Nebraska Soybean Board
Final Report

DS VB

1/26/2018

Project Title: 90957 - Development of next generation sequencing applications for improving soybean

Contractor & Principal Investigator: Dr. David Hyten

1. What was the focus of the research project or educational activity?

This proposal was part of Dr. David Hyten's start-up funding to focus on developing novel methods that utilize new sequencing technology to better detect markers in the soybean genome.

• First objective was to develop a marker technology that will detect specific markers across the soybean genome in a cost effective and high-throughput manner.

• Second objective was to develop a method that can assay markers across the soybean genome from single pollen grains to help understand how the genomes between two soybean parents recombine with each other during the breeding process.

• Third objective was to develop a method that can assay markers in a specific region of the genome to determine if the region of the DNA has a higher than average probability (hotspot) for recombining different genes from two parents during the breeding process.

2. What are the major findings of the research or impacts of the educational activity?

Objective 1: The sequencer in the UNL soybean genomics lab was set up and was able to produce usable data. The GRA position for the project was filled by the PhD student, Samantha McConaughy and the lab manager position was filled by John Wang. We developed a method to design probes targeted to specific markers in soybean for GBS. The 4K marker probe set was designed and tested. On-target sequence obtained was 70-73% for samples which is a very good on-target rate. This increase of on-target rate came from optimization experiments which focused on testing different hybridization conditions and amplification conditions. Out of the 4,000 SNPs, 3,535 were successfully identified and genotyped. This protocol will be able to run probe sets with higher number of SNP (up to 30,000) or lower number of SNPs. In this test, we were able to successfully implement custom barcodes, which allowed us to multiplex the 96 samples. The barcoding system implemented in this test has the capability to multiplex over 2,000 samples in a single run. We have already used these barcodes in sequencing runs that included 192 samples.

Objective 2: A new DNA extraction protocol for single pollen grains was tested. With this protocol, quality DNA was obtained from single pollen (tested) by adding paramagnetic beads, PEG8000 and NaCl into extraction buffer. Further, the hybridization and amplification results from the MIPs protocol demonstrated that the single pollen DNA prepared with this protocol might be good for detection of markers but would likely improve using a whole genome amplification step before running MIPs.

Objective 3: We were able to successfully develop long-range PCR protocols for soybean to target areas of interest that we wish to sequence. This will be useful technique for the lab when future projects call for targeting a large area for marker development or to detect structural variation. A significant hurdle was discovered with this method in using it to determine local recombination.

The purpose with using long-range PCR was to be able to genotype two markers at the end of the amplicon and know if there was a recombination between them in a pooled sample of pollen. We discovered that in long range PCR, chimeras formation can only be reduced to 6.5%. This is significantly higher than our expected recombination rate for investigating recombination hotspots and would greatly overestimate recombination. Until better methods are available to prevent chimeras during long-range PCR, other methods for this specific application will have to be used to confirm recombination hotspots.

3. Briefly summarize, in lay terms, the impact your findings have had, or will have, on improving the productivity of soybeans in Nebraska and the U.S.

Overall, this project helped to get the soybean genomics lab up and running by developing methods to utilize our sequencing capability that can be utilized by the soybean breeding program. This increase in sequencing competency will also help with future projects that have a focus on better understanding how the genome contributes to agronomically important traits and how we can best improve the genetics of soybean to maximize traits to increase the genetic potential of soybean.

Through the project, we were able to demonstrate, that targeted genotyping by sequencing can be successfully used to genetically characterize molecular markers in soybean. It can be developed into a cheaper and more accurate method compared to other methods that use sequencing to characterize molecular markers. This method is an improvement over current methods for marker genotyping at UNL and will maximize the soybean breeding program's breeding efficiency leading to accelerating the rate of yield gain.

This new protocol for marker genotyping will also be used by the larger soybean breeding community for improving soybean yield nationally as indicated by the additional funding we have been able to leverage with this current project.

Our findings have indicated that with some further protocol efficiency gains it will be possible to detect how the soybean's genome recombines between two parents in a high-throughput manner. This will enable UNL soybean researchers to further study how recombination is effected by the environment and how that affects the breeding program in creating new varieties.

4. Describe how your findings have been (or soon will be) distributed to (a) farmers and (b) public researchers. List specific publications, websites, press releases. etc.

Dr. Hyten has given multiple seminars to public researchers that highlighted the work being performed through this project and the current results in this project. In addition, Dr. Hyten will give a presentation sharing results about this project at the International Plant and Animal Genome Conference in January, 2018. During the duration of this project Dr. Hyten hosted a lab tour to the NSB and several other groups that described the genotyping and the benefits of the genotyping methods being developed from the current findings. A feature article written for the Agronomy & Horticulture Department 2016 Annual Newsletter was published in collaboration

with Dr. Graef and Dr. Clemente which describes the benefits of developing new molecular marker technology.

5. Did the NE soybean checkoff funding support for your project leverage any additional state or Federal funding support? (Please list sources and dollars approved.)

Yes.

USB project titled "Utilizing Unique Genetic Diversity to Combine Elevated Protein Concentration with High Yield in New Varieties and Experimental Lines" year 1 UNL awarded $79,970 out of $456,772 and year 2 UNL awarded $89,275 out of $515,426.

NCSRP titled, "Increasing the rate of genetic gain for yield in Soybean Breeding programs", 3 year project for a total of $3,125,564 with a total of $602,811 being awarded to UNL over the 3 years and $403,607 already received for the first two years.