



## **North Central Soybean Research Program**

### **Accelerating soybean yield and composition improvement through genomic selection**

*Brian Diers (Project Leader), Matthew Hudson and Patrick Brown (University of Illinois), Randall Nelson (USDA-ARS/University of Illinois), Katy Rainey (Virginia Tech), Bill Beavis and Asheesh Singh (Iowa State University), George Graef and Aaron Lorenz (University of Nebraska); [bdiers@illinois.edu](mailto:bdiers@illinois.edu)*

### **Final Report 2016**

In this project, we are applying new genetic technology and breeding methods to address two bottlenecks in the soybean breeding process. The first bottleneck is the difficulty in effectively selecting new, high yielding varieties early in the breeding process. This is being addressed by testing a new breeding strategy called genomic selection (GS). The second bottleneck is our lack of good information for deciding what combinations of parents should be crossed in a breeding program. To identify solutions for this, simulations will be run to predict the best cross combinations. The GS methods will be developed using information gained from the soybean checkoff -sponsored SoyNAM project.

### **Project Objectives**

1. Test genomic selection in breeding populations.

This first step is to use the dataset from the SoyNAM project to optimize methods for using genetic markers to predict yield, protein concentration, and maturity of soybean lines (this is called genomic prediction and when plants or lines are selected with this method, it is called genomic selection). In this objective, predictions for the traits can be made and compared to trait data that is available. The second step is to conduct a breeding experiment using the optimized genomic selection approaches in populations of breeding lines developed by each cooperating breeder. Within each population, lines selected using genomic selection and traditional approaches will be compared to determine which method was the most successful.

2. Use genomic data to decide what cross combinations should be made by breeders. We will develop methods for selecting specific combinations of parents for crossing in soybean breeding nurseries. Once these methods are developed, they will be used by breeders to help them decide what cross combinations to make.

### **How this research can benefit farmers**

Soybean farmers are paid based on the weight of grain that they sell, so their profitability is directly tied to the yield of their varieties. Farmers are generally frustrated by the slow rate of yield gain for soybean, especially compared to corn,

and they would like breeders to make faster yield gains with the varieties they develop.

The goal of our project is to use genetic resources that have been developed through the SoyNAM project, which was previously funded by the soybean checkoff, to develop and test new methods that have the potential to increase the rate of genetic gain for yield.

If we are successful in developing these methods, breeders will use them to predict which experimental lines in their programs have the greatest yield potential before they are actually tested. Breeding programs will then be able to focus their yield testing resources on those lines with the greatest yield potential.

## **Results**

New genetic technology and breeding methods have the potential for boosting soybean yield gains and in this project, we are testing the effectiveness of some of these methods. The objectives of the project are: (1) test the effectiveness genomic selection compared to standard breeding methods in soybean breeding populations (2) develop methods to use genomic data in making cross combination decisions in breeding programs.

When the project was initiated, computer simulations were done to compare different methods for conducting genomic selection in soybean. These simulations showed that the best method was determined by the inheritance of the trait being selected and if this inheritance is controlled by additive genetic effects (the simplest form of genetic effects) and many genes, the different methods of genomic selection work equally well. These results were then used to test how well the performance of breeding lines can be predicted with genomic data using results for yield, other agronomic traits, and genetic markers from the very large soyNAM project that was funded by USB. Different subsets of the soyNAM dataset were used to make the predictions and it was found that the best predictions can be made by using the entire soyNAM dataset. The next step in the analysis was then to test whether these predictions can be further enhanced by combining the soyNAM dataset together with data from the USB funded genotyping of the entire USDA soybean germplasm collection. Research in testing the combined datasets is currently underway.

Using the prediction methods identified in the first objective, genomic selection is being tested in new breeding populations developed at the University of Illinois, USDA-ARS, Purdue University, and the University of Nebraska. This experiment is still in progress so final results are not available. This research was started by genotyping 5,630 experimental lines in breeding populations with genetic markers using an inexpensive marker method called genotyping-by-sequencing (GBS). Over 480 billion base pairs of DNA sequence was generated and individual lines were scored for 6,652 genetic markers on average.

The 5,630 experimental lines tested with GBS were grown in non-replicated, single-row plots during the summer of 2014. These tests were grown at all four

cooperating universities and were scored for date of maturity, plant height, and plant lodging and seed yield. The data from these 2014 plots were used together with the marker data to make predictions of the performance of the experimental lines and lines were selected with different methods. The methods used were:

- 1) Random lines (Random) (This was included to have a baseline of unselected lines.)
- 2) Phenotypic selection based on plant row yield in 2014 (PhenoSel)
- 3) Genomic selection for yield using historical NAM data for training (GenoSel\_Yld)
- 4) Genomic selection for yield and protein using historical NAM data for training (GenoSel\_YldPro)
- 5) Genomic selection for yield using 2014 plant row data and historical NAM data for training (GenoSel\_YldPlus).

Approximately 24 lines from each category were selected using each of the five methods in 24 breeding populations. Selected lines were evaluated in field tests during 2015 and are currently being evaluated during 2016 in Iowa, Illinois, Nebraska and Indiana to compare the effectiveness of the methods. At each location, 2090 2-row yield plots were grown. Data on yield, maturity, plant height and lodging were collected in these trials. In an across-location analysis of the 2015 results, we found that progenies selected using GenoSel\_YldPlus yielded the same as those selected using PhenoSel. Both methods were better than Random. Geno\_YldPro resulted in the lowest yields among the selection methods as expected due to the negative correlations between protein and yield. All selection groups were equal in maturity for each maturity group.

This result suggests that genomic selection can achieve yield gains on par with selections based on phenotypic data from plant rows. Given advances in genotyping, selections based on genomic predictions are expected to be easier and less expensive than measuring yield on plant rows in the future. It is important to note that the quick turnaround needed to generate the selections on genomic data during the spring of 2015 prevented optimization of the selections and more work will be done to look at ways to better optimize these methods.

The final objective was to develop methods for using genomic information to help breeders select the best cross combinations in their breeding programs. This research was initiated but in 2015, other researchers developed these methods which can be used in soybean. In 2017, we plan to use these developed methods in soybean breeding programs to predict the best cross combinations.