shows a significant impact of *Sarracenia* on the evolutionary history of

Fletcherimyia. Event-based analysis attempts to describe the details of this impact (Figure 12); we recovered an ancestral *Fletcherimyia* associating with the eastern lineage of *Sarracenia* and an early cospeciation coinciding with the first cladogenesis of *Fletcherimyia* and the MPFPR pitcher clade. Divergence dating analysis puts the root of *Fletcherimyia* slightly after the proposed age of *Sarracenia* (~2.5 mya vs ~3 mya), supporting the event-based result that ancestral fly populations likely began their relationship with pitchers after some diversification within *Sarracenia*. Our reconstruction of fly biogeography correlates with these findings, suggesting an ancestral range for *Fletcherimyia* that includes the Blue Ridge Mountains and the Atlantic coastal plain, collectively home to the overwhelming majority of MPFPR pitcher clade populations. The only exceptions to this are *S. psittacina*, which does not have an affiliated fly species (due likely to unique morphology that is not conducive to fly larviposition), distributed from the Georgia coast west to Mississippi, and *S. rosea*, which exclusively occurs along the Gulf Coast.

Sarracenia rosea was once considered to be a subspecies of *S. purpurea* and is nearly identical save for a few subtle but consistent morphological distinctions (NACZI et al. 1999). More obvious is its distinct range, completely allopatric from *S. purpurea*. The geographic area that separates *S. rosea* from *S. purpurea*, generally around the Apalachicola basin in the Florida panhandle, is a well known phylogeographic break (Avise 1992) and has been periodically been impassable due to glaciation-related climatic changes over the Pleistocene and Pliocene epochs. This is almost certainly the cause for the speciation between *S. rosea* and *S. purpurea* (Godt and Hamrick 1998). It is at this point of the *Sarracenia* phylogeny that we recover the second instance of cospeciation with *Fletcherimyia*, specifically between *F. fletcheri* and *F. folkertsi*. Both fly species exhibit highly specialized host preferences; *F. fletcheri* only associates with pitchers of the *S. purpurea* subspecies, and *F. folkertsi* associates exclusively with pitchers of *S. rosea*. This cospeciation event was then almost certainly the result of vicariance. *Fletcherimyia papei*, sister to *F. fletcheri* + *F. folkertsi*, has similarly specific host preferences and only associates with two of the three monophyletic *S. rubra* subspecies (ssp. *rubra* and *wherryi*; *jonesii* is strictly montane

with much larger pitchers and is often considered to be a separate species). These species, like *S. purpurea* and S. *rosea*, have morphologically distinct interior spaces; these are by far the smallest *Sarracenia* species, and *F. papei* is the smallest species of *Fletcherimyia*. However, our results suggest this relationship originated not from cospeciation, but from a duplication (lineage split) and host switch. This seems to be the major force behind the diversification of host preferences in the genus and is common throughout the *rileyi* clade.

As noted earlier, the *rileyi* clade contains largely generalist species in terms of host selection. Our cophylogenetic reconstruction suggests that their ancestral population affiliated with the eastern *S. minor* pitcher clade and colonized the western *S. oreophila* clade through a series of cladogenetic and anagenetic host switches. Today *F. rileyi*, the basal taxon of the clade, affiliates solely with these same pitcher species. Jane suggests that this is a retained ancestral state; however, our biogeographic reconstruction places the ancestor of this clade in the west - this discordance requires more information to resolve. In any case, the ancestral *rileyi* clade was likely unaffiliated with species of the *S. oreophila* clade, despite the western distribution and preferences of three out of five extant taxa. The generalist nature of the *rileyi* clade likely enabled host switches that colonized western pitcher species; today these species still affiliate with >2 host species on average, with *F. abdita* routinely found in five. Generalist behavior may be a trait intrinsic to these flies but may also be due to overall similarity between their hosts. Hosts of *F. papei*, *F. fletcheri*, and *F. folkertsi* possess much deeper morphological differences; in contrast there may be no real ecological barrier, and therefore need to adapt, between most hosts of the *rileyi* clade. *Sarcophaga sarraceniae* uses all of these hosts indiscriminately but has never been recorded in *S. rubra* (spp. *rubra* and *wherryi*) or *S. rosea*, and though it is an affiliate of *S. purpurea*, we have found it to be much less common in those plants than *F. fletcheri*. This suggests that movement between the hosts of the *rileyi* clade may be ecologically simple. Regardless of the origin of the *rileyi* clade's generalist behavior, the outcome is the same.

The correlation between generalist preferences and host expansion is a well-studied phenomenon in phytophagous insects and is a key element of the Oscillation Hypothesis described by Janz and Nylin (2008). In this framework, a subset of a specialized phytophagous insect group either develops more generalized host preferences or is plastic enough to tolerate a novel host. This leads to a host expansion, which can coincide with a geographic expansion of insect ranges (as was likely the case with *Fletcherimyia* and *Sarracenia*). Finally, new host preferences or allopatry lead to local adaptation and/or specialization to a new host. This hypothesis posits an "oscillation" between specialization and generalization as a pathway to diversification. Though *Fletcherimyia* is kleptoparasitic (or kleptobiotic), it does share behavioral similarities with phytophagous insects. In both cases, host use is dictated by females, young are bound to the plants their mothers choose, host specificity varies, and host dependency is absolute. Studies of phytophagous butterflies (Nylin, Slove, and Janz 2014) and parasitoid tachinid flies (Stireman 2005) recovered a correlation between generalized host preferences and more speciose clades - a trend that appears in *Fletcherimyia*. Per the pathway outlined in the oscillation hypothesis, *F. oreophilae* may be an example of derived specialization in a generalist clade. However, it is not known if *F. oreophilae* has actually specialized to its hosts, as its host preferences may be a product of geographic isolation or local adaptation to a montane environment. Geographic isolation has likely been a major force in the more recent diversification of *Fletcherimyia*, more so than codiversification with *Sarracenia*. Extant taxa are almost all phylogenetically sister in allopatric couplets; sister species (or populations in the case of *F. papei*) are commonly divided between eastern and western ranges, or between coastal plains and mountains. This pattern may have been caused by the cyclical fragmentation and reunification of habitat in the southeastern United States which has produced similar phylogenetic patterns in many other taxa (Avise 1992).

Conservation

Destruction of habitat, agricultural runoff, poaching, and climate change have left only 3% of *Sarracenia*'s original range intact (G. W. Folkerts 1982). This devastation is further compounded by heavy fragmentation of an already highly specialized community, weakening dispersal routes for bog organisms and isolating populations. Six species of pitcher plant are already of conservation concern, as are their inquilines including *Exyra fax*, which was listed as threatened in Connecticut in 2015 (Connecticut Department of Energy and Environmental Protection). *Fletcherimyia* has yet to be included on these lists but is certainly at least as rare as its hosts. For example, *F. oreophilae* is restricted to two pitcher species, *S. oreophila* and *S. jonesii*, both of which are critically endangered. To compound this issue, we failed to recover *F. oreophilae* at some *S. oreophila* and *S. jonesii* sites, of which few remain. Even the type locality from which Dahlem and Naczi described *F. oreophilae* in 2006 failed to yield any specimens across our field seasons in 2019 and 2020. We have shown that the genus exhibits a history of host switching and that host preference plasticity has persisted throughout the radiation of the genus. This points to some resilience in *Fletcherimyia* to extirpations or extinctions of single pitcher species, but proper management and protection of *Sarracenia* and bog habitats will remain vital to preserving the diversity of these flies.

References

Adlassnig, Wolfram, Marianne Peroutka, and Thomas Lendl. 2011. “Traps of Carnivorous Pitcher Plants as a Habitat: Composition of the Fluid, Biodiversity and Mutualistic Activities.” *Annals of Botany* 107 (2): 181–94. https://doi.org/10.1093/aob/mcq238.

Aldrich, J.M. 1916. “Sarcophaga and Allies in North America.” *Entomological Society of America. Thomas Say Foundation. La Fayette, IN.* 1: 302.

Alves, Paulo C, José Melo-Ferreira, Hélder Freitas, and Pierre Boursot. 2008. “The Ubiquitous Mountain Hare Mitochondria: Multiple Introgressive Hybridization in Hares, Genus Lepus.” *Philosophical Transactions of the Royal Society B: Biological Sciences* 363 (1505): 2831–39. https://doi.org/10.1098/rstb.2008.0053.

Aslam, Muhammad L., Roberta Carraro, Anastasia Bestin, Sophie Cariou, Anna K. Sonesson, Jean-Sébastien Bruant, Pierrick Haffray, Luca Bargelloni, and Theo H. E. Meuwissen. 2018. “Genetics of Resistance to Photobacteriosis in Gilthead Sea Bream (Sparus Aurata) Using 2b-RAD Sequencing.” *BMC Genetics* 19 (1): 43. https://doi.org/10.1186/s12863-018-0631-x.

Avise, John C. 1992. “Molecular Population Structure and the Biogeographic History of a Regional Fauna: A Case History with Lessons for Conservation Biology.” *Oikos* 63 (1): 62–76. https://doi.org/10.2307/3545516.

Balbuena, Juan Antonio, Raúl Míguez-Lozano, and Isabel Blasco-Costa. 2013. “PACo: A Novel Procrustes Application to Cophylogenetic Analysis.” *PLOS ONE* 8 (4): e61048. https://doi.org/10.1371/journal.pone.0061048.

Booher, Douglas B., Joshua C. Gibson, Cong Liu, John T. Longino, Brian L. Fisher, Milan Janda, Nitish Narula, et al. 2021. “Functional Innovation Promotes Diversification of Form in the Evolution of an Ultrafast Trap-Jaw Mechanism in Ants.” *PLOS Biology* 19 (3): e3001031. https://doi.org/10.1371/journal.pbio.3001031.

