The American Dahlia Society

ADS Genome Project Update
4/16/2022

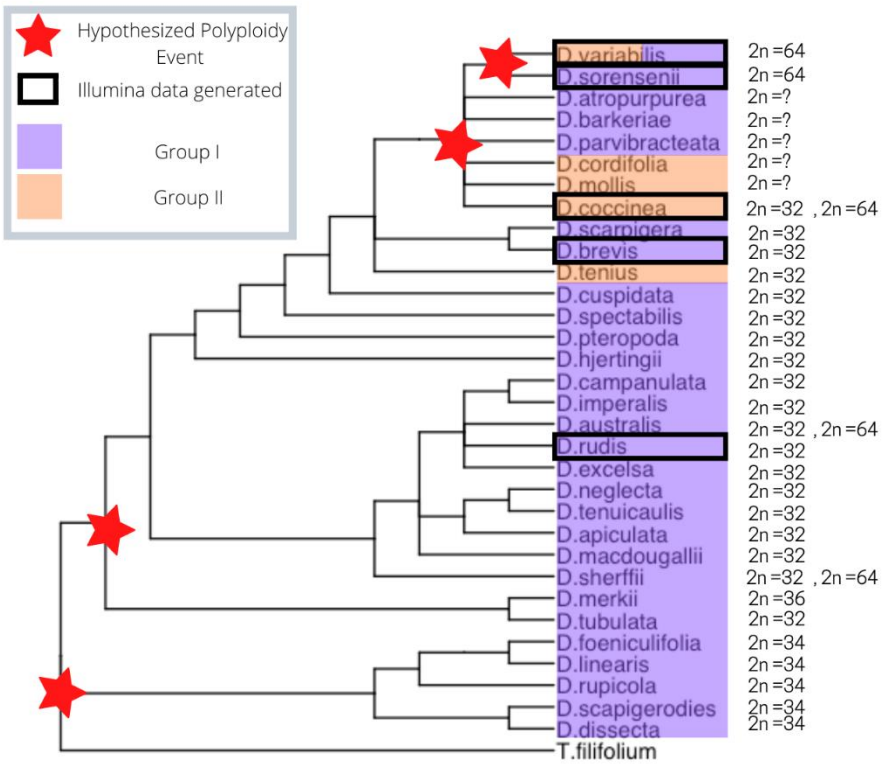
Resolving the phylogeny of the *Dahlia* genus and identifying the putative parents of the cultivated octoploids

In collaboration with the Freeman Herbarium at Auburn University, Zach is currently sourcing herbarium material from 48 unique *Dahlia* species and subspecies, in replicate, totaling 105 sample sheets from herbaria across the United States and Mexico. All tissues will be extracted using a standard CTAB method, checked for DNA quantity with the Qubit® 3.0 Fluorometer (Life Technologies, Carlsbad, CA, USA) and purity using the Denovix DS-11 FX (Denovix, Wilmington, Delaware, USA) and stored in a 4˚C refrigerator before library preparation starts.

The extracted libraries will be enriched with the Angiosperms353 baits. The libraries will be sequenced on a Illumina MiSeq. The sequence reads will be assembled using the Hybpiper pipeline. The fasta files created from Hybpiper will be aligned using MUSCLE and used as input for maximum likelihood tree estimation using RAxML.

The resulting gene trees will be used in ASTRAL-Pro to construct multi-gene coalescent species trees using single- and multi-copy gene families. This tree will be compared with the phylogenies created by Gatt and Saar[8]. With the addition of *Dahlia mixtecana, Dahlia tamaulipana, and Dahlia wixarika* to the genus and through the advancement in sequencing technology we expect this phylogeny to be a more accurate representation of the species relationships in te *Dahlia* genus.

Continue…



Using the updated *Dahlia* phylogeny, along with a k-mer based comparison pipeline, Zach will identify the most likely parents of *D. variabilis* and collect living samples from those species. Once grown, young leaf tissue will be collected and DNA will be extracted using a standard CTAB method before being Illumina whole genome shotgun sequenced. The data from the species will be added to the species I have already sequenced to >40X coverage: *D. brevis, D. coccinea, D. sorenesenii,* and *D. rudis* (Figure 1). K-mer lists for each species will be generated using jellyfish and mapped to each other using BWA to identify k-mers specific to each species. The species-specific k-mers will then be mapped to a cultivated dahlia (*D. variabilis*) to identify the putative ancestors to *D. variabilis*.

By generating an updated phylogeny of the *Dahlia* genus Zach will be able to utilize gene trees to map the polyploidization events using PUG. Using BiSSE he will test for increased rate of diversification following the polyploidization events in the genus.

Zach has already generated Illumina sequencing data for 5 Dahlia species and is developing an analysis pipeline to test if rudis, brevis, coccinea, and sorensenii might be putative parents to the D. variabilis species complex. Zach recently wrote a grant proposal to the Botanical Society of America to fund some of this work, to complement the ADS funds we are grateful to receive.