

The American Dahlia Society



ADS Genome Project Update February 2022

Alex Harkess

Many exciting updates have happened since the last board report. First, Zach Meharg transitioned from a Masters student to a PhD student. In my opinion, he was hungry for the challenge, and the experiments he was proposing were outside the scope of a 2 year Masters project. Consequently, we moved him to a PhD program in the same department, and his graduation timeline is now 5-6 years from now. He just completed his first semester.

Zach diligently collaborated with Kristine Albrecht and Marcie Holt to grow and sample tissue from *D. brevis*, *D. rudis*, *D. coccinea*, *D. sorensenii*, and cultivated Thomas Edison. Zach isolated high molecular weight and pure DNA, and generated Illumina PCR-free shotgun sequencing libraries. These 5 libraries were sequenced on a full lane of Illumina PE150 S4 NovaSeq6000 at HudsonAlpha.

Data amounts generated in Gigabases of DNA sequence (Gb):

D. brevis: 149.4 Gb
D. rudis: 116 Gb
D. sorensenii: 261 Gb
Thomas Edison: 174 Gb

One issue occurred – the DNA for *D. coccinea* is difficult to extract, and relatively impure. This caused the sequencing reaction to fail. Zach has cleaned this DNA using a variety of methods, and we will resequence this library. The Genome Sequencing core at HudsonAlpha is aware of the issue and also helping Zach to resolve it.

Zach is currently working with this raw Illumina sequencing data and has completed analyses to: 1) use GenomeScope to determine the level of heterozygosity and genome size and repeat content of each individual, and 2) run SmudgePlot to determine and validate the estimated ploidy of each individual. Based on these decisions, we will move forward with one of these individuals to build a chromosome-scale genome assembly. As of right now, all wild species

appear to be autotetraploids.

An example of GenomeScope and Smudgeplots for *D. sorensenii* is below, showing that the base haploid genome size is around 1 Gigabase. Further, the Smudgeplots show us that the species' genome looks appears to be an allotetraploid (hybridization of two different species or individuals), which has previously been suggested but never shown (e.g. Gatt et al. 1998).