Status update:
Reporting period: Final

MSR&PC Identification Number: 150-4120-18-02
PI: Cory Hirsch
Department: Department of Plant Pathology, University of Minnesota
Project title: High-throughput hyperspectral phenotyping to understand soybean cultivar response to diseases
Dates of reporting: 5/1/18 – 4/30/19

Project objectives and describe any activity for each:

1. *Determine the influence of soybean line on root disease hyperspectral profiles.*

To address this objective in this project we completed field-based experiments of brown stem rot and sudden death syndrome inoculations. In short, we conducted inoculations on 20 varieties at two different field sites for the two diseases. The sites also had paired non-inoculated plants for comparison purposes. Throughout the growing season we collected hyperspectral images and handheld hyperspectral data weekly. Collecting data on the number of genotypes and throughout the growing season is imperative to determine if hyperspectral profiles indicative of soybean plant resistance to root infection are consistent and accurate throughout a panel of soybean cultivars and breeding lines. Throughout the project we extracted and analyzed our hyperspectral images from the previously acquired images. We continued to improve the accuracy and speed of our hyperspectral data extraction, which is imperative to our data analysis pipeline when working with large amounts of data.

2. *Validate and confirm repeatability of hyperspectral observations of soybean root diseases.*

Throughout this project we conducted field experiments, spending time planning and prepping field sites, planting, inoculating, and collecting data (both hyperspectral data and field evaluations). We also collected hyperspectral and observational data from a set of 20 lines in a greenhouse inoculated with sudden death syndrome and non-inoculated plants. This data was be used to inform us about the accuracy and reproducibility of hyperspectral measured responses in different environments and to correlate controlled and field grown trials. Throughout the project we focused our data analysis to understand the hyperspectral profile response of plants to SDS and BSR. We did this by looking not only within individual trials, but also between different experiments that were collected. This analysis informed us about the accuracy and reproducibility of hyperspectral measured responses in different environments and to correlate controlled and field grown trials.

Specific project achievements accomplished during this reporting period:

This report covers progress for the entire project, from 5/1/18 to 4/30/19. This is a new project funded by the MSR&PC that builds on our knowledge and pipelines developed during our previous MSR&PC funded project the previous year. The main activities for this project were related to plant growth, data collection, data extraction, quality control, and data analysis. For this project, we are researching two different soybean pathogens, brown stem rot (BSR) and soybean sudden death syndrome (SDS). The two objectives of this project are: 1) *Determine the*

influence of soybean line on root disease hyperspectral profiles and 2) *Validate and confirm repeatability of hyperspectral observations of soybean root diseases*, regarding field grown plants. Towards these objectives we completed data acquisition and analyses from controlled environments and field sites to increase the number of genotypes that have data collected and analyzed to address both Objective 1 and 2.

Towards the first objective we completed replicated fields trials of BSR and SDS inoculations for 20 soybean cultivars. The selected cultivars were based on maturity rating and their disease score to BSR and SDS. We choose lines that would mature around the same time in St. Paul and Rosemount, MN and lines that had a range of response to the pathogens (from resistant to highly susceptible). This allowed us different levels of comparisons for the disease action on the plant and provide us the ability to understand the disease at a finer scale. The BSR trial was grown on the St. Paul Campus Fields, while the SDS trial was grown at the Rosemount Research and Outreach Center. This was due to limitations of applying different diseases at the field sites, with the St. Paul fields being more restrictive. Starting around 30 days after planting hyperspectral measurements were made approximately once a week. At the SDS trial in Rosemount we grew 20 cultivars planted in three treatments (non-inoculated, inoculated, non-inoculum substrate) and replicated five times for 300 total plots. Within each of these 300 plots we phenotyped 2 plants per plot, taking 2 measurements per plant. These measurements were taken using a handheld hyperspectral instrument to obtain hyperspectral profiles. At our BSR trials on the St. Paul Campus we planted 10 cultivar varieties on inoculated and non-inoculated plots with five replicates. The hyperspectral data collected at this site utilized the available plant phenotyping cart on the St. Paul Campus with images being collected approximately every week starting in early July and continuing until the end of the growing season. Unlike the previously mentioned handheld instrument, this cart is not limited to specific locations on the plant for measurements as it measures the entire plant (or plot) at a time.

