Molecular Phylogenetics and Evolution 85 (2015) 76–87

Contents lists available at ScienceDirect

Molecular Phylogenetics and Evolution

journal homepage: www.elsevier.com/locate/ympev

Resolving phylogenetic relationships of the recently radiated carnivorous plant genus *Sarracenia* using target enrichment

Jessica D. Stephens [a,*], Willie L. Rogers [a], Karolina Heyduk [a], Jennifer M. Cruse-Sanders [b], Ron O. Determann [b], Travis C. Glenn [c], Russell L. Malmberg [a]

[a] Department of Plant Biology, University of Georgia, Athens, GA 30602, United States
[b] Atlanta Botanical Garden, 1345 Piedmont Ave NE, Atlanta, GA 30309, United States
[c] Department of Environmental Health Science, University of Georgia, Athens, GA 30602, United States

ARTICLE INFO

Article history:
Received 5 October 2014
Revised 23 January 2015
Accepted 27 January 2015
Available online 14 February 2015

Keywords:
Gene capture
Coalescent
Phylogenomics
Pitcher plants
Species tree estimation

ABSTRACT

The North American carnivorous pitcher plant genus *Sarracenia* (Sarraceniaceae) is a relatively young clade (<3 million years ago) displaying a wide range of morphological diversity in complex trapping structures. This recently radiated group is a promising system to examine the structural evolution and diversification of carnivorous plants; however, little is known regarding evolutionary relationships within the genus. Previous attempts at resolving the phylogeny have been unsuccessful, most likely due to few parsimony-informative sites compounded by incomplete lineage sorting. Here, we applied a target enrichment approach using multiple accessions to assess the relationships of *Sarracenia* species. This resulted in 199 nuclear genes from 75 accessions covering the putative 8–11 species and 8 subspecies/varieties. In addition, we recovered 42 kb of plastome sequence from each accession to estimate a cpDNA-derived phylogeny. Unsurprisingly, the cpDNA had few parsimony-informative sites (0.5%) and provided little information on species relationships. In contrast, use of the targeted nuclear loci in concatenation and coalescent frameworks elucidated many relationships within *Sarracenia* even with high heterogeneity among gene trees. Results were largely consistent for both concatenation and coalescent approaches. The only major disagreement was with the placement of the *purpurea* complex. Moreover, results suggest an Appalachian massif biogeographic origin of the genus. Overall, this study highlights the utility of target enrichment using multiple accessions to resolve relationships in recently radiated taxa.

1. Introduction

The evolution of carnivory in angiosperms has long fascinated evolutionary biologists, with the most notable being Charles Darwin (1875). This interest partially stems from the complex trapping structures used in attraction, retention, and digestion of prey and subsequent absorption of nutrients (Albert et al., 1992; Juniper et al., 1989). These carnivorous adaptations to nutrient poor habitats have independently evolved six times in flowering plants, resulting in approximately 645 species, which often display tremendous morphological diversity at both the infrafamilial and infrageneric level (Albert et al., 1992; Ellison and Gotelli, 2009). Insight into the patterns of structural evolution and diversification across these groups requires an explicit understanding of their evolutionary relationships. Phylogenies currently exist for many carnivorous genera including *Utricularia* (bladderworts, Jobson et al., 2003; Müller and Borsch, 2005), *Drosera* (sundews, Rivadavia et al., 2003), and *Nepenthes* (Old World pitcher plants, Meimberg et al., 2001), yet the evolutionary relationships of one of the more well-studied genera, *Sarracenia* (New World pitcher plants), remain largely ambiguous.

Sarracenia is the most recently diverged group of the three extant genera within the family Sarraceniaceae (Ellison et al., 2012; Neyland and Merchant, 2006). All species within Sarraceniaceae are carnivorous with no geographical overlap among genera. The basal monotypic lineage, *Darlingtonia californica*, is restricted to serpentine seeps in Oregon and California, while the estimated 15 *Heliamphora* species are confined to the Guiana Highlands tepuis in South America (McPherson, 2007). *Sarracenia* is endemic to seepage slopes, wet pine savannas, and fens of North America, predominately the southeastern United States Coastal Plain with one subspecies, *purpurea* ssp. *purpurea*, extending into the

* Corresponding author.
E-mail addresses: jstephens@plantbio.uga.edu (J.D. Stephens), bonjour@uga.edu (W.L. Rogers), kheyduk@plantbio.uga.edu (K. Heyduk), JSanders@atlantabotanicalgarden.org (J.M. Cruse-Sanders), rdetermann@atlantabotanicalgarden.org (R.O. Determann), travisg@uga.edu (T.C. Glenn), russell@plantbio.uga.edu (R.L. Malmberg).

http://dx.doi.org/10.1016/j.ympev.2015.01.015
1055-7903/

northeastern United States and southern Canada. Unfortunately these habitats are being destroyed and estimates suggest less than 3% of historic *Sarracenia* habitat remains (Folkerts, 1982; Folkerts and Folkerts, 1993). This continued habitat loss has resulted in the U.S. Fish and Wildlife and Convention on International Trade in Endangered Species (CITES) listing of three endangered taxa within *Sarracenia* and one taxa considered a candidate for listing (www.cites.org). Complicating protection status of other members of this genus is the disagreement among sources in the number of recognized species, subspecies, and varieties with numbers ranging between 8–11 species and as many as 41 subspecies, varieties, and forms (Ellison et al., 2014).

Previous attempts at constructing a phylogeny for *Sarracenia* from nuclear (Ellison et al., 2012; Neyland and Merchant, 2006), chloroplast (Bayer et al., 1996; Ellison et al., 2012), and mitochondrial regions (Ellison et al., 2012) have been inconsistent, typically with numerous polytomies within the genus. In addition, the relatively short branch lengths dated at roughly 0.5–3 million years ago (mya) (Ellison et al., 2012) indicate that this group may have undergone a recent, rapid diversification. Further complicating phylogenetic resolution is frequent hybridization among sympatric species (Furches et al., 2013; Mellichamp and Case, 2009). Both short branches and hybridization can have dramatic effects on species tree estimation. In particular, a recent radiation increases the chance that genes retain ancestral polymorphisms, resulting in incomplete lineage sorting (Pamilo and Nei, 1988); additionally, hybridization can lead to reticulation within gene trees (Hennig, 1966). Using multilocus data and modeling differences in gene history with use of the multispecies coalescent model can mitigate these potential sources of gene tree discordance within the species tree (Degnan and Rosenberg, 2009; Knowles, 2009; Liu et al., 2009). Increasing loci is expected to produce more accurate model parameters and therefore increase nodal support values in phylogenetic analyses (Maddison, 1997; Song et al., 2012), and use of multispecies coalescence has repeatedly outperformed concatenation methods under simulated and empirical data (Kubatko and Degnan, 2007; McCormack et al., 2012; Song et al., 2012). Including multiple accessions per species can also decrease the variance around the effective population size parameter within the coalescent framework (Heled and Drummond, 2010).

To further our understanding of evolutionary relationships of *Sarracenia* we conducted target enrichment of nuclear genes from multiple accessions per species sequenced on an Illumina HiSeq platform. Target enrichment involves the use of oligonucleotide probes that retain selected genomic regions for sequencing while reducing non-selected DNA (Mamanova et al., 2010). Target enrichment is highly applicable for phylogenetics as it works well for non-model organisms, is cost-efficient, and allows for an increase in the number of species and individuals for phylogenetic analysis (Faircloth et al., 2012a; Lemmon and Lemmon, 2013). Here, we (1) assessed the utility of this method for a recently radiated, non-model genus, (2) compared the multispecies coalescent approach with a concatenation approach, and (3) determined the evolutionary relationships within *Sarracenia*. The resolved species level phylogeny is then discussed in regard to the current taxonomy, biogeography, and conservation status of this group. Taken together, this multilocus and multiaccessional approach represents the most robust attempt to resolve the *Sarracenia* phylogeny to date and has implications for other recently radiated groups.

2. Material and methods

2.1. Taxon sampling

The majority of leaf tissue was sampled from the Atlanta Botanical Garden, which maintains an extensive living collection of *Sarracenia* species from various localities for conservation and as a reference for the North American Plant Collections Consortium. The remaining samples were collected from plant stocks maintained at the University of Georgia Plant Biology greenhouse and field collections. Current estimates list between 8 to 11 species with many varieties and subspecies being designated to the species level based on differing taxonomic schemes (Ellison et al., 2014). We sampled 71 *Sarracenia* accessions covering putative species, varieties, and subspecies. These include the eleven species recognized by Mellichamp and Case (2009) (*alabamensis*, *alata*, *flava*, *jonesii*, *leucophylla*, *minor*, *oreophila*, *psittacina*, *purpurea*, *rosea*, *rubra*) with 1–8 localities spanning the southeastern range of each species (see Table A.1) and additional samples from Maryland, Nova Scotia, and Wisconsin for *purpurea* ssp. *purpurea*. The 71 accessions also include three subspecies/varieties from the *purpurea* complex (ssp. *purpurea*, ssp. *venosa*, ssp. *venosa* var. *montana*), two subspecies from the *rubra* complex (ssp. *gulfensis*, ssp. *wherryi*), one *minor* variety (var. *okefenokeensis*), and two *flava* varieties (var. *rugelii*, var. *rubricorpora*). These putative subspecies and varieties are based on a combination of taxonomic descriptions between Mellichamp and Case (2009) and McPherson and Schnell (2011). Taxonomic descriptions have frequently designated *alabamensis* and *jonesii* as subspecies within the *rubra* complex and *rosea* as *purpurea* ssp. *venosa* var. *burkii*. Three *Darlingtonia californica* and one *Heliamphora minor* (both within Sarraceniaceae) were used as outgroups for the genus. This coverage of varieties, subspecies, and range distribution of putative species allows for a comprehensive analysis of this genus. Voucher specimens were deposited in either the University of Georgia Herbarium (UGA) or the Texas A&M Herbarium (TAES) (Table A.1).

