

B-139

Genomic selection for soybean yield and seed composition improvement

*Benjamin Stewart-Brown**, Center for Applied Genetic Technologies, University of Georgia, Georgia, USA

Qijian Song, Soybean Genomics and Improvement Lab, USD-ARS, Maryland, USA

Zenglu Li, Institute of Plant Breeding, Genetics and Genomics, University of Georgia, Georgia, USA

Yield is a complex trait that is controlled by many genes, each of which has a small genetic effect. Genomic prediction has been shown to be a promising breeding methodology for predicting phenotypes from genotypic data using a reference population. The impact of proper implementation of genomic selection is that one can reduce the amount of time, money, and effort spent on phenotyping in the field and devote these resources elsewhere, while ideally maintaining comparable accuracy compared to phenotypic selection. Using SoySNP6k Infinium Chips, we genotyped four RIL populations and diverse advanced lines to develop a training population of 538 lines. These lines were yield tested at four locations across three years (not all lines grown in each location each year). Genomic prediction models for yield and seed composition traits were developed utilizing Ridge Regression BLUP. We achieved a cross validation prediction accuracy of 0.60 for yield, but this is inflated due to population structure. To negate this inflated prediction accuracy, we predicted each RIL population separately. Training populations were constructed using a RIL population to predict within itself as well as using all other breeding lines to predict a RIL population. Prediction accuracies increased with training population size, were maintained at lower marker densities, and varied depending upon the population being predicted. Prediction accuracies for more heritable seed composition traits were higher compared to yield. Finally, prediction accuracy for predicting “within population” was often greater than predicting “across populations” when the training population size was held constant. The advantage of predicting “across populations” is that one can increase the training population size, allowing accuracy to approach “within population” accuracy. This will not occur in all cases as some RIL populations may be too genetically distinct from other breeding materials to be properly predicted across populations.