

## The Macadamia Genome Project

C. Nock<sup>1</sup>, A. Baten<sup>1</sup>, K. Langon<sup>1</sup>, B. Barkla<sup>1</sup>, A. Furtado<sup>2</sup>, B. Topp<sup>2</sup>, C. Hardner<sup>2</sup>, R. Henry<sup>2</sup>, and G. King<sup>1</sup>

<sup>1</sup>Southern Cross University; <sup>2</sup>University of Queensland

[cathy.nock@scu.edu.au](mailto:cathy.nock@scu.edu.au)

The Macadamia genome project aims to sequence the first reference genome for macadamia and the early-diverging eudicot plant family Proteaceae. A draft assembly for *Macadamia integrifolia* cultivar HAES 741 'Mauka', constructed using genome and transcriptome sequence data was released in 2016. Although fragmented, its excellent coverage of the functional gene space enabled prediction of over 30,000 protein-coding genes with 90% of the models expressed in leaf, shoot or flower tissue.

Since this time, a second draft V2 has been developed using additional Illumina short-read and PacBio long-read sequence data, and a bioinformatics pipeline that incorporates transcriptome data in the de novo assembly process. These measures have dramatically improved the assembly, reducing fragmentation and extending scaffold lengths (4,416 scaffolds, 414 kb N50). A dense genetic linkage map for 741 is under construction utilising pseudo-testcross and half-sib populations in order to anchor and orient assembled genome scaffolds.

This collaborative project is providing genetic markers for cultivar identification, paternity testing, and studies to trace domestication origins and gene flow among threatened wild populations in Australian subtropical rainforests. While assembling the genome of a highly heterozygous tree species with a long generation time is problematic, this can be mitigated by concurrent advances in sequencing technologies and bioinformatic approaches. It is anticipated that the macadamia genome will enable QTL and association (GWAS) mapping and marker assisted selection for breeding, and will be a fundamental resource for understanding the genetic basis of important crop traits.