For the second objective we collected handheld hyperspectral data of a mirrored experiment of our Rosemount SDS trial of 300 plots grown in a greenhouse on the St. Paul Campus. These plants were inoculated at planting and measurements were taking approximately once a week starting at about 4 weeks after planting until maturity. In addition, we also planted and collected data for multiple replications of BSR inoculated plants in controlled growth chamber environments.

After the completion of the data collection we analyzed and dissected the data collected from of field grown and controlled environment infected plants. We generated a large amount of phenotyping data for this project and worked out important data storage and handling techniques necessary for successful completion of the project objectives.

We are able to process raw data to normalized utilized data in less than a day now. This is important to efficiently analyze current and future datasets generated for completion of the project objectives. We have extracted all the data from our field hyperspectral imaging. Initial data analysis during late vegetative and early reproductive stages of growth, we could detect unique spectral profiles for these two soybean diseases. Although further testing and validation needs to be completed, we are very encouraged by these initial results as our detections are occurring before visual foliar symptoms have appeared. We have also begun to use more sophisticated and hopefully robust ways of analyzing the data to fulfill the project objectives.

These include testing different regression methods and machine learning algorithms to test if we are able to robustly determine infection of SDS or BSR, and to be able to do it before visual symptoms occur. These techniques are heavily reliant on our hyperspectral data being of good quality, having variation within our disease response across genotypes, and also have end of season measurements to correlate to. By the of this project all of these pieces were in place for our field-based assay and we actively pursued these analytical techniques throughout the end of the project.

We are happy to report that using different machine learning techniques on our field grown BSR trial has yielded very positive results. We have successfully used support vector machine learning to build a model and prediction algorithm for BSR. This model is currently able to predict the disease status at plants at approximately 85% accuracy. We are very pleased with this result and the prospects of moving forward using the techniques we developed. This high level of accuracy is maintained even across different genotypes, which indicates our prediction model is fairly robust. We are currently applying these methods and analysis techniques to our SDS dataset. We are very hopeful we will receive similar encouraging results. Lastly, it is worth mentioning that our hyperspectral instruments collect data for 100s of wavelengths, but our predictions are being made with only 8 different wavelengths. This is important as we are getting close to the number of wavelengths that could be applied to a different type of sensor that would collect just those important wavelengths that would be much cheaper to implement and easier for data analysis as well.

Challenges encountered:

Our biggest challenge in this project is ensuring adequate and consistent inoculation of plants with the two diseases we are studying, brown stem rot and sudden death syndrome. This challenge is both prevalent in controlled and field conditions. We are taking all the necessary steps to mitigate this challenge as best we can, but we can't control everything within experiments or between them either. We are continuing to perfect our inoculation methods and growing conditions to limit inoculation variation as much as we can.

Information/data disseminated from research conducted during this reporting period:

In the first quarter of this project the research plan and initial progress was informally verbally communicated at the Minnesota Soybean Production Action Team Summer Meeting. This project was shared with the MSR&PC, soybean growers, and University researchers and extension agents.

In the second quarter of this project parts of our project was informally presented and discussed with Sentera (https://sentera.com). This Minnesota based company is interested in crop health assessments and seems to be interested in partnering with research teams. The project was presented and discussed with about 10 people from Sentera's Research and Development team, their Sensors team, and their Analytics team.

We are currently in the process of writing a manuscript centered around our BSR field trials and prediction models. We are hoping that manuscript will be disseminated to the community by the end of 2019.

Technology transfer:

There was no technology transferred in this reporting period.