2.2. Probe design

Targets for enrichment were initially identified by aligning *Sarracenia psittacina* and *S. purpurea* transcriptomes (Srivastava et al., 2011). All repeat-like regions were masked using RepeatMasker (http://www.repeatmasker.org/) prior to probe design. Targets with promising single nucleotide polymorphisms for phylogenetic analyses and at least two independent reads from each species were selected for further processing (~1000 contigs). Because paralogous sequences are not ideal for phylogenetic inference due to their independent evolutionary histories, potential targets were screened for paralogous signals using two methods. First, a within-species BLAST (Altschul et al., 1997) search of possible paralogous sequences was conducted with a stringent e-value cut off of $<3 \times 10^{-20}$. A reciprocal best BLAST(blastn) hit approach was then used on the subsequent targets to determine orthologous sequences between the two species. Targets that did not meet the cut off criterion were discarded from the potential target database; this resulted in 646 genes for target sequencing. Previous work suggests that *Sarracenia* may be a partial polyploid (Srivastava et al., 2011); however, we are confident that our stringent screening of paralogs prior to probe design and additional downstream removal of duplicates adequately addresses this possible source of conflict. Approximately three 120-mer oligonucleotide probes were designed for each gene per the manufacturer's probe design specifications. These probes were commercially synthesized by Mycroarray® into a custom MYbaits kit (http://www.mycroarray.com; Ann Arbor, MI).

2.3. DNA extraction, library preparation, sequencing

All leaves (i.e. pitchers) were cut near the base of the plant, sliced open, and cleaned of any insect residue, algae, soil, and other particulates. Areas of the leaf that were senescing, discolored, or greatly impacted from decomposing insect prey were removed

and discarded. The subsequent leaves were ground to powder using liquid nitrogen. Initially DNA extractions were conducted using 'option Y' from Peterson et al. (2000), but we were unable to extract high-quality DNA from older *Sarracenia* tissue and outgroups. Therefore, the majority of extractions were performed following the methods described in Lodhi et al. (1994) with slight modifications. Specifically, we replaced the 5 M sodium chloride solution with 3 M sodium acetate, and used two consecutive 2 mL treatments with 24 parts chloroform to 1 part octanol instead of a single 6 mL purification. We also used 1 volume cold iso-propanol in the final spin at 13,000 rpm to precipitate DNA. All DNA extractions were assessed for concentration and purity using a NanoDrop 2000 (Thermo Scientific, Wilmington, DE, USA). Samples were then sheared to approximately 180–500 bp lengths on a Bioruptor Standard (Diagenode Cat No. UCD-200, Woburn, MA). Library construction was completed using a protocol developed by Glenn et al. (submitted for publication). This protocol consists of removing overhangs created from shearing, phosphorylating the 5′ ends and adding a single adenosine to the 3′ end, ligating unique Illumina adapters with custom 10nt indexes to DNA fragments, and finally amplification of ligated DNA fragments with universal p5 and p7 primers (Faircloth and Glenn, 2012) to create uniquely indexed Illumina TruSeqHT compatible libraries. Samples with similar NanoDrop readings were combined at equal ratios resulting in two or three indexed individuals per tube. The MYbaits protocol was followed per manufacturer's instructions. To reduce daisy chaining during the hybridization, a blocking oligonucleotide with 10 inosines at the index location was used (c.f., Faircloth et al., 2012b). After target enrichment the subsequent libraries were sequenced on an Illumina HiSeq PE100 arranged by the Georgia Genomic Facility. All raw reads were deposited in NCBI Short Read Archive (accession numbers listed in Table A.1).

2.4. Assembly and alignment

All demultiplexed pooled reads were assessed for quality using FastQC v0.10.1 (http://www.bioinformatics.babraham.ac.uk/). Sequence reads were trimmed at the 3′ end with a Phred score of <20 to a minimum length of 40 bp using FastX v 0.013.2 (http://hannonlab.cshl.edu/fastx_toolkit/). Finally, all Illumina adapter contaminants were removed via FAR v2.15.

We used a combination of two approaches for target enrichment assembly similar to Heyduk et al. (in press). First, a *de novo* assembly method was conducted using Trinity version r2013-02-25 (Grabherr et al., 2011). The Trinity assembler was designed for RNA-seq, therefore Trinity tends to make multiple isoforms within contigs to take into account alternative splicing. To remove poorly supported isoforms (<1% per-component read support) the Trinity output file was parsed using RSEM in Bowtie 1.1.0 as described in the Trinity manual (Langmead et al., 2009). The second approach was a reference method, whereby the targeted genes used for probe design were used by the Columbus extension module (Zerbino, 2010) in VELVET to aid in assembly of reads. For each library, the k-mer length was optimized via KmerGenie (Chikhi and Medvedev, 2013) and then assembled with VELVET v 1.2.08 (Zerbino and Birney, 2008).

Contigs from both assembly methods that had at least 95% identity over 20 bp were merged using CAP3 v10/2011 (Huang and Madan, 1999). The resulting contigs were matched against the gene targets used for probe designs using BLAST (Altschul et al., 1997) and extracted for use via two steps. First, contigs were extracted if they had a 1:1 hit with the gene target. Second, a separate BLAST output containing instances where two contigs from an accession had best hits to the same target was created. If the two hits were non-overlapping (possibly due to intron regions) they were extracted and concatenated. The extracted contigs from each BLAST output were merged, renamed according to gene target, parsed into gene files, and aligned via Prank v100802 (Löytynoja and Goldman, 2008). The subsequent aligned gene files were filtered to remove poorly aligned regions using Gblocks v0.91b (Castresana, 2000). Finally, pairwise distances were calculated for each gene file. Alignments with an average pairwise distance of less than 0.35, at least 50% of the accessions present, showing no more than 45% missing data, and with at least one outgroup were used for subsequent downstream analyses.

2.5. Species tree analysis

Prior to species tree estimation, five genes were randomly selected to determine the model of evolution in JModelTest 2.1.1 using AIC (Darriba et al., 2012). All five genes had the same best-fitting model (GTRGAMMA) which was used for the whole dataset and implemented in RAxML v04/26/2012 (Stamatakis, 2006) for gene tree estimation using 500 bootstraps. These gene trees were then used as input data for species tree estimation. Species tree estimations were conducted using the Maximum Pseudolikelihood Estimation of the Species Tree v1.4 (MP-EST) method (Liu et al., 2010). MP-EST accounts for gene tree discordance resulting from incomplete lineage sorting and implements the triplet algorithm across gene trees to estimate the species tree topology. We used two approaches for phylogenetic analysis in MP-EST: one in which accessions were grouped into putative taxonomic designations by Mellichamp and Case (2009) and McPherson and Schnell (2011) and a second approach that treats all accessions as terminal taxa. The latter method was used to examine whether accessions exhibited reciprocal monophyly within their taxonomic groupings.

To test hypotheses of hybrid species within the genus, we used the Species Tree Estimation using Maximum Likelihood with hybridization (STEM-hy) method (Kubatko, 2009). This method uses a model-selection framework to evaluate hypotheses of hybridization in the presence of incomplete lineage sorting (Kubatko, 2009). STEM-hy requires estimates of theta (θ) and rate multipliers (*ri*). Theta (θ) was estimated and averaged from five gene trees using MIGRATE-n version 3.6.4 (Beerli, 2009), which estimates population parameters using maximum likelihood estimation under a coalescent framework (Beerli and Felsenstein, 2001). Estimation of rate multiplier (*ri*) values for each gene tree were calculated as the average divergence from the outgroups (Yang, 2002). The MP-EST species tree was used as the input species tree for STEM-hy. Tests of possible hybrid species were conducted on taxa that showed incongruence between the concatenated tree and MP-EST tree.

To assess discordance among genes, custom Perl scripts were used to query the presence and support values of nodes in consensus gene trees produced in RAxML. Five nodes of particular interest were queried based on conflict seen between the MP-EST accession tree estimation and the concatenation (see below) methodology. Nodes labeled in Fig. 2 were checked in all gene trees against either the MP-EST tree (nodes A and B) or the concatenated tree (nodes C, D, and E). Gene trees were classified as follows: a gene tree which had the node present with bootstrap support >80 "strongly agreed"; trees with support on the queried node between 50 and 80 "weakly agreed"; trees with support between 20 and 50 "weakly conflicted"; and trees with bootstrap support of less than 20 "strongly conflicted."

