

### Time line of the Dahlia genome project

When	Event	Comments
2016 April	Professor Walbot talk to the Santa Cruz Dahlia Society	Challenge issued to raise funds to sequence a dahlia genome and examine diversity in the genus
2016 April - September	Fundraising in ADS organized by Kristine Albrecht	Sufficient funds raised to proceed with planning
2017	NSF grant submitted to support a high quality dahlia genome project	And sample sequencing of modern cultivars and a number of species
2017 May	Professor Walbot receives USDA import permits for dahlia seed from Mexico	Pre-requisite to importing diverse species seed for genome work
2017 September	National ADS presentation in Chicago by Professor Walbot	Goals and timetables presented for the dahlia genome project, contingent on NSF funding
2017 October	Professor Walbot and Tim Culbertson collect dahlia species in Jalisco	Greatly assisted by Professor Eduardo Ruiz Sanchez and staff of the Univ. Guadalajara herbarium
2017 October	Stanford Professors Rodolfo Dirzo and Walbot collect dahlia species in Morelos Mexico	Greatly assisted by staff of the Univ. Morelos in Cuernavaca, Mexico. The 2 collecting trips yield 6 dahlia species, including <i>D. pugana</i> never grown before in the US, with diversity collections of the 2 proposed parents: <i>D. coccinea</i> and <i>D. sorensenii</i> .
2017 November	Professor Dirzo collects dahlia seeds in Querétaro	So late in the season, seed are collected “blind” with later identification after growout in summer 2018. Ultimately 5 additional species are identified, plus additional examples (and new flower colors) of <i>D. coccinea</i> and <i>D. brevis</i> .
2017 December	Decision to sequence leaf RNA, given the large number of species’ seed collected	RNA “transcriptomes” focus on the expressed genes and are the most efficient and cheapest way to get substantial data for the family tree
2018 January	Species seed started at Stanford and by Kristine Albrecht for summer gardens	
2018 Feb	We learn the NSF proposal is not approved for funding	Reviewers were extremely impressed and enthusiastic about our proposed outreach activity in classrooms, involving many ADS members

2018 April	Professor Walbot presents dahlia genome progress report at an ADS meeting	Examples of some species seedlings distributed and volunteers found to grow the “mystery” seed from Querétaro
2018 May - August	Flash frozen young leaves collected from EDNA C and Comet and from 15 dahlia species as soon as they can be identified (from vegetation and flowers)	Two trips to Santa Cruz and a trip by Professor Walbot, Tim Culbertson, and Kristine Albrecht to Seattle area to collect from Wayne Lobaugh, Martin Kral, and Brad Freeman gardens, yielding 4 more dahlia species including <i>D. rupicola</i> , an “outgroup” species distant from likely parents of modern dahlia.
2018 September	Professor Walbot and Kristine Albrecht collect 8 modern dahlia samples in Santa Cruz CA	Jomanda, Mexico, Apopa Gold, Eden Benary, Rhonda, Thomas Edison, Emory Paul and Pam Howden (Edna C, Comet and Rio Riata had already been collected)
2018 August - September	Professor Walbot processes the frozen dahlia leaves for shipment for RNA preparation at Novogene	Back up samples sent to Dr. Alex Harkness for RNA prep; Dr. Harkness will be analyzing and organizing the data
2018 September	Frozen, powdered leaves sent to Novogene for RNA sequencing	
Expected 2018 Oct-Nov	Transcriptome data back on 15 species and 11 modern dahlias	Dr. Harkness will analyze and with Professor Walbot answer the questions in our goals (see below)
Expected 2018 Dec	Pick one species for a high quality genome project	
Expected Mar 2019	Genome analysis	
Fall 2019, if needed	Collecting trip to Mexico to acquire additional species	Required if the existing 15 species do not include the parent(s) of modern dahlia. It is possible that different parents were crossed to initiate some dahlia types.

## Goals and Applications

1. Organize a phylogeny of dahlia species based on gene sequences and of a number of modern cultivars to determine the parent or parents of modern horticultural dahlias. A phylogeny is a family tree: because dahlia breeding dates back just a few hundred years, one or two species of dahlia are expected to be a close match to modern dahlias.
2. Generate a high quality assembled genome of one Dahlia species, the top choice is one of the parents of modern dahlias. The genome will stimulate scientific research on dahlias and can aid the search for virus sequences integrated into the dahlia genome.

