#063

Genomic strategies for optimisation of flowering time in canola

***Shannon Dillon*[1]**

Alexandre Boyer[1]
Alec Zwart[1]
Andrew Gock[3]
Jeremy Whish[2]
Bangyou Zheng[2]
Jing Wang[2]
Matthew Nelson[4]
Julianne Lilley[1]
Chris Helliwell[1]

[1] CSIRO Agriculture and Food, Canberra, Australia,

[2] CSIRO Agriculture and Food, St Lucia, Australia,

[3] CSIRO Environment, Canberra, Australia,

[4] CSIRO Agriculture and Food, Floreat, Australia

Background:

Targeting phenology to the optimal flowering window, thereby minimising the risk of yield impacts due to frost or extreme heat, is essential to maximising canola productivity and profitability. There is significant demand for flexible tools that leverage our understanding of genetic and environmental effects and their interactions to better inform management decisions on farm and in breeding.

Objective:

We developed a hybrid model that fuses ML and process-based crop growth models to estimate flowering time for any variety based on its genome, reducing the dependence of phenology optimisation on field-based assessments. In parallel we applied novel multi-'omics approaches to resolve genomic factors and environmental interactions underpinning variation in flowering time, to support breeding strategies for better adapted canola.

Methods:

We assessed phenology and generated multi-omic data in a diverse set of modern Australian and globally relevant canola varieties across Australian growing environments (18 combinations of site, year and TOS), as well as controlled environments (long days, short days and short days + vernalisation). We evaluated several approaches for prediction of phenology across environments and compared performance, from a base line crop model to genomic prediction. 'Omic associations were assessed via ensemble ML with adjustment for structure in a two-step process to account for correlation among features. Interactions with environment were evaluated overall and for SNP and transcript markers.

Results:

The genomic-APSIM hybrid model predicted flowering for new genotypes with high accuracy (R=0.87). Australian lines outperformed international lines, reflecting composition of the training panel. Modelling flowering time from genomic and environmental data directly was more promising (R>0.94), however less generalisable owing to the dependence on empirical data for training. Multi-environment association analyses identified locus level GxE and validation of QTL across experiments. Results were integrated across analysis to resolve a set of high confidence QTL and SNP and transcript markers tagging these, which can be used for tracking phenology traits in Australian breeding programs. The combined analysis indicate regulation by many small to moderate effect loci with redundancy, with variance in flowering explained by individual SNPs or transcripts not exceeding R2=0.2. Among the larger effect loci are known regulators of flowering such as the SOC1 transcription factor and homologues and FLC, but also putatively novel flowering loci.

Conclusions:

The genomic-APSIM-NG solution was the most promising, having a high level of accuracy for prediction of new genotypes and can be generalised to a wide range of environments. This makes it a practical option as a tool for management of productivity of new canola varieties in their region based on their phenology much earlier than is currently possible. Genome wide association analysis across field and controlled environments identified genomic loci and tagging SNP markers for tracking sources of variation in flowering time in Australian breeding programs. Multi-environment association analyses identified significant locus level GxE in both controlled and field conditions, and QTL were validated across multiple experiments. Our analysis identified previously known phenology QTL and genes, as well as multiple novel genomic sources of variation, some with suggested redundancy.