DS VB 11/1/2017

Nebraska Soybean Board
Year-End Research Findings Report

Please use this form to summarize the practical benefits of your research project and what has been accomplished. Your answers need to convey why the project is important and how the results impact soybean production.

Project Title: *Soybean Phytophthora root rot disease control by identifying protein interactions of virulence effectors in Phytophthora Sojae to soybean proteins*

Contractor & Principal Investigator: Chi Zhang

Please check/fill in appropriate box: ☐ ***Continuation research project***
☑ ***Year 1 of 2 research project (for example: Year 1 of 2)***

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

After identification of potential virulence effectors in Phytophthora Sojae (P. Sojae), we conducted research to disocver protein-protein interactions between P. Sojae virulence effectors and proteins in soybean to understand how virulence effectors cause the Phytophthora root rot disease. P. Sojae, a notorious oomycete pathogen to soybeans, secretes virulence effectors as major 'weapons' to attack the target host organisms, and understanding the biological functions of virulence effector is the key step to reveal the mechanism of oomycete pathogenicity to hosts and control the disease. Therefore, we are conducting the so-far blank research area of protein-protein interactions between P. Sojae and soybean by developing a statistical model based on protein-domain interaction prediction. As a result, the mechanism under which P. Sojae interacts with and infects soybean can be elucidated, which may directly lead to feasible and efficient methods to prevent soybean from the Phytophthora root rot disease.

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

We were working on the development of a new computational method to identify the interactions between virulence factor P. Sojae and soybean proteins based on the protein-domain interaction network and Bayesian statistics. From protein sequences that we collected for P. Sojae and soybean, we conducted domain analysis for >80,000 protein sequences for both species, and > 220,000 protein domains were identified. We were developing a promising model that was designed on Bayesian theorem and graph theory. To improve the accuracy, a directed graph was employed by this model, and pathogenic and host protein domains were distinguished in the network, which is novel in this type of algorithms. Based the preliminary evaluation, this new improvement can make the prediction accuracy better than other existing methods, such as methods based on sequence similarity, position weighted matrix (PWM), and hidden Markov model (HMM). Further development and evaluation are still under construction.

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.

The long-term goal of our project is to discover the mechanism under which P. Sojae interacts with and infects soybean, and the knowledge of interaction mechanism between P. Sojae and soybean may directly lead to feasible and efficient methods to prevent soybean from the Phytophthora root rot disease. The interactions between P. Sojae and soybean are executed by protein interactions, especially by interactions between P. Sojae virulence effectors and soybean proteins. We designed a statistical model to identify the protein interactions based on protein-domain interaction networks.

*****This form must be completed and submitted with the fourth quarter report.***

A protein domain is the basic building block of a protein, and one protein may have multiple domains. The domain interaction network was trained from known domain interaction database. Once our algorithm can identify the protein-domain interactions between P. Sojae and soybean, we further obtain the protein-protein interaction information and hence to confirm the predicted virulence effectors and predict the functions of these virulence effectors.

With our developed computational methods, my lab also collaborated with Drs. Tom Clemente and Bin Yu to analyzed the nanopore sequence data to identify the insertion position of transgene in soybean.

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.

Based on our research, we co-authored four manuscripts: three of them have been published, and the other is under preparation.

1. Y. Dou, S. Li, W. Yang, K. Liu, Q. Du, G. Ren, B. Yu, **Chi Zhang**. Genome-wide Discovery of Circular RNAs in the Leaf and Seedling Tissues of Arabidopsis Thaliana. *Current Genomics* (2017); 18(4):360-365.
2. S. Li, K. Liu, S. Zhang, X. Wang, K. Rogers, G. Ren, **Chi Zhang**, B. Yu. STV1, a ribosomal protein, binds primary microRNA transcripts to promote their interaction with the processing complex in Arabidopsis. *Proc Natl Acad Sci* USA (2017); 114(6):1424-1429.
3. H. Wang, Y. Dou§, **Chi Zhang**, B. Yu, Y. Liu, T. Heng-Moss, G. Lu, M. Wachholtz, J. Bradshaw, P. Twigg, G Sarath. Insect and plant-derived miRNAs in greenbug (Schizaphis graminum) and yellow sugarcane aphid (Sipha flava) revealed by deep sequencing. *Genes (2016)*; 599:68-7.
4. T. Lu, J. Alfano, **Chi Zhang**. Pathogen-host protein-protein interaction of *Phytophthora* based on directional domain-domain interaction. (Preparation)

We attend several notational wide meetings to present our discoveries:

- Oral presentation: Kan Liu, Qian Du, Guodong Ren, Bin Yu, Chi Zhang. “Identification of Different Alternative Splicing Events of Plant and Human RNA-Seq Data with Beta-Distribution Model”. IEEE International Conference on Electro Information Technology (eit2017), Lincoln, NE, 2017
- Poster Presentation: Qian Du, Weilong Yang, Chi Zhang. “Identification of protein domain interactions with directed graph model based on Bayesian statistics.” Next Generation Dx SUMMIT; August 15-18, 2017; Washington DC

We also designed a webpage to show relevant information there:
http://sysbio.unl.edu/resources

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.)

I submitted several proposes to NSF, DOE, and USDA the past year, as PI or co-PI. For example, a proposal, as a PI, was submitted to the MathBioSys center grant supported by NSF and Simons Foundation on September 27th. We request about 8.9M dollars to develop a center for Mathematical Systems Biology to develop predictive model for linking genome to phenome. This proposal is pending. As a PI, a proposal to DOE JGI program for sequencing a plant genome was funded and as a Co-PI, a grant supported by the EPSCoR RII Track-2 FEC program in NSF was funded.

Nebraska Soybean Board
Year-End Research Findings Report

Please use this form to summarize the practical benefits of your research project and what has been accomplished. Your answers need to convey why the project is important and how the results impact soybean production.

*****This form must be completed and submitted with the fourth quarter report.***