Bouckaert, Remco, Timothy G. Vaughan, Joëlle Barido-Sottani, Sebastián Duchêne, Mathieu Fourment, Alexandra Gavryushkina, Joseph Heled, et al. 2019. “BEAST 2.5: An Advanced Software Platform for Bayesian Evolutionary Analysis.” *PLOS Computational Biology* 15 (4): e1006650. https://doi.org/10.1371/journal.pcbi.1006650.

Bradshaw, William E., and Robert A. Creelman. 1984. “Mutualism between the Carnivorous Purple Pitcher Plant and Its Inhabitants.” *The American Midland Naturalist* 112 (2): 294–304. https://doi.org/10.2307/2425436.

Buenaventura, Eliana. 2021. “Museomics and Protein-Encoding Ultraconserved Elements Illuminate the Evolution of Life History and Phallic Morphology of Flesh Flies (Diptera: Sarcophagidae).” Preprint. In Review. https://doi.org/10.21203/rs.3.rs-274320/v1.

Buenaventura, Eliana, and Thomas Pape. 2018. “Phylogeny, Evolution and Male Terminalia Functionality of Sarcophaginae (Diptera: Sarcophagidae).” *Zoological Journal of the Linnean Society* 183 (4): 808–906. https://doi.org/10.1093/zoolinnean/zlx070.

Buenaventura, Eliana, Krzysztof Szpila, Brian K. Cassel, Brian M. Wiegmann, and Thomas Pape. 2019. "Anchored Hybrid Enrichment Challenges the Traditional Classification of Flesh Flies (Diptera: Sarcophagidae)." *Systematic Entomology*, October, syen.12395. https://doi.org/10.1111/syen.12395.

Conow, Chris, Daniel Fielder, Yaniv Ovadia, and Ran Libeskind-Hadas. 2010. "Jane: A New Tool for the Cophylogeny Reconstruction Problem." *Algorithms for Molecular Biology* 5 (1): 16. https://doi.org/10.1186/1748-7188-5-16.

Crotti, Marco, Christopher D. Barratt, Simon P. Loader, David J. Gower, and Jeffrey W. Streicher. 2019. "Causes and Analytical Impacts of Missing Data in RADseq Phylogenetics: Insights from an African Frog (Afrixalus)." *Zoologica Scripta* 48 (2): 157–67. https://doi.org/10.1111/zsc.12335.

Cruaud, Astrid, and Jean-Yves Rasplus. 2016. "Testing Cospeciation through Large-Scale Cophylogenetic Studies." *Current Opinion in Insect Science* 18 (December): 53–59. https://doi.org/10.1016/j.cois.2016.10.004.

Dahlem, G. A., and R. F. C. Naczi. 2006. "Flesh Flies (Diptera: Sarcophagidae) Associated with North American Pitcher Plants (Sarraceniaceae), with Descriptions of Three New Species." *Annals of the Entomological Society of America* 99 (2): 218–40. https://doi.org/10.1603/0013-8746(2006)099[0218:FFDSAW]2.0.CO;2.

Davey, John W., and Mark L. Blaxter. 2010. "RADSeq: Next-Generation Population Genetics." *Briefings in Functional Genomics* 9 (5–6): 416–23. https://doi.org/10.1093/bfgp/elq031.

Desdevises, Y. 2007. "Cophylogeny: Insights from Fish-Parasite Systems." *Parassitologia* 49 (October): 5.

Downes, W. L. Jr. 1965. "Family Sarcophagidae." In *A. Stone, C. W. Sabrosky, W. W. Wirth, R. H. Foote, and J. R. Coulson [Eds.], A Catalog of the Diptera of America North of Mexico. Agriculture Handbook*, 933–61. 276. Washington, DC: Agricultural Research Service, USDA.

Drummond, Alexei J., and Andrew Rambaut. 2007. "BEAST: Bayesian Evolutionary Analysis by Sampling Trees." *BMC Evolutionary Biology* 7 (1): 214. https://doi.org/10.1186/1471-2148-7-214.

Eaton, Deren A. R., and Isaac Overcast. n.d. "Ipyrad: Interactive Assembly and Analysis of RADseq Datasets." *Bioinformatics*. Accessed January 26, 2020. https://doi.org/10.1093/bioinformatics/btz966.

Eaton, Deren A. R., Elizabeth L. Spriggs, Brian Park, and Michael J. Donoghue. 2017. "Misconceptions on Missing Data in RAD-Seq Phylogenetics with a Deep-Scale Example from Flowering Plants." *Systematic Biology* 66 (3): 399–412. https://doi.org/10.1093/sysbio/syw092.

Ellison, Aaron M., Elena D. Butler, Emily Jean Hicks, Robert F. C. Naczi, Patrick J. Calie, Charles D. Bell, and Charles C. Davis. 2012. "Phylogeny and Biogeography of the Carnivorous Plant Family Sarraceniaceae." *PLOS ONE* 7 (6): e39291.

https://doi.org/10.1371/journal.pone.0039291.

Folkerts, Debbie R. 1999. “Pitcher Plant Wetlands of the Southeastern United States.” In *Invertebrates in Freshwater Wetlands of North America: Ecology and Management*, by Darold P. Batzer, Russell B. Rader, and Scott A. Wissinger, 247–75. New York, NY: John Wiley & Sons.

Folkerts, George W. 1982. “The Gulf Coast Pitcher Plant Bogs [Mississippi, Alabama].” *American Scientist*. v. 70(3) p. 260-267

Furches, M. Steven, Randall L. Small, and Anna Furches. 2013. “Hybridization Leads to Interspecific Gene Flow in Sarracenia (Sarraceniaceae).” *American Journal of Botany* 100 (10): 2085–91. https://doi.org/10.3732/ajb.1300038.

Galili (https://www.r-statistics.com), Tal, Yoav Benjamini, Gavin Simpson, Gregory Jefferis (imported code from his dendroextras package), Marco Gallotta (a k a: marcog), Johan Renaudie (https://github.com/plannapus), The R. Core Team (Thanks for the Infastructure, et al. 2019. *Dendextend: Extending “dendrogram” Functionality in R* (version 1.13.2). https://CRAN.R-project.org/package=dendextend.

Godt, Mary Jo W., and J. L. Hamrick. 1998. “Genetic Divergence among Infraspecific Taxa of Sarracenia Purpurea.” *Systematic Botany* 23 (4): 427–38. https://doi.org/10.2307/2419375.

Guo, Yadong, Lagabaiyila Zha, Weitao Yan, Pei Li, Jifeng Cai, and LiXiang Wu. 2014. “Identification of Forensically Important Sarcophagid Flies (Diptera: Sarcophagidae) in China Based on COI and Period Gene.” *International Journal of Legal Medicine* 128 (1): 221–28. https://doi.org/10.1007/s00414-013-0923-7.

Guo, Yu, Hui Yuan, Dongming Fang, Lianbo Song, Yan Liu, Yong Liu, Lu Wu, et al. 2014. “An Improved 2b-RAD Approach (I2b-RAD) Offering Genotyping Tested by a Rice (Oryza Sativa L.) F2 Population.” *BMC Genomics* 15 (1): 956. https://doi.org/10.1186/1471-2164-15-956.

Herrera, Santiago, and Timothy M. Shank. 2016. “RAD Sequencing Enables Unprecedented Phylogenetic Resolution and Objective Species Delimitation in Recalcitrant Divergent Taxa.” *Molecular Phylogenetics and Evolution* 100 (July): 70–79. https://doi.org/10.1016/j.ympev.2016.03.010.

Hutchinson, Matthew C., E. Fernando Cagua, Juan A. Balbuena, Daniel B. Stouffer, and Timothée Poisot. 2017. “Paco: Implementing Procrustean Approach to Cophylogeny in R.” *Methods in Ecology and Evolution* 8 (8): 932–40. https://doi.org/10.1111/2041-210X.12736.

Iyengar, Erika V. 2008. “Kleptoparasitic Interactions throughout the Animal Kingdom and a Re-Evaluation, Based on Participant Mobility, of the Conditions Promoting the Evolution of Kleptoparasitism.” *Biological Journal of the Linnean Society* 93 (4): 745–62. https://doi.org/10.1111/j.1095-8312.2008.00954.x.

Janz, Niklas, and Sören Nylin. 2008. *The Oscillation Hypothesis of Host-Plant Range and Speciation. Specialization, Speciation, and Radiation.* University of California Press.

https://california.universitypressscholarship.com/view/10.1525/california/9780520251328.001.0001/upso-9780520251328-chapter-15.

Jones, Frank Morton. 1904. “Pitcher-Plant Insects.” In *Entomological News, and Proceedings of the Entomological Section of the Academy of Natural Sciences of Philadelphia.*, 15:14–17.

Keller-Schmidt, Stephanie, Nicolas Wieseke, Konstantin Klemm, and Martin Middendorf. 2011. “Evaluation of Host Parasite Reconciliation Methods Using a New Approach for Cophylogeny Generation.” *University of Leipzig, Tech. Rep.*, 26.

Kelly, Joseph B., and Robert W. Thacker. 2020. “Microbiomes and Host Genetics Provide Evidence for Ecological Diversification among Caribbean Members of the Sponge Genus Ircinia Nardo, 1833.” *BioRxiv*, September, 2020.09.04.282673. https://doi.org/10.1101/2020.09.04.282673.

Krawchuk, Margaret A., and Philip D. Taylor. 1999. “Roosting behaviour by *Fletcheromyia fletcheri* (Diptera: Sarcophagidae) in *Sarracenia purpurea* (Sarraceniacea).” *The Canadian Entomologist* 131 (6): 829–30. https://doi.org/10.4039/Ent131829-6.

Kubatko, Laura S., H. Lisle Gibbs, and Erik W. Bloomquist. 2011. “Inferring Species-Level Phylogenies and Taxonomic Distinctiveness Using Multilocus Data in Sistrurus Rattlesnakes.” *Systematic Biology* 60 (4): 393–409. https://doi.org/10.1093/sysbio/syr011.