2.6. Concatenation analysis of nDNA and cpDNA

To compare with the coalescent analysis, we concatenated all 199 nuclear genes into a 'supergene' of 128,110 bp. This concatenated dataset was used to estimate an accession tree in RAxML with 1000 bootstraps. In addition to the targeted nuclear genes,

an average of 178,748 trimmed reads mapped back to a reference *Vitis vinifera* plastid genome (Jansen et al., 2006) using Bowtie2 v2.2.1 (Langmead and Salzberg, 2012). *Vitis vinifera* plastid genome was used as the reference due to high sequence similarity between *Vitis* and *S. psittacina*/*purpurea* found in Srivastava et al. (2011); furthermore, the placement of the Ericales is contested (Soltis and Soltis, 2004), though high sequence similarity between *Vitis* and *Sarracenia* species indicates it may be closer to rosids than asterids. These mapped reads were extracted from the trimmed reads for each accession using SAMtools (Li et al., 2009). Assembly of the chloroplast reads was conducted in the reference-based assembler YASRA (Ratan, 2009) with the *Vitis vinifera* plastid genome as a reference. Contigs were then BLASTed back to the reference, corrected for strandedness, and finally concatenated for each accession. The resulting concatenated sequence for each accession was aligned with the reference using MAFFT v7.029-e (Katoh and Standley, 2013) and poorly aligned regions were filtered using default setting with Gblocks v0.91b (Castresana, 2000) followed by visual inspection in Geneious v7.0.6 (http://www.geneious.com/). An accession chloroplast tree was estimated in RAxML v04/26/2012 (Stamatakis, 2006) under a GTRGAMMA model with 1000 bootstraps.

3. Results

3.1. Assembly and gene trees

Each accession had roughly 3.5 million trimmed reads which resulted in an average of 7124 contigs after Trinity assembly and 67,894 contigs from the Velvet assembly (Table A.1). These assemblies were subsequently merged into an average of 5608 contigs per accession with 546 contigs matching the 646 gene targets. The contigs on target had an average N50 of 503 bp with approximately 11× coverage.

In total 199 genes were used for subsequent phylogenetic analyses after poor alignments, genes with >50% missing data, and those missing an outgroup were discarded. Sixty-three (32%) of the 199 gene trees had all putative ingroup species represented, while 76 (38%) gene trees had one missing species, 44 (22%) had two missing species, 14 (7%) had three missing species, and 2 (1%) had four missing species. Among the 199 genes, an average of 56 accessions were present with a length of 642 bp per gene. This totaled 128,110 bp used for nuclear phylogenetic analyses. Of the 128,110 bp, 11,202 bp (8.7%) were variable with only 5066 bp (4%) being parsimony-informative within the ingroups. From the trimmed sequencing reads, we were able to recover 42,031 bp of the plastome, which contained 783 variable sites (~1.9%) and 216 parsimony-informative sites (~0.5%) within the ingroups. The 42,031 bp recovered consisted of intron and exon regions within the Long Single Copy and Short Single Copy segment of the chloroplast. Resulting gene trees, gene alignments, and species trees have been deposited in Dryad repository http://dx.doi.org/10.5061/dryad.nn153.

3.2. MP-EST species and accession tree

The MP-EST analyses supported the monophyly of the *Sarracenia* clade with *Darlingtonia californica* as basal to *Heliamphora* and *Sarracenia* (Figs. 1 and 2), which is consistent with the concatenated nDNA tree and plastid tree (Figs. 2 and 3). Additionally, MP-EST resolved many of the phylogenetic relationships within *Sarracenia* with high bootstrap support (Figs. 1 and 2). Specifically, both MP-EST analyses support an "*oreophila* clade" consisting of *oreophila* as sister to the *alata*, *leucophylla*, and the *rubra* complex (i.e. *alabamensis*, *jonesii*, ssp. *wherryi*, ssp. *gulfensis*) and another clade comprising *flava*, *minor*, *psittacina*, and the *purpurea* complex (ssp. *venosa* var. *montana*, ssp. *venosa*, ssp. *purpurea*, and *rosea*). Within the *oreophila* clade there was high bootstrap support for *alabamensis* sharing a more recent common ancestor with *leucophylla* (Fig. 1). This result suggests that the *rubra* complex is a polyphyletic group. It should be noted that there is low bootstrap support (<65) for relationships between *leucophylla*, *alata* and the other members of the *rubra* complex. In addition, the MP-EST accession tree (Fig. 2) shows a polytomy among *alabamensis*, *alata*, *leucophylla*, and the *rubra* complex with the exception of *rubra* as sister to *jonesii*. All *jonesii* accessions formed a monophyletic clade with 80 bootstrap support. The relationships between *alata* accessions were unresolved as well as their relationships to members of *rubra* complex.

The other subclade recovered within *Sarracenia* consists of *flava*, *minor*, *psittacina*, and the *purpurea* complex with 89 bootstrap support (Fig. 1). All members of the *purpurea* complex were monophyletic with *purpurea* ssp. *venosa* var. *montana* sister to all other *purpurea* subspecies (bootstrap value = 100; Fig. 1). In addition, infraspecific relationships within the *purpurea* clade suggest that the *venosa* subspecies are paraphyletic. Moreover, *rosea* (*purpurea* ssp. *venosa* var. *burkii*) is placed within the MP-EST *purpurea* clade (Fig. 1). All *purpurea* accession relationships were unresolved, with the exception of *purpurea* ssp. *venosa* var. *montana* and *purpurea* ssp. *purpurea* accessions (bootstrap = 50, 81, respectively; Fig. 2). Conversely, all accessions for *flava*, *psittacina*, and *minor* had 100 bootstrap support, however varieties within species were not monophyletic. The species and accession tree analyses supports the placement of the *purpurea* complex as sister to the *flava*, *psittacina*, and *minor* clade.

3.3. Concatenated nDNA and plastid tree

The RAxML concatenated nDNA accession tree has a similar overall topology to the MP-EST accession tree (Fig. 2). The tree shows strong support for *oreophila* as sister to *alata*, *leucophylla*, and the *rubra* complex (bootstrap value = 100). Additionally, it supports a polyphyletic *rubra* complex with *alata* as being a part of the *rubra* polytomy. Many relationships within the *rubra* complex are unresolved with the exception of *rubra* and *jonesii*, though two accessions of *rubra* are placed closer to the *jonesii* clade (Fig. 2). Unlike the MP-EST accession tree, concatenation supported the placement of *leucophylla* as sister to a polytomy of *alata* and the *rubra* complex, albeit with low support (bootstrap value = 58).

The concatenation tree supports the *purpurea* complex as sister to the *oreophila* clade with high support (bootstrap value = 91). Moreover, it places *psittacina* as sharing a more recent common ancestor with *minor*, although this relationship is poorly supported (bootstrap value = 55; Fig. 2). Accessions within these clades were monophyletic at 100 bootstrap support. Similar to the MP-EST accession analysis both *purpurea* ssp. *venosa* var. *montana* and *purpurea* ssp. *purpurea* are supported as monophyletic taxa, however monophyly is not supported for *purpurea* ssp. *venosa* and *rosea* (*purpurea* ssp. *venosa* var. *burkii*).

In comparison to both nuclear analyses, the plastid tree had very low resolution with polyphyletic relationships across most species and in both the *purpurea* and *rubra* complexes (Fig. 3). *Sarracenia jonesii* (*rubra* ssp. *jonesii*) and *purpurea* ssp. *venosa* var. *montana* comprised one clade within the plastid accession tree with a 100 percent bootstrap support (Fig. 3). This result may indicate introgression of the chloroplast between these two species as these species have overlapping distributions (Fig. 4a and c).

3.4. Gene tree discordance

For all nodes queried (Fig. 2, A–E), the majority of gene trees showed some degree of conflict; no node had more than 3 gene

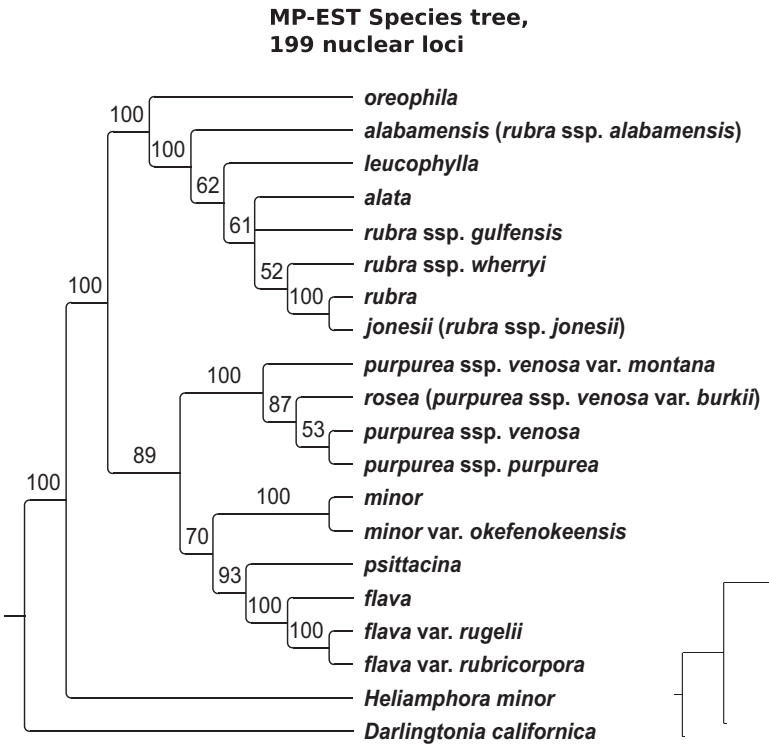

Fig. 1. *Sarracenia* MP-EST species tree based on 199 nuclear genes. Bootstrap support values are listed above respective branches. Nodes with <50 bootstrap support are collapsed. MP-EST phylogram representing branch lengths in coalescence units ($2\tau/\theta$), whereby theta (θ) is the population size estimator and tau (τ) is the parameterized branch length. Branch lengths at the terminal tips are not estimated.

trees that agreed. Conflict at nodes A–E (Fig. 2) may be the result of factors other than incomplete lineage sorting, such as gene reticulation resulting from hybridization. The hypothesis that *minor*, *psittacina* and the *purpurea* complex may be the result of hybridization between sister taxa was not supported in STEM-hy.