3. From several sequenced examples of some species dahlias, determine how “diverse” several species are. For those interesting in crossing species dahlias with modern dahlias, knowledge of where “genetic diversity” occurs can guide which types to cross.
4. From sequenced examples of 11 modern dahlias determine how “diverse” the germplasm is in breeding types (formal decorative, informal decorating, ball, anemone, single). Do some floral forms have more diverse alleles and hence are more promising in developing additional dahlia cultivars?

Fundraising total as of September 2018 \$47,000 plus \$5000 pledge from Dr. Pappu

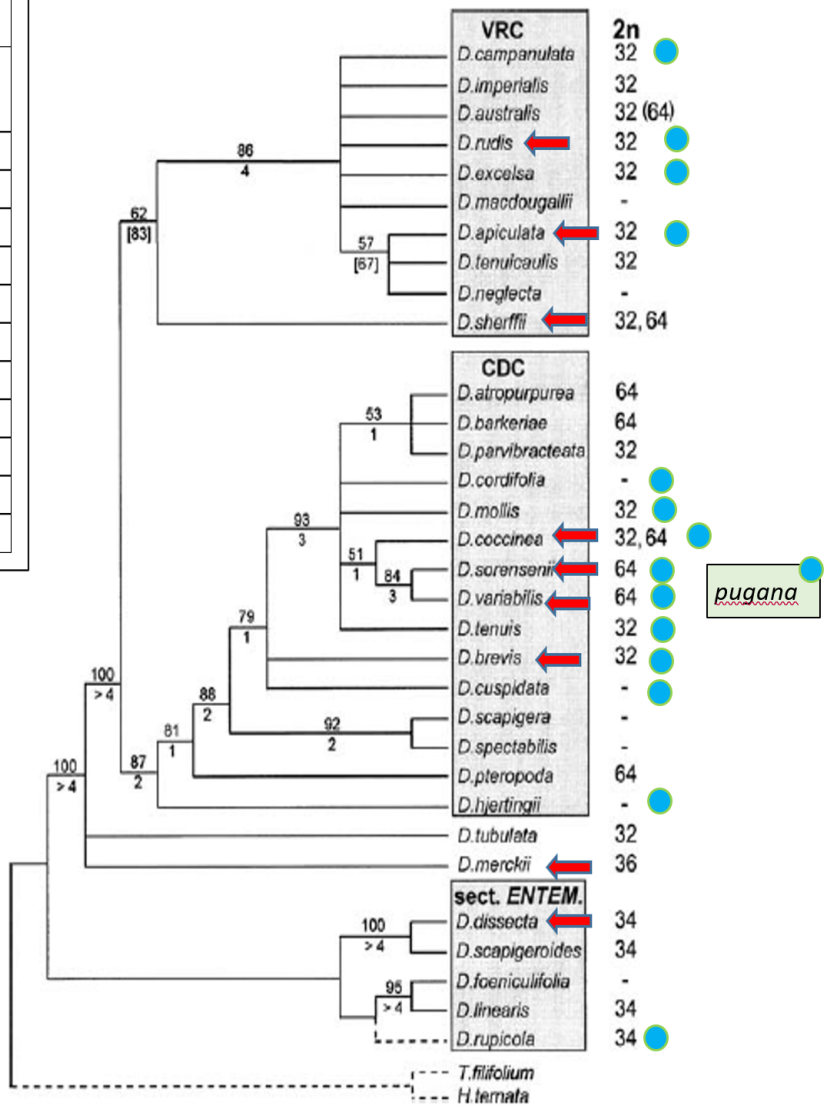
#### Expenses

Collecting trips, about \$4500 RNA transcriptomes for 30 total samples \$9,800

Expected high quality genome cost \$30,000 Cost of a second genome \$10,000

In the future more modern dahlias could be evaluated for about \$500 each, provided tubers are sent to California to permit collecting young leaves into liquid nitrogen (flash frozen sample collecting).

S,  
l.  
ib



On this chart the red arrows are what we promised sequencing in the NSF grant proposal, and the blue dots indicate the species we have collected in 2017-2018 that will have the leaf transcriptome analyzed and from which we will sequence one Dahlia species genome.