Lamb, Trip, Teresa C. Justice, Michael S. Brewer, Paul E. Moler, Heidi Hopkins, and Jason E. Bond. 2018. “A Biogeographical Profile of the Sand Cockroach Arenivaga Floridensis and Its Bearing on Origin Hypotheses for Florida Scrub Biota.” *Ecology and Evolution* 8 (11): 5254–66. https://doi.org/10.1002/ece3.3885.

Lamb, Trip, and Elizabeth L. Kalies. 2020. “An Overview of Lepidopteran Herbivory on North American Pitcher Plants (Sarracenia), with a Novel Observation of Feeding on Sarracenia Flava.” *The Journal of the Lepidopterists’ Society* 74 (3): 193–97. https://doi.org/10.18473/lepi.74i3.a7.

Lanfear, Robert, Paul B. Frandsen, April M. Wright, Tereza Senfeld, and Brett Calcott. 2017. “PartitionFinder 2: New Methods for Selecting Partitioned Models of Evolution for Molecular and Morphological Phylogenetic Analyses.” *Molecular Biology and Evolution* 34 (3): 772–73. https://doi.org/10.1093/molbev/msw260.

Liu, Haiyang, Beide Fu, Meixia Pang, Xiu Feng, Xiaomu Yu, and Jingou Tong. 2017. “A High-Density Genetic Linkage Map and QTL Fine Mapping for Body Weight in Crucian Carp (Carassius Auratus) Using 2b-RAD Sequencing.” *G3 Genes|Genomes|Genetics* 7 (8): 2473–87. https://doi.org/10.1534/g3.117.041376.

Lopes, H.S. 1971. “Notes on Tephromyiini with Description of a New Species of Mantidophaga Townsend (Diptera, Sarcophagidae).” *Rev. Bras. Biol.*, no. 31: 3–13.

Maddison, Wayne P, and L Lacey Knowles. 2006. “Inferring Phylogeny Despite Incomplete Lineage Sorting.” *Systematic Biology* 55 (1): 21–30. https://doi.org/10.1080/10635150500354928.

Manzello, Derek P., Mikhail V. Matz, Ian C. Enochs, Lauren Valentino, Renee D. Carlton, Graham Kolodziej, Xaymara Serrano, Erica K. Towle, and Mike Jankulak. 2019. "Role of Host Genetics and Heat-Tolerant Algal Symbionts in Sustaining Populations of the Endangered Coral Orbicella Faveolata in the Florida Keys with Ocean Warming." *Global Change Biology* 25 (3): 1016–31. https://doi.org/10.1111/gcb.14545.

Marussich, Wendy A., and Carlos A. Machado. 2007. "Host-Specificity and Coevolution among Pollinating and Nonpollinating New World Fig Wasps." *Molecular Ecology* 16 (9): 1925–46. https://doi.org/10.1111/j.1365-294X.2007.03278.x.

Matzke, Nicholas. 2012. "Founder-Event Speciation in BioGeoBEARS Package Dramatically Improves Likelihoods and Alters Parameter Inference in Dispersal–Extinction–Cladogenesis (DEC) Analyses." *Frontiers of Biogeography* 4 (January).

McGuire, Jimmy A., Charles W. Linkem, Michelle S. Koo, Delbert W. Hutchison, A. Kristopher Lappin, David I. Orange, Julio Lemos-Espinal, Brett R. Riddle, and Jef R. Jaeger. 2007. "Mitochondrial Introgression and Incomplete Lineage Sorting Through Space and Time: Phylogenetics of Crotaphytid Lizards." *Evolution* 61 (12): 2879–97. https://doi.org/10.1111/j.1558-5646.2007.00239.x.

Meier, Rudolf, Marion Kotrba, and Paul Ferrar. 1999. "Ovoviviparity and Viviparity in the Diptera." *Biological Reviews* 74 (3): 199–258. https://doi.org/10.1111/j.1469-185X.1999.tb00186.x.

Meiklejohn, Kelly A., James F. Wallman, and Mark Dowton. 2011. "DNA-Based Identification of Forensically Important Australian Sarcophagidae (Diptera)." *International Journal of Legal Medicine* 125 (1): 27–32. https://doi.org/10.1007/s00414-009-0395-y.

Merkle, Daniel, and Martin Middendorf. 2005. "Reconstruction of the Cophylogenetic History of Related Phylogenetic Trees with Divergence Timing Information." *Theory in Biosciences* 123 (4): 277–99. https://doi.org/10.1016/j.thbio.2005.01.003.

Merz, Clayton, Julian M. Catchen, Victor Hanson-Smith, Kevin J. Emerson, William E. Bradshaw, and Christina M. Holzapfel. 2013. "Replicate Phylogenies and Post-Glacial Range Expansion of the Pitcher-Plant Mosquito, Wyeomyia Smithii, in North America." Edited by William J. Etges. *PLoS ONE* 8 (9): e72262. https://doi.org/10.1371/journal.pone.0072262.

Miller, Thomas E., William E. Bradshaw, and Christina M. Holzapfel. 2018. *Pitcher-Plant Communities as Model Systems for Addressing Fundamental Questions in Ecology and Evolution*. Vol. 1. Oxford University Press. https://doi.org/10.1093/oso/9780198779841.003.0024.

Mlynarek, Julia J., and Terry A. Wheeler. 2018. "Chloropid Flies (Diptera, Chloropidae) Associated with Pitcher Plants in North America." *PeerJ* 6 (March): e4491. https://doi.org/10.7717/peerj.4491.

Mulieri, Pablo Ricardo, Eduardo Federico Schaefer, Marta Inés Duré, and Cynthya Elizabeth González. 2018. "A New Flesh Fly Species (Diptera: Sarcophagidae) Parasitic on Leptodactylid Frogs." *Parasitology Research* 117 (3): 809–18. https://doi.org/10.1007/s00436-018-5755-4.

Naczi, Robert F.C., Eric M. Soper, Frederick W. Case, and Roberta B. Case. 1999. "*Sarracenia rosea* (Sarraceniaceae), a new species of pitcher plant from the southeastern United States." *SIDA, Contributions to Botany* 18 (4): 1183–1206.

Nguyen, Lam-Tung, Heiko A. Schmidt, Arndt von Haeseler, and Bui Quang Minh. 2015. "IQ-TREE: A Fast and Effective Stochastic Algorithm for Estimating Maximum-Likelihood Phylogenies." *Molecular Biology and Evolution* 32 (1): 268–74. https://doi.org/10.1093/molbev/msu300.

Nylin, Sören, Jessica Slove, and Niklas Janz. 2014. "Host Plant Utilization, Host Range Oscillations and Diversification in Nymphalid Butterflies: A Phylogenetic Investigation." *Evolution* 68 (1): 105–24. https://doi.org/10.1111/evo.12227.

Pape, T. 1990. "Revisionary Notes on American Sarcophaginae (Diptera: Sarcophagidae)." *Tijdschrift Voor Entomologie* 133 (1): 43–74.

Pape, T. 1994. "The World Blaesoxipha Loew, 1861 (Diptera : Sarcophagidae)." *Entomologica Scandinavica, Supplement* 45: 1–247.

Paradis, Emmanuel, and Klaus Schliep. 2019. "Ape 5.0: An Environment for Modern Phylogenetics and Evolutionary Analyses in R." *Bioinformatics* 35 (3): 526–28. https://doi.org/10.1093/bioinformatics/bty633.

Pecoraro, Carlo, Massimiliano Babbucci, Adriana Villamor, Rafaella Franch, Chiara Papetti, Bruno Leroy, Sofia Ortega-Garcia, et al. 2016. "Methodological Assessment of 2b-RAD Genotyping Technique for Population Structure Inferences in Yellowfin Tuna (Thunnus Albacares)." *Marine Genomics* 25 (February): 43–48. https://doi.org/10.1016/j.margen.2015.12.002.

Puritz, Jonathan B., Mikhail V. Matz, Robert J. Toonen, Jesse N. Weber, Daniel I. Bolnick, and Christopher E. Bird. 2014. "Demystifying the RAD Fad." *Molecular Ecology* 23 (24): 5937–42. https://doi.org/10.1111/mec.12965.

Rambaut, Andrew, Alexei J Drummond, Dong Xie, Guy Baele, and Marc A Suchard. 2018. "Posterior Summarization in Bayesian Phylogenetics Using Tracer 1.7." *Systematic Biology* 67 (5): 901–4. https://doi.org/10.1093/sysbio/syy032.

Rango, Jessamy J. 1999. "Resource Dependent Larviposition Behavior of a Pitcher Plant Flesh Fly, Fletcherimyia Fletcheri (Aldrich) (Diptera: Sarcophagidae)." *Journal of the New York Entomological Society* 107 (1): 82–86.

Reeves, Will, Thomas Pape, and Peter Adler. 2000. "Biological Notes on New World Sarcophagidae (Diptera)." *Studia Dipterologica* 7 (January): 497–500.

Riley, C.V. 1847. "Descriptions and Natural History of Two Insects Which Brave the Dangers of Sarracenia Variolaris." *Trans. Acad. Sci. St. Louis*, no. 3: 235–40.

Roback, Selwyn S. 1954. *The Evolution and Taxonomy of the Sarcophaginae Diptera, Sarcophagidae*. Urbana.

Robledo, Diego, Christos Palaiokostas, Luca Bargelloni, Paulino Martínez, and Ross Houston. 2018. “Applications of Genotyping by Sequencing in Aquaculture Breeding and Genetics.” *Reviews in Aquaculture* 10 (3): 670–82. https://doi.org/10.1111/raq.12193.

Sanders, Kate L., Michael S. Y. Lee, Mumpuni, Terry Bertozzi, and Arne R. Rasmussen. 2013. “Multilocus Phylogeny and Recent Rapid Radiation of the Viviparous Sea Snakes (Elapidae: Hydrophiinae).” *Molecular Phylogenetics and Evolution* 66 (3): 575–91. https://doi.org/10.1016/j.ympev.2012.09.021.

“Sanger Sequencing Services | Psomagen.” n.d. Accessed July 5, 2021. https://psomagen.com/research-sequencing-services/sanger-sequencing/.