4. Discussion

4.1. Target enrichment with recently radiated taxa

There have been three attempts at resolving the relationships within *Sarracenia* with little resolution or agreement in species relationships (Bayer et al., 1996; Ellison et al., 2012; Neyland and Merchant, 2006). Incongruence among previous attempts that used few genes highlights the difficulties of inferring phylogenies of recently radiated groups. These groups often have not accumulated enough polymorphisms to overcome the signals left by incomplete lineage sorting. To circumvent this issue, we used 199 genes for phylogenetic analyses. This is a 28-fold increase of loci for analysis when compared to previous phylogenetic studies of this group (Ellison et al., 2012), and garnered additional parsimony-informative sites (~5000) for a more robust resolution of relationships within *Sarracenia*. This study emphasizes the utility of target enrichment for discerning relationships among recently diverged taxa.

4.2. Comparison of phylogenetic approaches

The use of next generation data requires that the methods adequately model the complexities inherent in multilocus datasets (Lemmon and Lemmon, 2013). The multispecies coalescent has been shown through simulations and theory to handle incomplete lineage sorting and produce accurate species trees when compared to concatenation methods (Degnan and Rosenberg, 2009; Edwards, 2009; Liu et al., 2008, 2009, 2010). Our analysis shows few conflicts between the MP-EST accession tree and the concatenation tree. This result may be due to robust taxon sampling in the genus (and possibly the use of multiple accessions) as previous simulations have shown that concatenation often gives incorrect topology as the number of missing taxa increases (Song et al., 2012). Additionally, 70% of the gene trees had either all taxa or just one missing taxon. This combination of conservative filtering of genes with complete taxon sampling may have contributed to the overall congruence between methods. Another possible reason for congruence between analyses could be due to lack of an "anomaly zone" (i.e., a highly probable gene topology that conflicts with the species tree) (Degnan and Rosenberg, 2006). Concatenation analyses are particularly susceptible to anomalous gene trees as this approach estimates the species tree based on the commonly observed gene tree (Degnan and Rosenberg, 2006; Liu and Edwards, 2009). Examination of gene trees in our study show high levels of gene tree heterogeneity with no dominant topology, possibly decreasing the likelihood of anomalous gene trees.

To further compare the concatenation tree and MP-EST accession tree, we examined the two major nodes that conflicted between the methods: the placement of the *purpurea* complex and *psittacina*. Alternative topologies were counted at these nodes to see if the conflicting topology was more frequent. Due to high levels of variable gene trees we were unable to discern a common, alternative topology for all nodes queried. These topological conflicts, which notably occur at short internodes, most likely resulted from retention of ancient polymorphisms resulting in high levels of incomplete lineage sorting. In addition to examining gene tree discordance, we tested hypotheses of speciation for *minor*, *psittacina* and the *purpurea* complex resulting from hybridization between sister taxa. These hypotheses were rejected from STEM-hy; however, speciation within this genus may be the result of hybridization

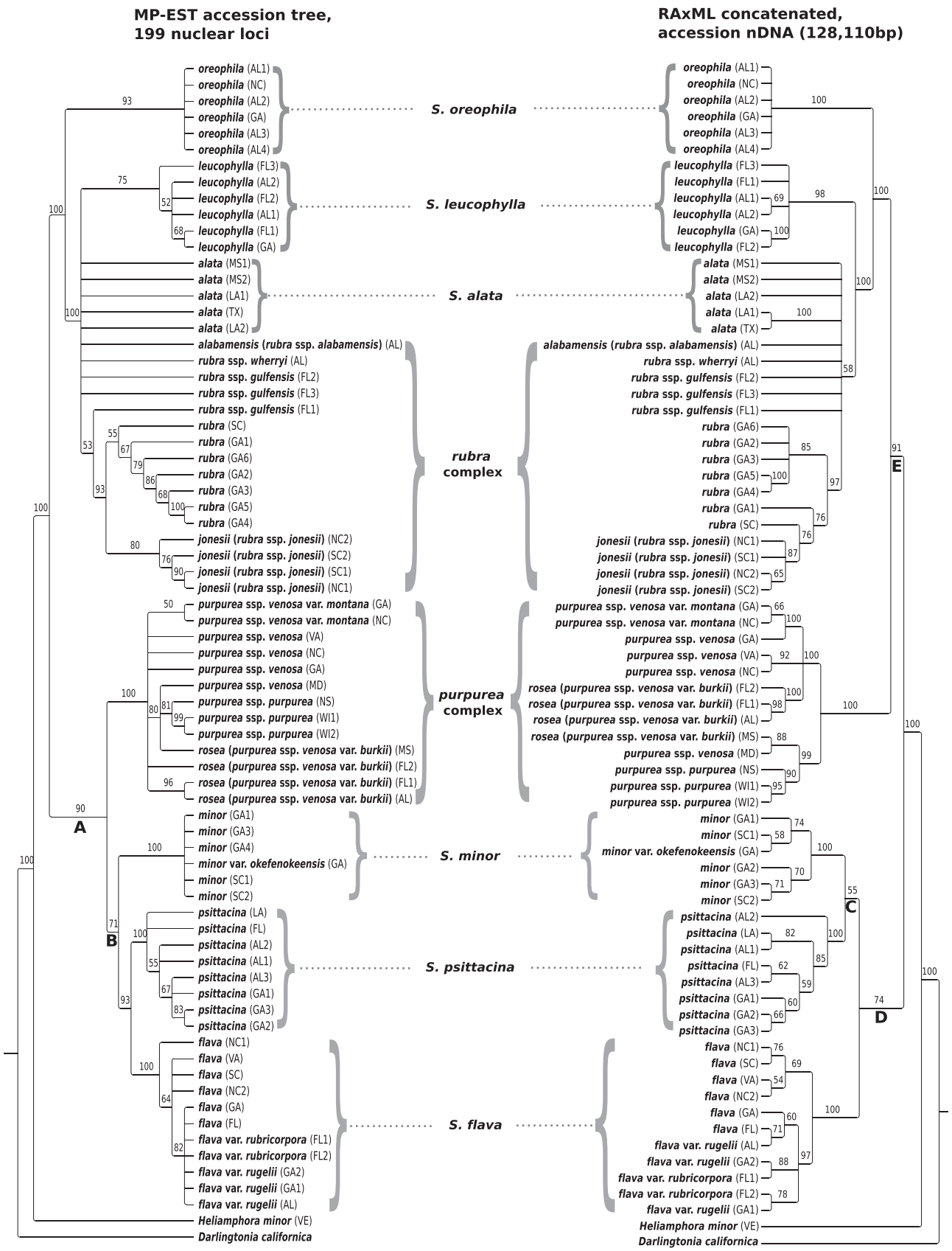

Fig. 2. *Sarracenia* MP-EST accession tree based on 199 nuclear genes and concatenated nDNA accession tree (128,110 bp) estimated from RAxML. Bootstrap support is indicated at the nodes; <50 bootstrap support nodes collapsed. Letters indicate areas of disagreement between the two analyses. Gene tree discordance was assessed at these nodes and letters correspond to results listed in the text.

