

Systematic study

Stephens et al. (2015): Resolving phylogenetic relationships of the recently radiated carnivorous plant genus *Sarracenia* using target enrichment. *Molecular Phylogenetics and Evolution* 85, 76–87.



Reasons for the study

- elucidate evolutionary relationships of *Sarracenia* (New World pitcher plants)
- *Sarracenia* – recently diverged group
- previous attempts at construction phylogeny failed
- assess the utility of target enrichment of nuclear genes for a recently radiated, non-model genus

Study genus – *Sarracenia*

- 8-11 species
- 41 subspecies, varieties and forms
- wet savannas and fens of N America
- endangered – due to destroyed habitats (less than 3% remained)
- recently radiated (0.5-3 million years ago)
- *Heliamphora* – sister genus
- *Darlingtonia* – basal lineage in Sarraceniaceae

Methods – probes, libraries

- 71 accessions
- probe design – comparing two *Sarracenia* transcriptomes
 - ca. 1, 000 contigs
 - within-species BLAST – search for potential paralogues
 - reciprocal best BLAST (blastn) – determine orthologues between species
 - resulted in 646 genes for target enrichment
- DNA sheared to fragments (180-500 bp)
- Illumina TruSeqHT compatible libraries
- target enrichment using MYbaits
- sequenced on Illumina HiSeq PE100

Methods – read analysis

- quality check – FastQC
- quality trimming – FastX Toolkit
- Illumina adapter removal – FAR
- assembly – two approaches
 - de novo – Trinity
 - reference-based – Columbus extension module in VELVET
- identical contigs merged (CAP3) and matched against gene targets (BLAST) – 1:1 hits retained
- contigs merged, aligned (Prank), poorly aligned regions removed (Gblocks)
- selection of loci with
 - less than 0.35 average pairwise distance
 - at least 50% of the accessions present
 - no more than 45% of missing data
 - at least one outgroup

Methods – tree analyses

- gene trees – ML in RAxML, 500 bootstrap replicates
- MP-EST species tree
- STEM-hy – test hypotheses of hybrid species (on taxa showed incongruence between MP-EST and concatenated trees)
- concatenated analysis – RAxML, 1000 bootstrap replicates

- plastome tree
 - reads mapped to *Vitis* plastome (Bowtie2), extracted (SAMtools), assembled with YASRA (reference-based assembler)
 - contigs concatenated
 - aligned using MAFFT, poorly aligned regions removed (Gblocks)
 - ML tree (RAxML, 1000 bootstrap replicates)

Results

- each accession – roughly 3.5 million reads
- 7,124 Trinity and 67,894 Velvet contigs – merged into 5,608 contigs per assembly
- 546 contigs matched 646 genes
- 11× coverage
- 199 genes – average length 642 bp, 128,110 bp in total
 - 8.7% variable characters
 - 4% parsimony informative characters
- plastome – 42,031 bp
 - 1.9% variable characters
 - 0.5% parsimony informative characters

Results

- MP-EST – supports monophyly of *Sarracenia*, many relationships within the genus
- *rubra* complex – polyphyletic
- *flava*, *minor*, *psittacina*, *purpurea* monophyletic
- concatenated tree – similar topology, but
 - *purpurea* sister to *oreophila*
 - *flava* sister to *minor* and *psittacina*
- plastid tree – very low resolution
 - *jonesii* and *purpurea* subsp. *venosa* var. *montana* – one clade, indication for introgression
- *minor*, *psittacina* and *purpurea* – hybridization between sister taxa not supported by STEM-hy

Discussion

- difficulties of inferring phylogenies of recently radiated groups
- MP-EST and concatenation – only few conflicts
 - robust taxon sampling
 - lack of “anomaly zone” (i.e., highly probable gene topology that conflicts with the species tree)
 - two major conflicts – placement of the *purpurea* complex and *psittacina* – no dominant topology among gene trees
- conflict between nuclear and plastid tree – supports the role of hybridization and ILS – common in recently radiated groups

Discussion - biogeography

- diversification of *Sarracenia* – less than 3 mya
- two species growing on ancient Appalachian soils (*oreophila* and *purpurea* ssp. *venosa* var. *montana*) are basal to other species – common ancestor in southern Appalachian massif
- first scenario
 - ancestors of two subclades migrated into the Gulf and Atlantic
 - *purpurea/minor* clade – to Atlantic Coastal Plain
 - *oreophila* ancestor – ACF river drainage to the Gulf Coastal Plain
- second scenario
 - diversification around the Appalachian region – here is the highest overlap of species
 - migration along the Gulf Coast with fragmentation caused by glaciation

Conservation implications

- less than 3% of suitable habitat currently remains
- numerous species threatened, 3 federally listed as endangered
- confusion in nomenclature – serious consequences for the protection status of species
- complete reevaluation of nomenclature suggested
- taxonomic reevaluation – less confusion for management

Conclusions

- utility of target enrichment for phylogenetic resolution of recently diverged taxa
- 199 loci across 75 individuals
- 42 kb of cpDNA-derived sequences – unable to resolve relationships