Santichaivekin, Santi, Qing Yang, Jingyi Liu, Ross Mawhorter, Justin Jiang, Trenton Wesley, Yi-Chieh Wu, and Ran Libeskind-Hadas. 2020. “EMPRess: A Systematic Cophylogeny Reconciliation Tool.” *Bioinformatics*, no. btaa978 (November). https://doi.org/10.1093/bioinformatics/btaa978.

Satler, Jordan D., and Bryan C. Carstens. 2016. “Phylogeographic Concordance Factors Quantify Phylogeographic Congruence among Co-Distributed Species in the *Sarracenia alata* Pitcher Plant System: Community Phylogeography In A Pitcher Plant System.” *Evolution* 70 (5): 1105–19. https://doi.org/10.1111/evo.12924.

Satler, Jordan D., and Brayan C Carstens. 2019. “The *Sarracenia Alata* Pitcher Plant System and Obligate Arthropod Inquilines Should Be Considered an Evolutionary Community.” *Journal of Biogeography* 46 (2): 485–96. https://doi.org/10.1111/jbi.13498.

Schnell, D. E. 2002. “Carnivorous Plants of the United States and Canada.” *Carnivorous Plants of the United States and Canada.*, no. Ed.2. https://www.cabdirect.org/cabdirect/abstract/20023101134.

Seetharam, Arun S., and Gary W. Stuart. 2013. “Whole Genome Phylogeny for 21 Drosophila Species Using Predicted 2b-RAD Fragments.” *PeerJ* 1 (December): e226. https://doi.org/10.7717/peerj.226.

“Sequencher DNA Sequence Analysis Software from Gene Codes Corporation.” n.d. Accessed July 5, 2021. http://www.genecodes.com/.

Shaw, Kerry L. 2002. “Conflict between Nuclear and Mitochondrial DNA Phylogenies of a Recent Species Radiation: What MtDNA Reveals and Conceals about Modes of Speciation in Hawaiian Crickets.” *Proceedings of the National Academy of Sciences* 99 (25): 16122–27.

Shewell, G. E. 1987. “Sarcophagidae.” In J. F. McAlpine, B. V. Peterson, G. E. Shewell, H. J. Teskey, J. R. Vockeroth, and D. M. Wood [Eds.], *Manual of Nearctic Diptera.*, 2:1159–86. 28. Biosystematic Research Institute.

Simon, Chris, Francesco Frati, Andrew Beckenbach, Bernie Crespi, Hong Liu, and Paul Flook. 1994. “Evolution, Weighting, and Phylogenetic Utility of Mitochondrial Gene Sequences and a Compilation of Conserved Polymerase Chain Reaction Primers.” *Annals of the Entomological Society of America* 87 (6): 651–701. https://doi.org/10.1093/aesa/87.6.651.

Stamper, Trevor, Gregory A. Dahlem, Clifford Cookman, and Ronald W. Debry. 2013. "Phylogenetic Relationships of Flesh Flies in the Subfamily Sarcophaginae Based on Three MtDNA Fragments (Diptera: Sarcophagidae)." *Systematic Entomology* 38 (1): 35–44. https://doi.org/10.1111/j.1365-3113.2012.00646.x.

Stephens, Jessica D., Willie L. Rogers, Karolina Heyduk, Jennifer M. Cruse-Sanders, Ron O. Determann, Travis C. Glenn, and Russell L. Malmberg. 2015. "Resolving Phylogenetic Relationships of the Recently Radiated Carnivorous Plant Genus Sarracenia Using Target Enrichment." *Molecular Phylogenetics and Evolution* 85 (April): 76–87. https://doi.org/10.1016/j.ympev.2015.01.015.

Stevens, Jamie R., and James F. Wallman. 2006. "The Evolution of Myiasis in Humans and Other Animals in the Old and New Worlds (Part I): Phylogenetic Analyses." *Trends in Parasitology* 22 (3): 129–36. https://doi.org/10.1016/j.pt.2006.01.008.

Stireman, J. O. 2005. "The Evolution of Generalization? Parasitoid Flies and the Perils of Inferring Host Range Evolution from Phylogenies." *Journal of Evolutionary Biology* 18 (2): 325–36. https://doi.org/10.1111/j.1420-9101.2004.00850.x.

Stucky, Brian J. 2015. "Infection Behavior, Life History, and Host Parasitism Rates of Emblemasoma Erro (Diptera: Sarcophagidae), an Acoustically Hunting Parasitoid of the Cicada Tibicen Dorsatus (Hemiptera: Cicadidae)." *Zoological Studies* 54 (1): 30. https://doi.org/10.1186/s40555-015-0105-z.

Swenson, Nathan G., and Daniel J. Howard. 2005. "Clustering of Contact Zones, Hybrid Zones, and Phylogeographic Breaks in North America." *The American Naturalist* 166 (5): 581–91. https://doi.org/10.1086/491688.

Townsend, C.H.T. 1917. "Genera of the Dipterous Tribe Sarcophagini." *Proceedings of the Biological Society of Washington* 30: 189–98.

Tripp, Erin A., Yi-Hsin Erica Tsai, Yongbin Zhuang, and Kyle G. Dexter. 2017. "RADseq Dataset with 90% Missing Data Fully Resolves Recent Radiation of Petalidium (Acanthaceae) in the Ultra-Arid Deserts of Namibia." *Ecology and Evolution* 7 (19): 7920–36. https://doi.org/10.1002/ece3.3274.

Underwood, Angela Spano. 2009. "Larvae of Sarcophagidae (Insecta: Diptera) and Their Relationship with the Pitcher Plants (Sarraceniaceae: Sarracenia) of Southeastern U.S. Bogs." Thesis. https://etd.auburn.edu//handle/10415/1889.

Venkatachalam, Balaji, Jim Apple, Katherine St. John, and Daniel Gusfield. 2010. "Untangling Tanglegrams: Comparing Trees by Their Drawings." *IEEE/ACM Transactions on Computational Biology and Bioinformatics* 7 (4): 588–97. https://doi.org/10.1109/TCBB.2010.57.

Vollrath, F. 1984. "Kleptobiotic Interactions in Invertebrates." In *Barnard CJ (Ed) Producers and Scroungers; Strategies of Exploitation and Parasitism.*, 61–94. New York, NY: Chapman and Hall.

Wagner, Catherine E., Irene Keller, Samuel Wittwer, Oliver M. Selz, Salome Mwaiko, Lucie

Greuter, Arjun Sivasundar, and Ole Seehausen. 2013. “Genome-Wide RAD Sequence Data Provide Unprecedented Resolution of Species Boundaries and Relationships in the Lake Victoria Cichlid Adaptive Radiation.” *Molecular Ecology* 22 (3): 787–98. https://doi.org/10.1111/mec.12023.

Wang, Shi, Pingping Liu, Jia Lv, Yangping Li, Taoran Cheng, Lingling Zhang, Yu Xia, Hongzhen Sun, Xiaoli Hu, and Zhenmin Bao. 2016. “Serial Sequencing of Isolength RAD Tags for Cost-Efficient Genome-Wide Profiling of Genetic and Epigenetic Variations.” *Nature Protocols* 11 (11): 2189–2200. https://doi.org/10.1038/nprot.2016.133.

Wang, Shi, Eli Meyer, John K. McKay, and Mikhail V. Matz. 2012. “2b-RAD: A Simple and Flexible Method for Genome-Wide Genotyping.” *Nature Methods* 9 (8): 808–10. https://doi.org/10.1038/nmeth.2023.

Wells, J. D., T. Pape, and F. a. H. Sperling. 2001. “DNA-Based Identification and Molecular Systematics of Forensically Important Sarcophagidae (Diptera).” *Journal of Forensic Science* 46 (5): 1098–1102. https://doi.org/10.1520/JFS15105J.

Wiens, John J. 2003. “Missing Data, Incomplete Taxa, and Phylogenetic Accuracy.” *Systematic Biology* 52 (4): 528–38. https://doi.org/10.1080/10635150390218330.

Wray, D. L., and C. S. Brimley. 1943. “The Insect Inquilines and Victims of Pitcher Plants in North Carolina.” *Annals of the Entomological Society of America* 36 (1): 128–37. https://doi.org/10.1093/aesa/36.1.128.

Zhang, Changquan, Xiaoliang Fu, Kai Xie, Weitao Yan, and Yadong Guo. 2015. “MtDNA Analysis for Genetic Identification of Forensically Important Sarcophagid Flies (Diptera: Sarcophagidae) in China.” *Journal of Medical Entomology* 52 (6): 1225–33. https://doi.org/10.1093/jme/tjv131.

Zhang, Jiajie, Kassian Kobert, Tomáš Flouri, and Alexandros Stamatakis. 2014. “PEAR: A Fast and Accurate Illumina Paired-End ReAd MergeR.” *Bioinformatics* 30 (5): 614–20. https://doi.org/10.1093/bioinformatics/btt593.

Tables and Figures

***Fletcherimyia* species**	**Host *Sarracenia* species**	**States within fly range**
F. abdita	*S. leucophylla, S. alata, S. flava, S. rubra wherryi, S. minor*	AL, FL, GA, MS
F. celarata	*S. leucophylla*	AL, FL, MS, TX
F. fletcheri	*S. purpurea*	NC, northwards through southern Canada
F. folkertsi	*S. rosea*	AL, FL
F. jonesi	*S. minor, S. flava*	AL, FL, GA, SC, NC
F. oreophilae	*S. oreophila*	AL, NC
F. papei	*S. rubra, S. rubra wherryi*	AL, NC, TX
F. rileyi	*S. minor, S. flava*	AL, FL, GA, MS, NC, SC

Table 1. Current species constructs within the flesh fly genus *Fletcherimyia* and the species of *Sarracenia* that each utilizes (Dahlem and Naczi 2006). Also included are state-level ranges detailed in the same study.