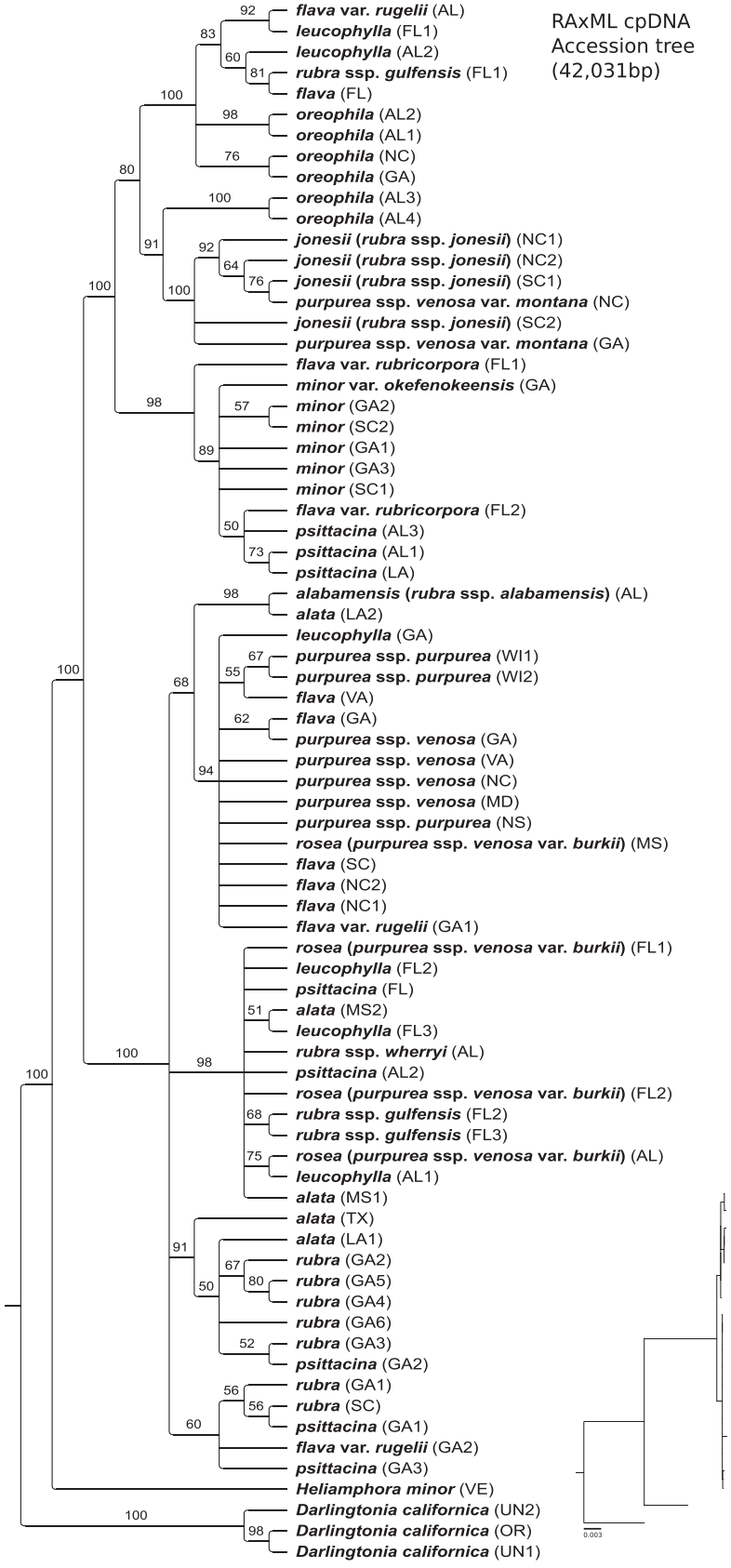

Fig. 3. *Sarracenia* plastid accession tree (42,031 bp) estimated from RAxML. Bootstrap support is indicated on the cladogram at the nodes; <50 bootstrap support nodes collapsed. Phylogram representing nucleotide substitutions per site.

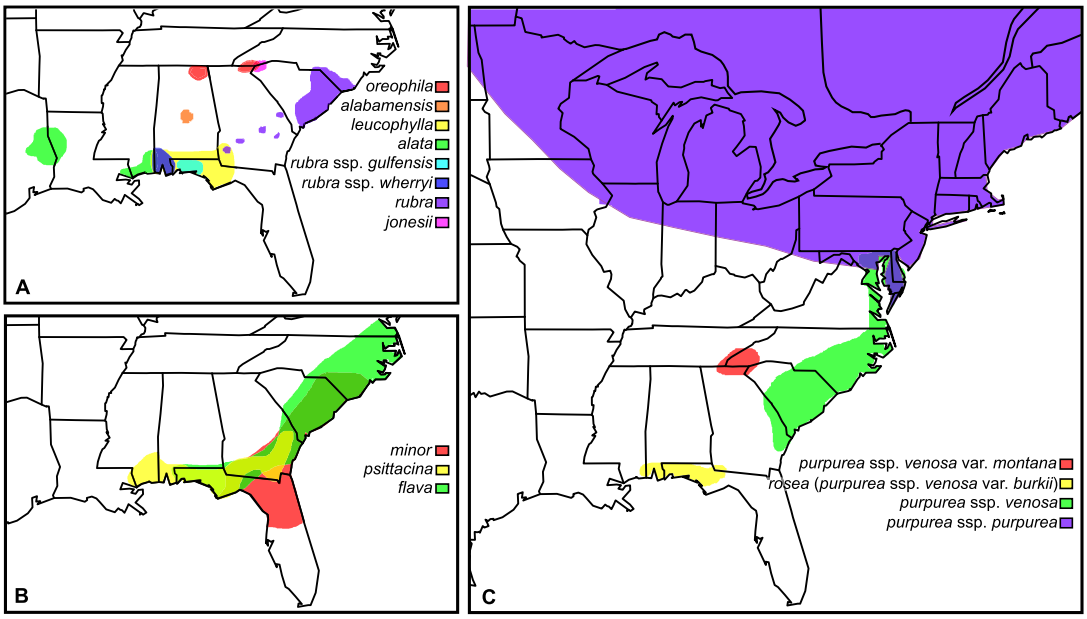

Fig. 4. Range maps for *Sarracenia* species. (A) Species and subspecies ranges from the *oreophila* clade in the MP-EST species tree. (B) Ranges for the *flava–minor–psittacina* clade; varieties are not shown. (C) Ranges of the subspecies within the *purpurea* complex.

among numerous taxa for which we were unable to test within STEM-hy. Current hybridization has been well documented for *Sarracenia* with nineteen known hybrids occurring in wild populations and these are not limited to hybridization between sister taxa (Mellichamp and Case, 2009). While care was taken to select accessions that exhibited no phenotypic signs of hybridization, ancient hybridization among species has most likely influenced speciation within this genus and contributed to incongruent topologies among gene trees.

In addition, the conflicting topologies between the nuclear phylogeny and the plastid tree further support the role of hybridization and incomplete lineage sorting within this genus. For example, the cpDNA tree supports a monophyletic clade consisting of all *jonesii* (*rubra* ssp. *jonesii*) and *purpurea* ssp. *venosa* var. *montana* accessions, a result contradicted by both coalescent and concatenated nuclear trees. This suggests possible introgression of the maternally-inherited plastome. These two species are largely isolated in mountain bogs in Georgia, North Carolina, and western South Carolina and are known to hybridize where they occur in sympatry (Fig. 4a and c; Mellichamp and Case, 2009). In contrast, Ellison et al. (2012) found signals of introgression between *purpurea* ssp. *venosa* var. *montana* and *oreophila* within the chloroplast. While we did not recover this result, many *oreophila* individuals were found to be sister to the *purpurea* ssp. *venosa* var. *montana* and *jonesii* clade possibly resulting from the proposed introgression.

Further complicating resolution using cpDNA is the lack of informative sites. In general, all accessions are not reciprocally monophyletic and there is little resolution across the tree. This is not unexpected as the chloroplast genome is more slowly evolving than the nuclear genome (Wolfe et al., 1987). Given the estimated radiation of this genus at 0.5–3 mya (Ellison et al., 2012) there has been little time for the chloroplast to accumulate enough informative polymorphisms. This lack of resolution is similar to a previous attempt to resolve these relationships using chloroplast data (Ellison et al., 2012). In addition, recently radiated taxa are often known to have conflicting cpDNA or mtDNA trees compared to nDNA trees (Sanders et al., 2013; Shaw, 2002; Zhang et al., 2014).

4.3. Evolutionary relationships within Sarracenia

4.3.1. Sarracenia oreophila clade

Similar to previous attempts, *oreophila*, *alata*, *leucophylla*, and the *rubra* complex share a close affinity with each other (Ellison et al., 2012; Neyland and Merchant, 2006), with *oreophila* as sister to the rest of the clade. In addition, there are a number of ambiguous relationships within the clade, specifically involving members of the *rubra* complex. This is not surprising, as the relationships and numbers of species/subspecies within the *rubra* complex have been highly debated due to considerable phenotypic variation maintained across disjunct populations within the range of *rubra* (Bell, 1949; Case and Case, 1976; McDaniel, 1971; Schnell, 1977, 1978). However, MP-EST and concatenation analyses did show strong support for the relationship between *rubra* and *jonesii* (*rubra* ssp. *jonesii*). These two members of the complex are closer geographically than the other subspecies within this complex (Fig. 4a).

Similar to *jonesii*, *alabamensis* (*rubra* ssp. *alabamensis*) is found in isolated populations (Fig. 4a) and is phenotypically different from other members of the *rubra* complex (Schnell, 1977). In addition, *alabamensis* grows phyllodia (i.e. non carnivorous leaves), a trait which is absent in other *rubra* complex members and *alata* and present in *oreophila* (Ainsworth and Ainsworth, 1996). This character supports the MP-EST species tree placement of *alabamensis* as more closely related to *oreophila* suggesting a polyphyletic *rubra* complex. However, in both accession trees *alabamensis* is in a polytomy with *alata*, *leucophylla*, and other members of the *rubra* complex. The inclusion of additional accessions of *alabamensis* may have supported monophyly within the species, but unfortunately were not successfully sequenced.

The other subspecies (*rubra* ssp. *gulfensis*, *rubra* ssp. *wherryi*) relationships within the *rubra* complex remain unresolved in the accession trees and have low bootstrap support in the MP-EST species tree. Both are found in the Gulf Coastal Plain where they are sympatric with numerous *Sarracenia* species (Fig. 4). Interestingly, *rubra* ssp. *gulfensis* and *alata* form a polytomy in the species tree, and both subspecies are in a polytomy with *rubra* ssp. *wherryi* in the accession tree. This result suggests a very close affinity between *alata* and the *rubra* complex, which has been suggested

by previous taxonomic descriptions of this group based on similar pitcher morphology, petal shape, size of flowers, and degree of reflexion in pitcher lid (Case and Case, 1976; McDaniel, 1966; Schnell, 1976, 1978; Sheridan, 1991). Sister to the clade containing *alata* and the *rubra* complex is *leucophylla*, which is morphologically distinct from all other members within this clade (i.e. it is the only *Sarracenia* species with white coloration in leaves). Even with its distinct morphology, there was low bootstrap support for its placement within the MP-EST species tree. Additionally, the *leucophylla* accession clade fell within a polytomy with the clade containing *alata* and the *rubra* complex in the MP-EST accession tree, but was supported as sister to those species in the concatenation tree. Locally abundant hybrids among *leucophylla*, *rubra*, and *alata* with complex backcrosses are common where species are in sympatry (McPherson and Schnell, 2011) and show signs of genetic admixture when in sympatry (Furches et al., 2013).

