**Exploring the mungbean pan-genome**

Mens C1, Grundy E1,2, Kharabian Masouleh A.1, Henry R.J.1, Mudge J.3, Farmer A.3, Udvardi M1

*c.mens@uq.edu.au*

1 Queensland Alliance for Agriculture and Food Innovation, The University of Queensland, Brisbane, Australia

2 Integrative Legume Research Group, School of Agriculture and Food Sustainability, The University of Queensland, Brisbane, Australia

3  National Center for Genome Resources, Santa Fe, New Mexico, United States

Mungbean (*Vigna radiata*) is a short-season tropical pulse grown on over six million hectares worldwide, mostly in Southern Asia with increasing importance in Australia and East Africa. Current commercial varieties show little genetic diversity with breeding focussed on a limited number of agronomic traits like disease resistance and flowering time. Luckily, we have access to diverse core collections and an Australian NAM (Nested Association Mapping) population.

Eight genetically diverse NAM parent accessions were selected based on available DArTseq SNP data including commercial varieties and an Australian wild accession (*V. radiata* var. *sublobata*). High molecular weight DNA was extracted for long-read HiFi PacBio sequencing (71-136x coverage). These reads were assembled into high quality genomes (contig N50 17.6-32.8 Mb) with a total assembly length varying from 512.8 to 577.2 Mb and a BUSCO completeness score of over 99% for each of the eight accessions. RNAseq data was generated using Illumina paired-end sequencing for shoot, root, and nodule tissues to allow for evidence-based gene annotation and differential gene expression analysis. Overall, this approach allowed for improved assembly statistics relative to the publicly available genomes VC1937A and Vrad\_JL7.

This high-quality pan-genome makes it possible to investigate diverse topics from crop improvement to evolutionary biology. It can be utilised by researchers and breeders to explore and improve complex traits like nitrogen fixation, yield stability, seed quality and adaptability to climate stresses. Integration of this sequence data with other data types like diverse phenotypes will help us move towards genome-based predictive breeding and crop improvement.