Figure 1. *Fletcherimyia* species in their natural habitats. (A) *F. abdita* - *S. alata*, LA (B) *F. celarata* - *S. alata*, TX (C) *F. jonesi* - *S. minor*, SC (D) *F. fletcheri* - *S. purpurea* ssp. *purpurea*, OH (E) *F. rileyi* - *S. minor*, GA (F) *F. papei* - *S. rubra* ssp. *rubra*, SC (G) *F. folkertsi* - *S. rosea*, AL (H) *F. papei in copula* - near *S. rubra* ssp. *rubra*, SC (I) *F. celarata* larva in dissected pitcher of *S. alata*, TX. Not pictured is *F. oreophilae*.

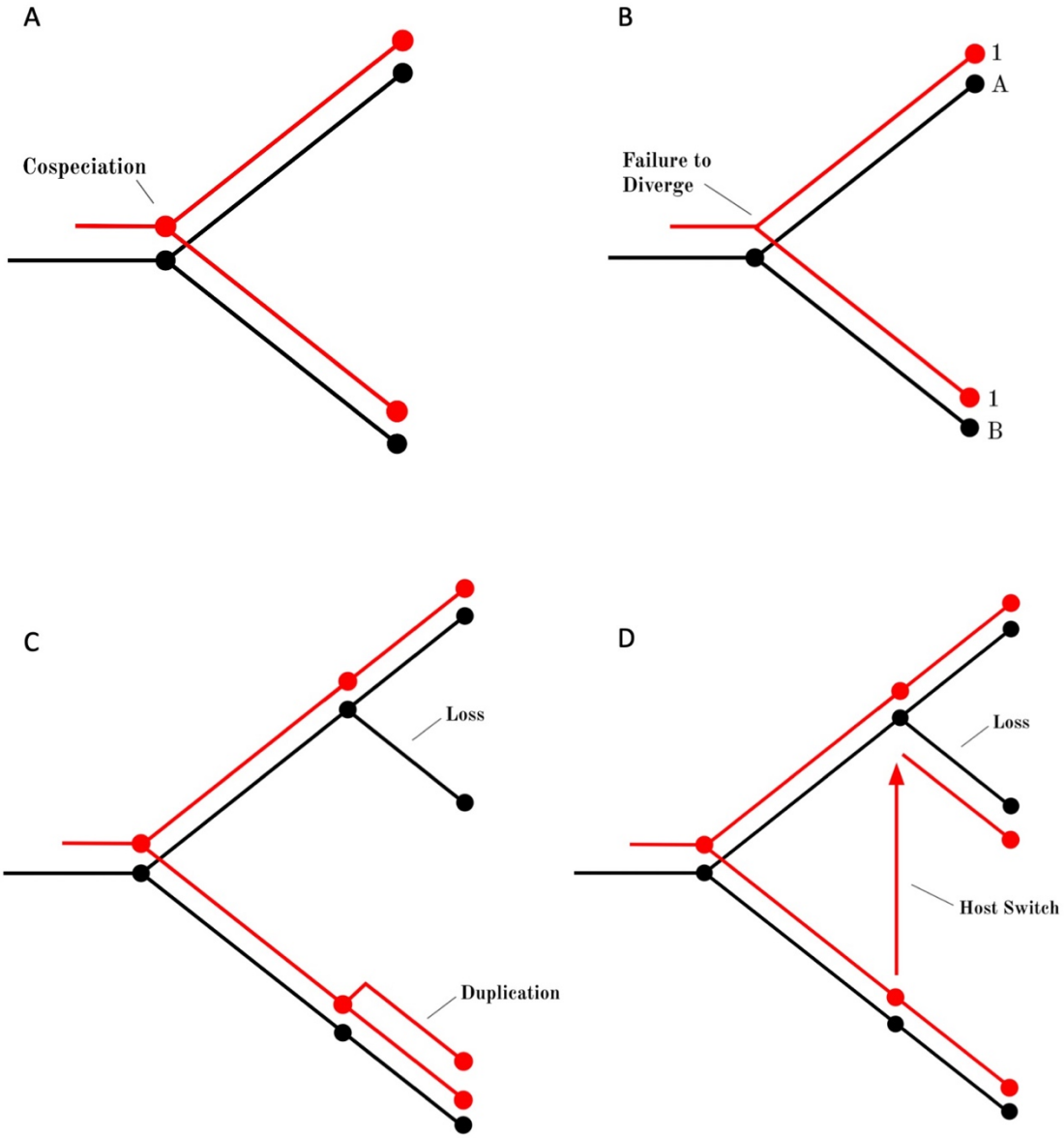


Figure 2. The five major events in cophylogenetic reconstructions. Shown in this figure are cospeciation (A) failure to diverge (B) loss (C,D) duplication (C) and host switching (D).

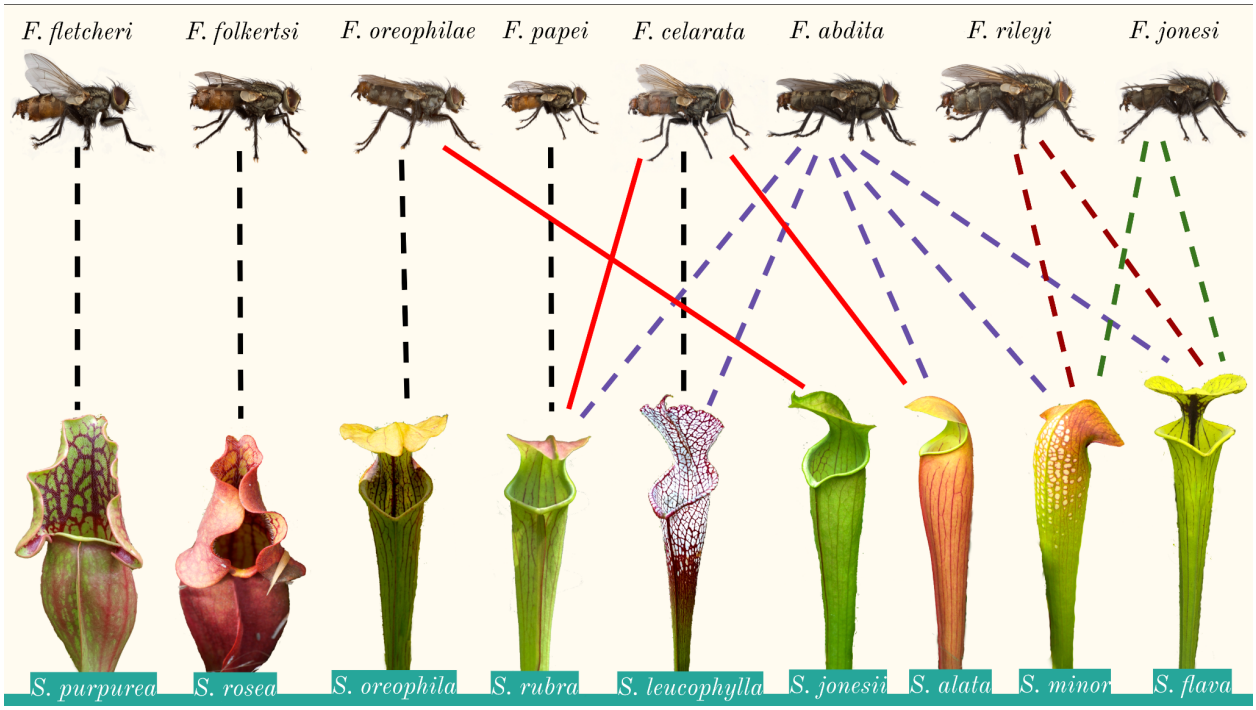


Figure 3. Known *Fletcherimyia* - *Sarracenia* affiliations. Dotted lines denote relationships described by Dahlem and Naczi (2006); solid red lines are new relationships recovered from mtDNA barcoding of larvae. Of note, the line between *F. celarata* and *S. rubra* represents the fly's relationship with *S. rubra* spp. *gulfensis*, which is paraphyletic with other *rubra* subspecies and possibly a distinct lineage.