We did not find any phylogenetic structure for populations of *alata* sampled from either side of the Mississippi (these populations are separated by roughly 300 km with the western populations being allopatric) in the MP-EST analyses, but accessions west of the Mississippi were grouped together in the concatenation analysis. Combined with the unresolved relationships with members of the *rubra* complex, our analyses suggest that *alata* could be considered a subspecies within the *rubra* complex. The potential for *alata* to be a subspecies is in contrast to the result from the species delimitation approach conducted by Carstens and Satler (2013), suggesting that *alata* consists of two cryptic species on either side of the Mississippi River. In either case there is phenotypic variation among the *rubra* complex with geographic isolation of numerous members suggesting that this group may be in the midst of speciation.

4.3.2. Sarracenia purpurea complex

Sarracenia purpurea is the most widespread species within the genus, extending from the Gulf Coastal Plain into Newfoundland and across to British Columbia (Fig. 4c; Fernald, 1937). The infraspecific designations within the *purpurea* complex are the product of discontinuity in the distribution of this species. Both the Gulf Coastal Plain *purpurea* (*rosea*/*purpurea* ssp. *venosa* var. *burkii*) and *purpurea* ssp. *venosa* var. *montana* are geographically disjunct from the other portions of the range. The latter variety is found in isolated seep bogs in northern Georgia and the western Carolinas (Schnell and Determann, 1997). The more contiguous portion of the range consists of two named subspecies delineated near Maryland; *purpurea* ssp. *venosa* in the south and *purpurea* ssp. *purpurea* north of Maryland and across Canada. There is not a complete geographic break between the two subspecies and these species form a hybrid zone at the delineation point (Ainsworth and Ainsworth, 1996). Similar to Ellison et al. (2012), we found *rosea* (*purpurea* ssp. *venosa* var. *burkii*) as sister to *purpurea* ssp. *venosa* and *purpurea* ssp. *purpurea*. However, unlike previous results suggesting that the *purpurea* complex is sister to all other *Sarracenia* species (Ellison et al., 2012; Neyland and Merchant, 2006), our results suggest that the *purpurea* complex is sister to the clade containing *minor*, *psittacina*, and *flava* (MP-EST) or *oreophila* clade (concatenation). Examination of the accession trees did show reciprocal monophyly for *purpurea* ssp. *venosa* var. *montana* and *purpurea* ssp. *purpurea*, but not for *purpurea* ssp. *venosa* or *rosea* (*purpurea* ssp. *venosa* var. *burkii*). The latter result, more specifically *rosea* (*purpurea* ssp. *venosa* var. *burkii*), is in contrast to population level analyses of subspecies/varieties found in Godt and Hamrick (1999) and Sheridan (2010). Overall, the geographic isolation, phenotypic differences, and population genetics of the subspecies/varieties within the *purpurea* complex suggest this group may be diversifying, similar to the *rubra* complex.

4.3.3. Sarracenia minor–psittacina–flava clade

Both concatenation and MP-EST analyses strongly support the relationships between *minor*, *psittacina*, and *flava* described previously (Bayer et al., 1996; Neyland and Merchant, 2006). In addition, the MP-EST relationships of *psittacina* and *flava* as sister taxa were strongly supported by Ellison et al. (2012). Morphologically these two species are remarkably different; *psittacina* is the smallest species within the genus and has decumbent pitchers, while *flava* pitchers can reach over 80 centimeters in height. This relationship highlights the extreme range of morphological variation in trapping structures across sister species in this recently radiated group. Lastly, the relationships of *flava* varieties within the species tree are well supported, however are not reciprocally monophyletic in the accession trees. Varieties within *flava* are generally designated by anthocyanin presence. For example, *flava* var. *rubricorpora* has an almost completely red pitcher, while *flava* var. *rugelii* is characterized by the red coloration at the throat of the pitcher. Anthocyanin presence can be highly variable, responding dramatically to ecological conditions and therefore, the designation of variety based on coloration may be unwarranted.

4.4. Biogeographic hypotheses for Sarracenia

Recent estimates suggest the majority of *Sarracenia* diversification occurred less than 3 million years ago during the Pleistocene epoch (Ellison et al., 2012). The Pleistocene has been documented as having a large influence on the distribution and diversification of many southeastern United States species, most likely caused by interglacial activities (see Avise, 1996; Soltis et al., 2006). This constant climatic oscillation likely influenced *Sarracenia* speciation. Interestingly, two out of the three *Sarracenia* species found on ancient Appalachian soils (i.e. *oreophila* and *purpurea* ssp. *venosa* var. *montana*; Fig. 4a and c) are basal to other species within their respective clades. This suggests the common ancestor of *Sarracenia* may have originated from the southern Appalachian massif, which is a known area of antiquity and endemism. Resulting diversification and speciation may have occurred through two possibilities. In the first scenario ancestors to the two *Sarracenia* subclades (Fig. 1) migrated from the Appalachian massif by drainages into the Gulf and Atlantic (Godt and Hamrick, 1999) as *Sarracenia* seeds are primarily water dispersed (Schnell, 1976). Under this scenario, the ancestor of the *purpurea*/*minor* clade may have migrated along drainages leading to the Atlantic Coastal Plain while the *oreophila* ancestor may have followed the Apalachicola-Chattahoochee-Flint (ACF) River drainage to the Gulf Coastal Plain with secondary contact between the two clades occurring at the Apalachicola region. Current species ranges seem to support this hypothesis: species within the *oreophila* clade primarily occur along the Gulf Coastal Plain with the exception of more recently diverged species (*rubra* and *jonesii*) shifting along the Fall Line into Georgia and South Carolina (Fig. 4a). In addition, the MP-EST species tree indicates that *oreophila* shares a most recent common ancestor with *alabamensis* (restricted to central Alabama), possibly as a result of dispersal via drainages. Moreover, the ranges of *flava*, *minor*, and *purpurea* mostly occur along the Atlantic Coastal Plain with the exception of *psittacina* and *purpurea* ssp. *venosa* var. *burkii* (Fig. 4b and c). In the latter case, the *purpurea* complex may have covered a larger continuous range with *purpurea* ssp. *venosa* var. *montana* becoming disjunct followed by isolation of *purpurea* ssp. *venosa* var. *burkii* from *purpurea* ssp. *venosa*. In addition, the geography of the *flava* clade also suggests dispersal to the Atlantic Coastal Plain, as the basal *flava* occupies the northern limits of the its range, while the more recently diverged varieties occur at the southern limits, with *flava* var. *rubricorpora* restricted to the Florida panhandle. This biogeography hypothesis of speciation within this group is supported by numerous phylogeography

studies of this region (Soltis et al., 2006; Pauly et al., 2007). Worth noting is that the *Sarracenia* obligate symbiont *Exyra semicrocea* (pitcher plant moth) exhibits a similar genetic break in this region (Stephens et al., 2011), while *E. fax* and *E. ridingsii* are restricted to the Atlantic Coastal Plain due to their specialization on *purpurea* and *flava*, respectively.

In contrast, speciation and diversification of *Sarracenia* could have centralized around the Apalachicola region, which is a known biodiversity hotspot and has the highest overlap of *Sarracenia* species. Under this scenario, the ancestor of *Sarracenia* migrated along the Gulf Coast drainages from the Appalachian massif with successive interglacial activity fragmenting populations (J.L. Hamrick, personal communication). Subsequent speciation may have occurred through movement east and west of this area, which is supported by the current ranges of the two subclades of *Sarracenia* (Fig. 4). Future phylogeographic work on various *Sarracenia* species may help to elucidate the biogeographic history of these species and provide further insights into *Sarracenia* speciation.

4.5. Conservation implications

Sarracenia species are generally restricted to open wet pine savannas in the Coastal Plain that are maintained through frequent fires. Unfortunately, less than 3% of suitable habitat currently remains, as a cumulative result of fire restrictions, urbanization, forestry, and agriculture (Folkerts, 1982; Folkerts and Folkerts, 1993). This has lead to numerous *Sarracenia* species listed as state threatened or endangered with three (*oreophila*, *rubra* ssp. *jonesii*, and *rubra* ssp. *alabamensis*) listed as federally endangered and one (*purpurea* ssp. *venosa* var. *montana*) is a candidate for protection at the federal level (Department of the Interior, 2014). All species are listed in the Convention on International Trade in Endangered Species, Appendix II (www.cites.org) due to collection pressures. Confusion in nomenclature designations for *Sarracenia* confounds conservation practices. For example, Mellichamp and Case (2009) recognize 11 species with six subspecies, while McPherson and Schnell (2011) identify 8 species and 41 infraspecific designations. As pointed out by Ellison et al. (2014), this confusion can have serious consequences for the protection status of species within this genus. In lieu of a more well resolved phylogeny, we suggest a complete reevaluation of nomenclature across the genus. Ellison et al. (2014) recommends abandoning the use of "variety" in plant systematics so that "subspecies" is the only infraspecific designation below the rank of species. Following these recommendations, *minor* var. *okefenokeensis* and *flava* varieties will need to be reevaluated. Additionally, all taxonomic designations within the *purpurea* and *rubra* complexes should be reassessed. Given our results, *alata* should be included in the taxonomic revision of the *rubra* complex as well as a more thorough description of *jonesii* (*rubra* ssp. *jonesii*) and *alabamensis* (*rubra* ssp. *alabamensis*). Overall, a taxonomic reevaluation of this group is warranted and will hopefully lead to less confusion for management and conservation officials in charge of protecting these rare and endangered species.