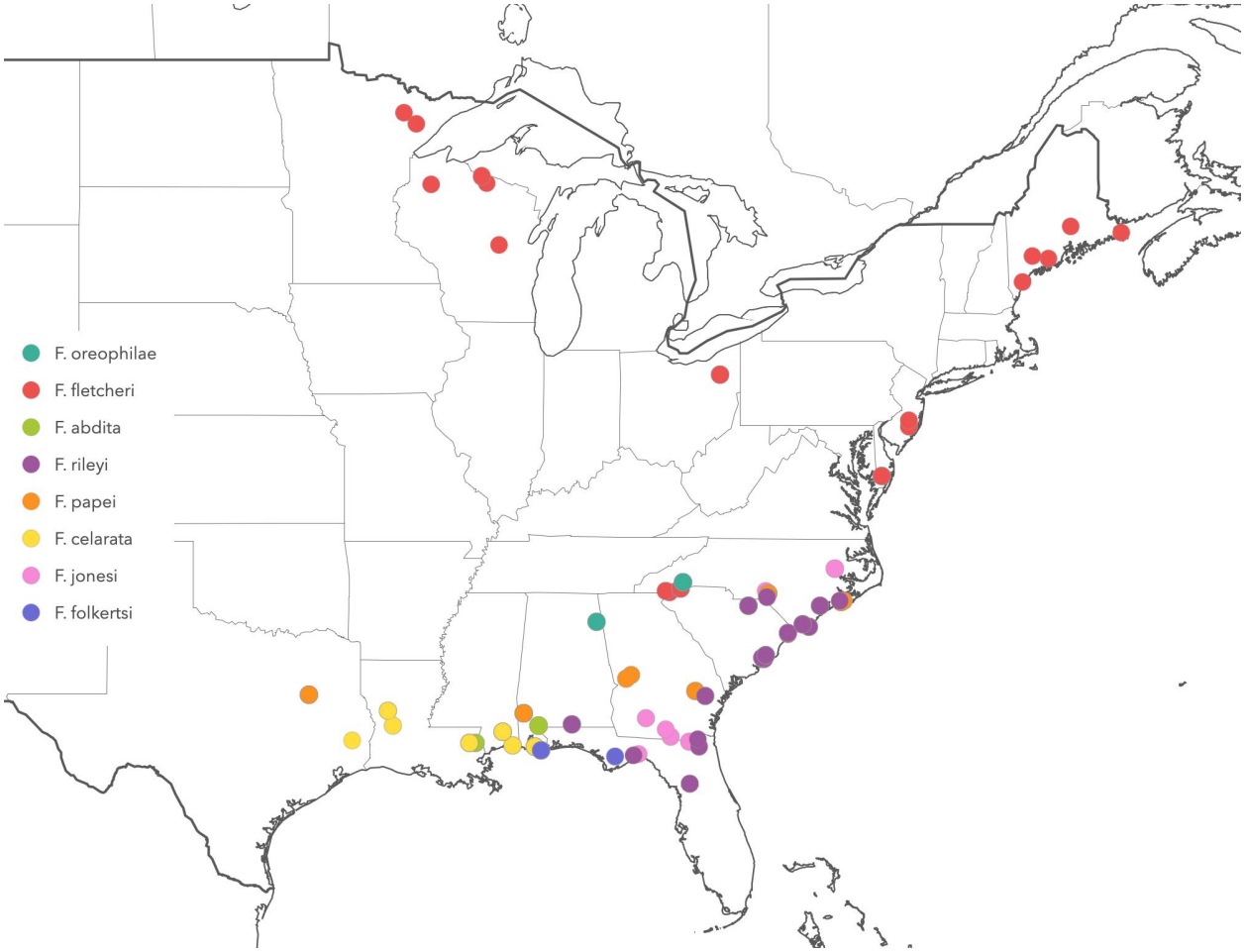


Figure 4. Map of sampling localities during the 2018-20 field seasons. Color denotes the most commonly recovered *Fletcherimyia* species at each site.

Figure 5. Pinned abdomen of adult male *Fletcherimyia rileyi* showing spread genitalia, exposing diagnostic structures (particularly cerci and phallus).

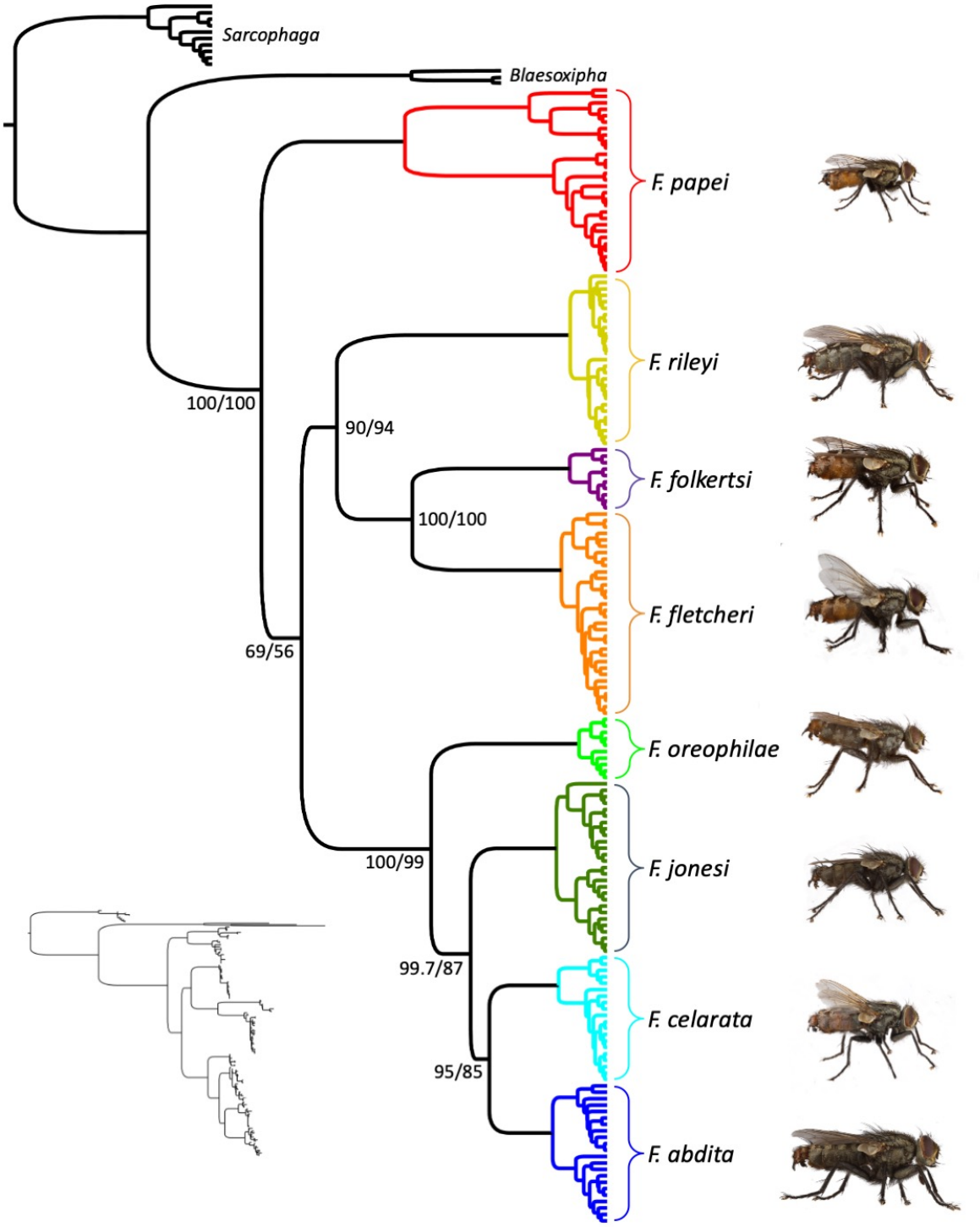


Figure 6. Left: Bayesian Inference phylogram of *cox1* sequence data from field specimens. Node labels denote Posterior Probability support / ML Bootstrap Support. Support values for individual species clades were 100% in each matrix and are not shown. ML topology is shown in the bottom left to illustrate branch lengths. Note the deep divergence between eastern and western populations of *F. papei* relative to population variation within other species. Right: Pinned adult males with genitalia spread.

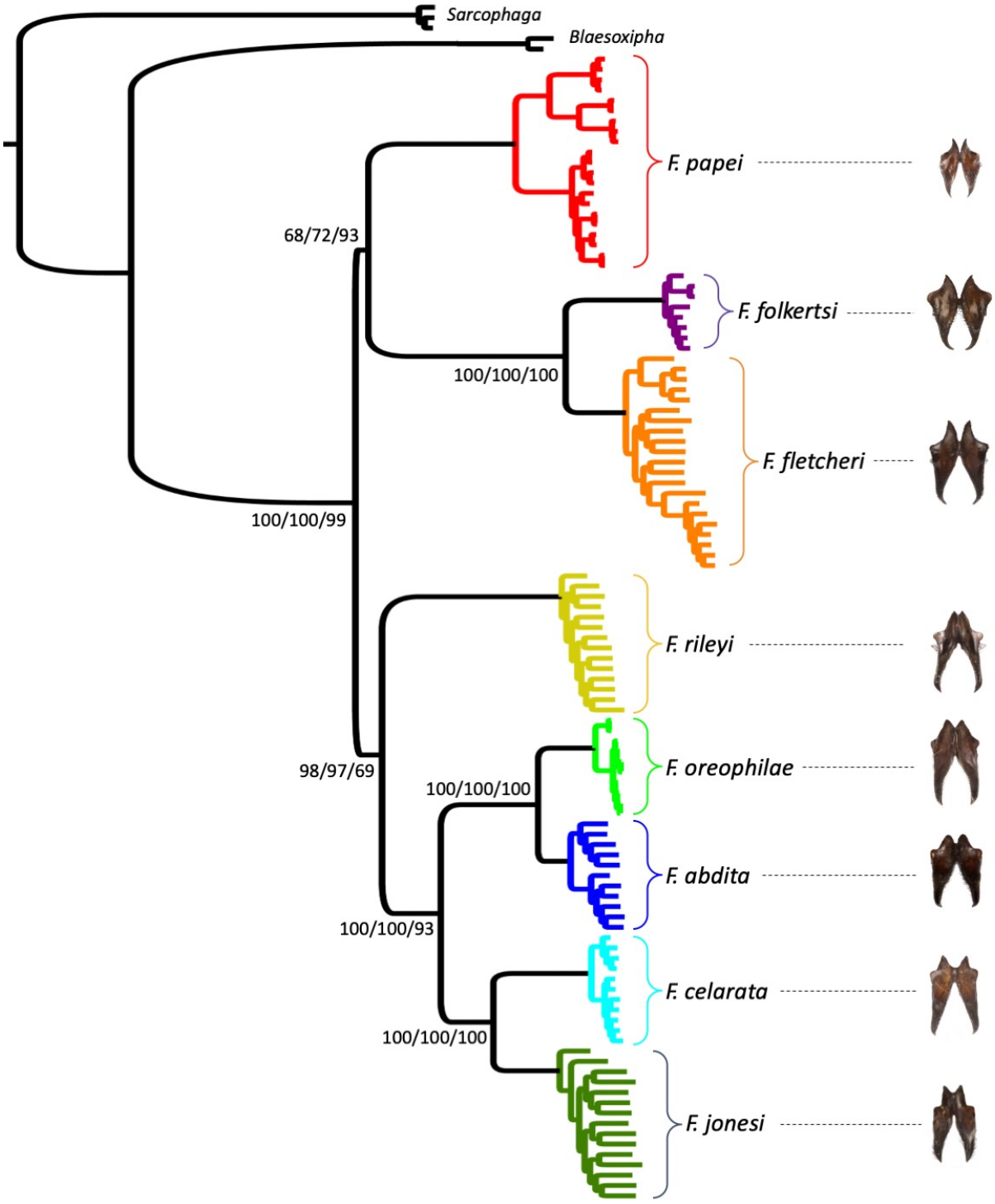


Figure 7. Left: Maximum Likelihood phylogram of 2bRAD sequence data from field specimens. Node labels denote bootstrap support (x1000) using alignment matrices with 55/86/91.7% missing data. Support values for individual species clades were 100% in each matrix and are not shown. Right: Male cerci, spread and dried - one of the main diagnostic structures between species.

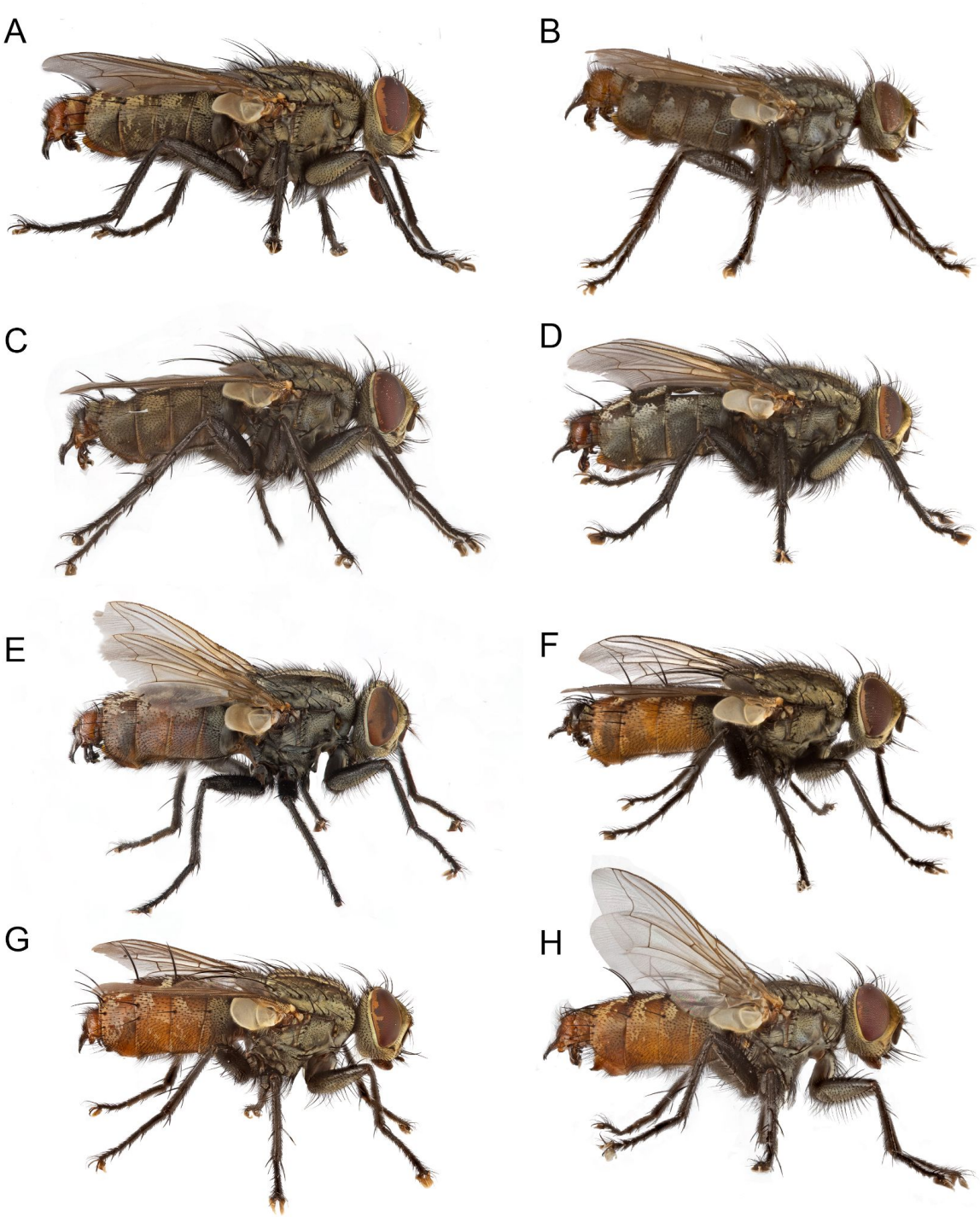


Figure 8. Pinned adult male *Fletcherimyia*, not shown at relative scale: (A) *F. abdita* (B) *F. oreophilae* (C) *F. jonesi* (D) *F. rileyi* (E) *F. celarata* (F) *F. papei* (G) *F. folkertsi* (H) *F. fletcheri*

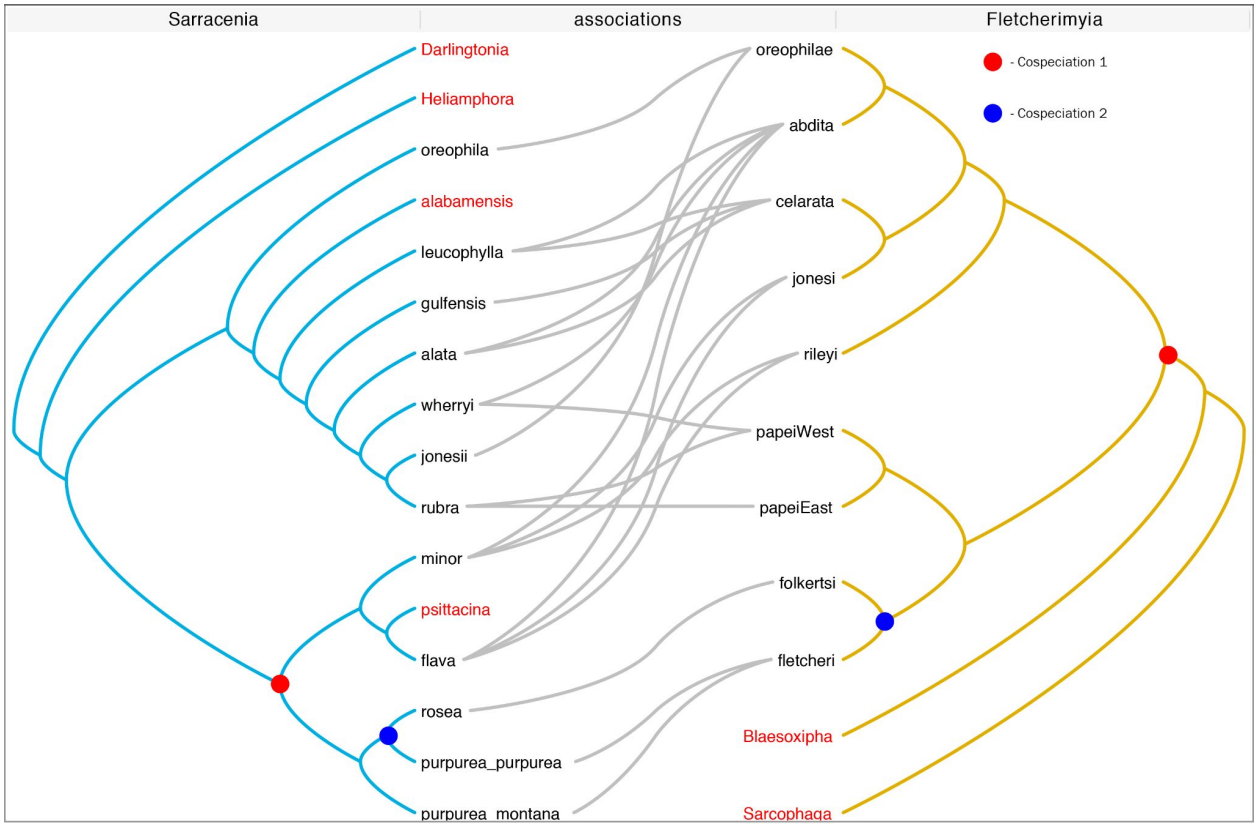


Figure 9. TreeMap tanglegram illustrating *Fletcherimyia* - *Sarracenia* interactions; cospeciation events as recovered by Jane are shown as circles on both phylogenies.

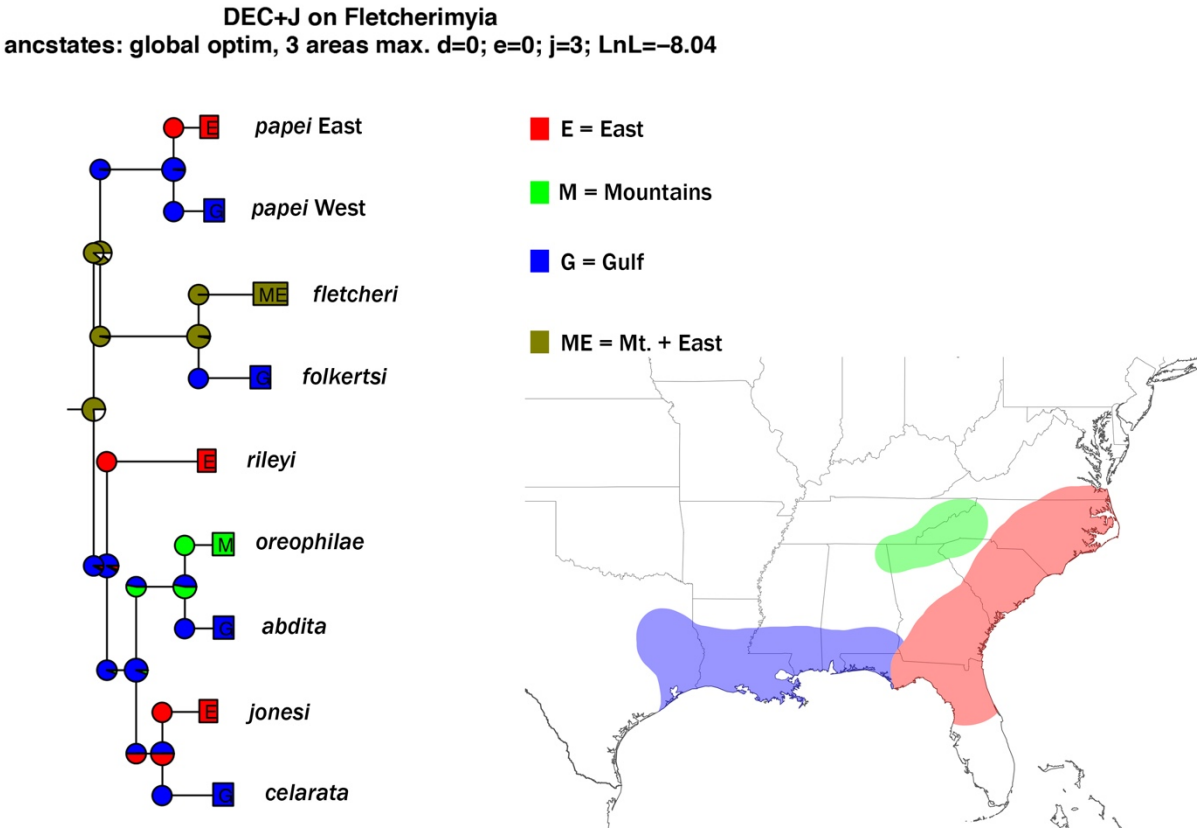


Figure 10. Historical biogeography of *Fletcherimyia* as reconstructed using BioGeoBears. The three modern geographic areas used to calibrate the analysis are shown on the right.