5. Conclusions

We demonstrate the utility of using target enrichment and coalescent-based approaches for phylogenetic resolution of recently diverged taxa. Using target enrichment, we were able to successfully use 199 loci across 75 individuals to elucidate relationships within this genus. In addition, we were able to pull out 42 kb of cpDNA-derived sequences for a plastid tree analysis, however the plastid analysis was unable to resolve relationships. Overall, this study has resolved numerous relationships within this genus, which has important implications on the protection status of these species. Understanding the evolutionary history of *Sarracenia* lays the foundation for examining questions pertaining to evolution and speciation in this group.

Funding

This research was supported by University of Georgia Department of Plant Biology Palfrey Grant and a Grants-In-Aid of Research from the National Academy of Sciences, administered by Sigma Xi, The Scientific Research Society.

Acknowledgments

We are extremely grateful to Deanna Boensch, Stephanie Burgess, Aaron Carlson, Craig Cornu, Alex Entrup, Phil Sheridan, and Pauline Singleton for their help collecting field samples. We would also like to thank the following agencies for facilitating logistics associated with this study: Atlanta Botanical Garden, The Nature Conservancy, North Carolina Natural Heritage Program, Florida Park Service, US Forest Service, US National Park Service. This study was conducted in accordance with all United States laws and internationally accepted protocols. Permits for field collections were granted for The Nature Conservancy sites in North Carolina and Georgia and Big Thicket National Preserve Site (Permit # BITH-2012-SCI-0013). Permission to collect was approved by Kisatchie National Forest, Crotan National Forest, Francis Marion National Forest, Apalachicola National Forest, Coos Bay NERRS site, and Conecuh National Forest. Lastly, we would like to thank Jim Hamrick and the Leebens-Mack and Donovan labs for helpful comments that greatly improved this manuscript and the Burke and Zhang lab for use of their lab equipment.

Appendix A. Supplementary material

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.ympev.2015.01.015.

References

Ainsworth, J., Ainsworth, J., 1996. *Sarracenia*: North American Pitcher Plants, The National Council for the Conservation of Plants and Gardens, UK.
Albert, V.A., Williams, S.E., Chase, M.W., 1992. Carnivorous plants: phylogeny and structural evolution. Science 257, 1491–1495.
Altschul, S.F., Madden, T.L., Schäffer, A.A., Zhang, J., Zhang, Z., Miller, W., Lipman, D.J., 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. 25, 3389–3402.
Avise, J.C., 1996. Towards a regional conservation genetic perspective: phylogeography of faunas in the southeastern United States. In: Avise, J.C., Hamrick, J.L. (Eds.), Conservation Genetics; Case Histories from Nature. Chapman & Hall, New York, pp. 431–470.
Bayer, R.J., Hufford, L., Soltis, D.E., 1996. Phylogenetic relationships in Sarraceniaceae based on rbcL and ITS sequences. Syst. Bot. 21, 121–134.
Beerli, P., 2009. How to use migrate or why are markov chain Monte Carlo programs difficult to use? In: Bertorelle, G., Bruford, M.W., Hauffe, H.C., Rizzoli, A., Vernesi, C. (Eds.), Population Genetics for Animal Conservation. Conservation Biology, vol. 17. Cambridge Press, Cambridge, UK, pp. 42–79.
Beerli, P., Felsenstein, J., 2001. Maximum likelihood estimation of a migration matrix and effective population sizes in n subpopulations by using a coalescent approach. Proc. Natl. Acad. Sci. USA 98, 4563–4568.
Bell, C.R., 1949. A cytotaxonomic study of the Sarraceniaceae of North America. J. Elisha Mitchell Sci. Soc. 65, 137–166.
Carstens, B.C., Satler, J.D., 2013. The carnivorous plant described as *Sarracenia alata* contains two cryptic species. Biol. J. Linn. Soc. 109, 737–746.
Case, F.W., Case, R.B., 1976. The *Sarracenia rubra* complex. Rhodora 78, 270–325.
Castresana, J., 2000. Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. Mol. Biol. Evol. 17, 540–552.
Chikhi, R., Medvedev, P., 2013. Informed and automated k-Mer size selection for genome assembly. Bioinformatics. http://dx.doi.org/10.1093/bioinformatics/btt310.
Darriba, D., Taboada, G.L., Doallo, R., Posada, D., 2012. JModelTest 2: more models, new heuristics and parallel computing. Nat. Methods 9, 772.

Darwin, C., 1875. Insectivorous Plants. John Murray, London.
Degnan, J.H., Rosenberg, N.A., 2006. Discordance of species trees with their most likely gene trees. PLoS Genet. http://dx.doi.org/10.1371/journal.pgen.0020068.
Degnan, J.H., Rosenberg, N.A., 2009. Gene tree discordance, phylogenetic inference and the multispecies coalescent. Trends Ecol. Evol. 24, 332–340.
Department of Interior, 2014. U.S. Fish & Wildlife Service: Endangered Species Program. <http://www.fws.gov/endangered/> (accessed 14.09.14).
Edwards, S.V., 2009. Is a new and general theory of molecular systematics emerging? Evolution. http://dx.doi.org/10.1111/j.1558-5646.2008.00549.x.
Ellison, A.M., Gotelli, N.J., 2009. Energetics and the evolution of carnivorous plants-Darwin's "most wonderful plants in the world". J. Exp. Bot. 60, 19–42.
Ellison, A.M., Butler, E.D., Hicks, E.J., Naczi, R.F.C., Calie, P.J., Bell, C.D., Davis, C.C., 2012. Phylogeny and biogeography of the carnivorous plant family Sarraceniaceae. PLoS One. http://dx.doi.org/10.1371/journal.pone.0039291.
Ellison, A.M., Davis, C.C., Calie, P.J., Naczi, R.F.C., 2014. Pitcher plants (*Sarracenia*) provide a 21st-century perspective on infraspecific ranks and interspecific hybrids: a modest proposal for appropriate recognition and usage. Syst. Bot. 39, 939–949.
Faircloth, B.C., Glenn, T.C., 2012. Not all sequence tags are created equal: designing and validating sequence identification tags robust to indels. PLoS One 7, e42543.
Faircloth, B.C., Glenn, T.C., McCormack, J.E., Harvey, M.G., Sorsenson, L., Alfaro, M., White, N., 2012a. Target Enrichment of Illumina Libraries. <http://ultraconserved.org> (accessed 03.10.14).
Faircloth, B.C., McCormack, J.E., Crawford, N.G., Harvey, M.G., Brumfield, T.C., Glenn, T.C., 2012b. Ultraconserved elements anchor thousands of genetic markers for target enrichment spanning multiple evolutionary timescales. Syst. Biol. 61, 717–726.
Fernald, M.L., 1937. Plants of the Inner Coastal Plain of Virginia. Rhodora 39, 386–387.
Folkerts, G.W., 1982. The Gulf Coast pitcher plant bogs. Am. Sci. 70, 260–267.
Folkerts, G.W., Folkerts, D.R., 1993. Southeastern pitcher plant bogs: a natural history sketch. Tipularia 8, 10–16.
Furches, M.S., Small, R.L., Furches, A., 2013. Hybridization leads to interspecific gene flow in *Sarracenia* (Sarraceniaceae). Am. J. Bot. 100, 2085–2091.
Glenn, T.C., Faircloth, B.C., Nilsen, R., Kieran, T.J., Finger Jr., J.W., Pierson, T.W., García-De-Leon, F.J., del Rio Portilla, M.A., Reed, K., Anderson, J.L., Meece, J.K., Alabady, M., Belanger, M., submitted for publication. Adapterama I: universal stubs and primers for thousands of dual-indexed Illumina Nextera and TruSeqHT compatible libraries (iNext & iTru). Mol. Ecol. Resour.
Godt, M.J.W., Hamrick, J.L., 1999. Genetic divergence among infraspecific taxa of *Sarracenia purpurea*. Syst. Bot. 23, 427–438.
Grabherr, M.G. et al., 2011. Full-length transcriptome assembly from RNA-Seq data without a reference genome. Nat. Biotechnol. 29, 644–652.
Heled, J., Drummond, A.J., 2010. Bayesian inference of species tree from multilocus data. Mol. Biol. Evol. 27, 570–580.
Hennig, W., 1966. Phylogenetic Systematics. Illinois Press, Urbana-Champaign (IL).
Heyduk, K., Trapnell, D.W., Barrett, C.F., Leebens-Mack, J., in press. Estimating relationships within *Sabal* (Arecaceae) through multilocus analyses of sequence capture data. Biol. J. Linn. Soc.
Huang, X., Madan, A., 1999. CAP3: a DNA sequence assembly program. Genome Res. 9, 868–877.
Jansen, R.K., Kaittanis, C., Saski, C., Lee, S.B., Tomkins, J., Alverson, A.J., Daniell, H., 2006. Phylogenetic analyses of *Vitis* (Vitaceae) based on complete chloroplast genome sequences: effects of taxon sampling and phylogenetic methods on resolving relationships among rosids. BMC Evol. Biol. http://dx.doi.org/10.1186/1471-2148-6-32.
Jobson, R.W., Playford, J., Cameron, K.M., Albert, V.A., 2003. Molecular phylogenetics of Lentibulariaceae inferred from plastid *rps*16 intron and *trn*L-F DNA sequences: implications for character evolution and biogeography. Syst. Bot. 28, 157–171.
Juniper, B.E., Robins, R.J., Joel, D.M., 1989. The Carnivorous Plants. Academic Press, London.
Katoh, K., Standley, D.M., 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol. Biol. Evol. 30, 772–780.
Knowles, L.L., 2009. Estimating species trees: methods of phylogenetic analysis when there is incongruence across genes. Syst. Biol. 58, 463–467.
Kubatko, L.S., 2009. Identifying hybridization events in the presence of coalescence via model selection. Syst. Biol. 58, 478–488.
Kubatko, L.S., Degnan, J.H., 2007. Inconsistency of phylogenetic estimates from concatenated data under coalescence. Syst. Biol. 56, 17–24.
Langmead, B., Salzberg, S.L., 2012. Fast gapped-read alignment with Bowtie 2. Nat. Methods 9, 357–359.
Langmead, B., Trapnell, C., Pop, M., Salzberg, S.L., 2009. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biol. http://dx.doi.org/10.1186/gb-2009-10-3-r25.
Lemmon, E.M., Lemmon, A.R., 2013. High-throughput genomic data in systematics and phylogenetics. Annu. Rev. Evol. Syst. 44, 99–121.
Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R., 1000 Genome Project Data Processing Subgroup, 2009. The sequence alignment/map (SAM) format and SAMtools. Bioinformatics 25, 2078–2079.
Liu, L., Edwards, S.V., 2009. Phylogenetic analysis in the anomaly zone. Syst. Biol. http://dx.doi.org/10.1093/sysbio/syp034.
Liu, L., Pearl, D.K., Brumfield, R.T., Edwards, S.V., 2008. Estimating species trees using multiple-allele DNA sequence data. Evolution 62, 2080–2091.
Liu, L., Yu, L., Pearl, D.K., Edwards, S.V., 2009. Estimating species phylogenies using coalescence times among sequences. Syst. Biol. 58, 468–477.
Liu, L., Yu, L., Edwards, S.V., 2010. A maximum pseudo-likelihood approach for estimating species trees under the coalescent model. BMC Evol. Biol. 10, 302.
Lodhi, M.A., Guang-Ning, Y., Weeden, N.F., Reisch, B.I., 1994. A simple and efficient method for DNA extraction from grapevine cultivars, *Vitis* species and *Ampelopsis*. Plant Mol. Biol. Rep. 12, 6–13.
Löytynoja, A., Goldman, N., 2008. Phylogeny-aware gap placement prevents errors in sequence alignment and evolutionary analysis. Science. http://dx.doi.org/10.1126/science.1158395.
Maddison, W.P., 1997. Gene trees in species trees. Syst. Biol. 46, 523–536.
Mamanova, L., Coffey, A.J., Scott, C.E., Kozarewa, I., Turner, E.H., Kumar, A., Howard, E., Shendure, J., Turner, D.J., 2010. Target-enrichment strategies for next-generation sequencing. Nat. Methods 7, 111–118.
McCormack, J.E., Faircloth, B.C., Crawford, N.G., Gowaty, P.A., Brumfield, R.T., Glenn, T.C., 2012. Ultraconserved elements are novel phylogenomic markers that resolve placental mammal phylogeny when combined with species tree analyses. Genome Res. 22, 746–754.
McDaniel, S., 1966. Distribution of *S. purpurea* in Canada. In: A Taxonomic Revision of *Sarracenia* (Sarraceniaceae), Ph.D. Dissertation, The Florida State University, Tallahassee, Florida, USA.
McDaniel, S.T., 1971. The genus *Sarracenia* (Sarraceniaceae). Bull. Tall Timbers Res. Station 9, 1–36.
McPherson, S., 2007. Pitcher Plants of the Americas. The McDonald & Woodward Publishing Company, Blacksburg, VA, USA.
McPherson, S., Schnell, D., 2011. Sarraceniaceae of North America. Redfern Natural History Productions Ltd., Poole, Dorset, England.
Meimberg, H., Wistuba, A., Dittrich, P., Heubl, G.R., 2001. Molecular phylogeny of Nepenthaceae based on cladistic analysis of plastid trnK intron sequences. Plant Biol. 3, 164–175.
Mellichamp, T.L., Case, F.W., 2009. Sarracenia. In: Flora of North America Editorial Committee (Eds.), Flora of North America North of Mexico, vol. 8. Oxford Univ. Press, Oxford, U.K. and New York, NY, pp. 348–363.
Müller, K., Borsch, T., 2005. Phylogenetics of *Utricularia* (Lentibulariaceae) and molecular evolution of the *trn*K intron in a lineage with high substitutional rates. Plant Syst. Evol. 250, 39–67.
Neyland, R., Merchant, M., 2006. Systematic relationships of Sarraceniaceae inferred from nuclear ribosomal DNA sequences. Madroño 53, 223–232.
Pamilo, P., Nei, M., 1988. Relationships between gene trees and species trees. Mol. Biol. Evol. 5, 568–583.
Pauly, G.B., Piskurek, O., Shafer, H.B., 2007. Phylogeographic concordance in the southeastern United States: the flatwoods salamander, *Ambystoma cingulatum*, as a test case. Mol. Ecol. 16, 415–429.
Peterson, D.G., Tomkins, J.P., Frisch, D.A., Wing, R.A., Paterson, A.H., 2000. Construction of plant bacterial artificial chromosom (BAC) libraries: an illustrated guide. J. Agric. Genom. 5, <www.ncgr.org/research/jag>.
Ratan, A., 2009. Assembly Algorithms for Next Generation Sequence Data. Ph.D. Dissertation, Pennsylvania State University, State College, Pennsylvania, USA.
Rivadavia, F., Kondo, K., Kato, M., Hasebe, M., 2003. Phylogeny of the sundews, *Drosera* (Droseraceae), based on chloroplast and nuclear 18S ribosomal DNA sequences. Am. J. Bot. 90, 123–130.
Sanders, K.L., Lee, M.S.Y., Mumpuni, Bertozzi, T., Rasmussen, A.R., 2013. Multilocus phylogeny and recent rapid radiation of the viviparous sea snakes (Elapidae: Hydrophiinae). Mol. Phylogenet. Evol. 66, 575–591.
Schnell, D.E., 1976. Carnivorous Plants of the United States and Canada. John F. Blair, Publisher, Winston-Salem, NC.
Schnell, D.E., 1977. Infraspecific variation in *Sarracenia rubra* Walt.: some observations. Castanea 42, 149–170.
Schnell, D.E., 1978. Systematic flower studies of *Sarracenia* L. Castanea 43, 211–220.
Schnell, D.E., Determann, R.O., 1997. *Sarracenia purpurea* L. subspecies *venosa* (Raf.) Wherry variety *montana* Schnell & Determann (Sarraceniaceae): a new variety. Castanea 62, 60–62.
Shaw, K.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. Proc. Natl. Acad. Sci. 99, 16122–16127.
Sheridan, P.M., 1991. What is the identity of the west Gulf coast pitcher plant, *Sarracenia alata* Wood? Carniv. Plant Newslett. 20, 102–110.
Sheridan, P.M., 2010. Ecological and Genetic Status of the Purple Pitcher Plant, *Sarracenia purpurea* L., in Maryland and Virginia. Ph.D. Dissertation, Old Dominion University, Virginia.
Soltis, P.S., Soltis, D.E., 2004. The origin and diversification of angiosperms. Am. J. Bot. 91, 1614–1626.
Soltis, D.E., Morris, A.B., McLachlan, J.S., Manos, P.S., Soltis, P.S., 2006. Comparative phylogeography of unglaciated eastern North America. Mol. Ecol. 14, 671–688.
Song, S., Liu, L., Edwards, S.V., Wu, S., 2012. Resolving conflict in eutherian mammal phylogeny using phylogenomics and the multispecies coalescent model. Proc. Natl. Acad. Sci. http://dx.doi.org/10.1073/pnas.1211733109.
Srivastava, A., Rogers, W.L., Breton, C.M., Cai, L., Malmberg, R.L., 2011. Transcriptome analysis of *Sarracenia*, an insectivorous plant. DNA Res. http://dx.doi.org/10.1093/dnares/dsr014.
Stamatakis, A., 2006. RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. Bioinformatics. http://dx.doi.org/10.1093/bioinformatics/btl446.

Stephens, J.D., Santos, S.R., Folkerts, D.R., 2011. Genetic differentiation, structure, and a transition zone among populations of the pitcher plant moth *Exyra semicrocea*: implications for conservation. PLoS One 6, e22658.

Wolfe, K.H., Li, W.H., Sharp, P.M., 1987. Rates of nucleotide substitution vary greatly among plant mitochondrial, chloroplast, and nuclear DNAs. Proc. Natl. Acad. Sci. 84, 9054–9058.

Yang, Z., 2002. Likelihood and Bayes estimation of ancestral population sizes in hominoids using data from multiple loci. Genetics 162, 1822–1823.

Zerbino, D.R., 2010. Using the Columbus extension to Velvet. <http://bioweb2.pasteur.fr/docs/modules/velvet/1.1.02/Columbus_manual.pdf> (accessed 12.06.14).

Zerbino, D.R., Birney, E., 2008. Velvet: algorithms for de novo short read assembly using de Bruijn graphs. Genome Res. 18, 821.

Zhang, J.Q., Meng, S.Y., Allen, G.A., Wen, J., Rao, G.Y., 2014. Rapid radiation and dispersal out of the Qinghai-Tibetan Plateau of an alpine plant lineage Rhodiola (Crassulaceae). Mol. Phylogenet. Evol. 77, 147–158.