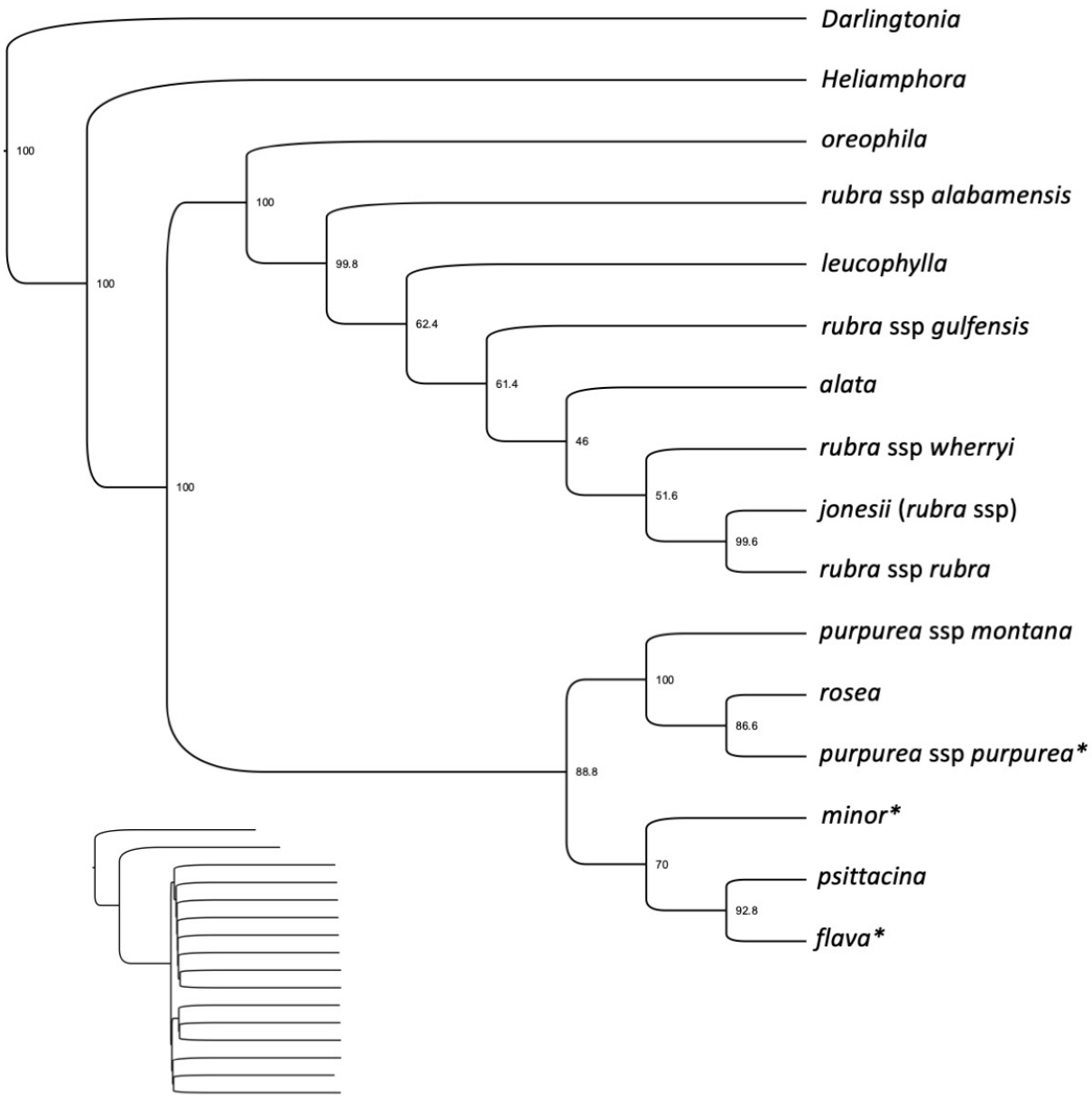


Figure 11. Maximum Pseudo-likelihood Estimated Species Tree (MP-EST) of *Sarracenia* from Stephens et. al. (2015), here shown as a cladogram to better illustrate species relationships. True branch lengths are shown in the bottom left; all *Sarracenia* species are separated by very short internodes indicative of their rapid diversification. This is likely responsible for relatively poor bootstrap support values, shown at the cladogram nodes. Asterisks denote collapsed monophyletic subspecies clades.

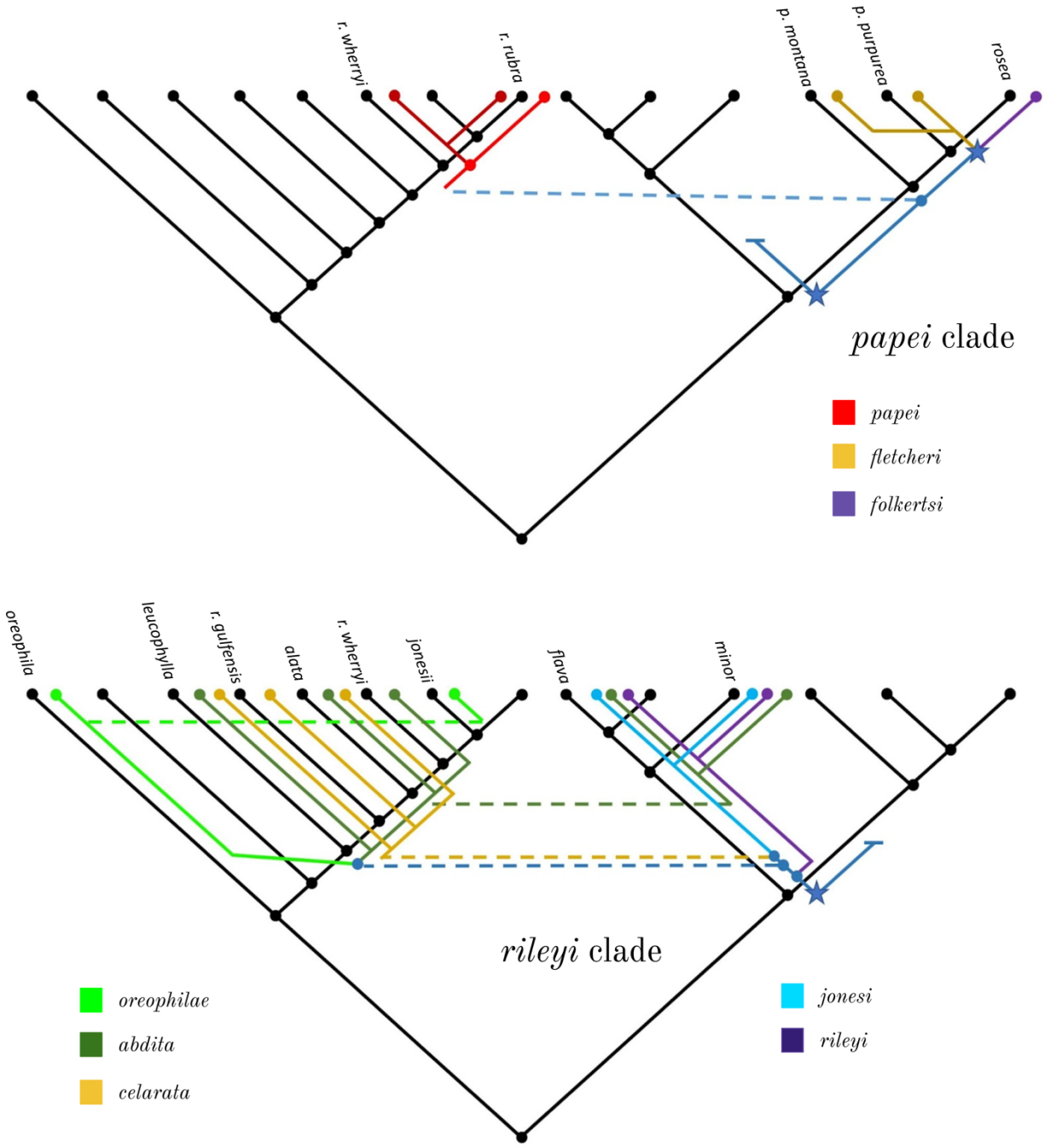


Figure 12. Event-based cophylogenetic reconstruction of the *Sarracenia-Fletcherimyia* system, based on results from Jane 4.0, split between the two major fly clades. Sarracenia trees are shown in black, Fletcherimyia trees are colored by species as labeled. Starred nodes denote cospeciation between flies and pitchers, and dotted lines denote host-switching events. Fly nodes without circles denote failures to diverge.