

Nebraska Soybean Board
FINAL Research Report Form



Note: Submit this report no later than 90 days after the NSB-funded project officially terminates.

This post-project 90-day time-frame will allow the Lead PI time to complete any final data analysis and a final technical report, plus the drafting of any articles for submission to scientific journals. Note that this completed report will be provided to the National Soybean Checkoff Research Database, (soybeanresearchdata.com).

Project # and Title: #1739: Dissect Carbon Capture and Partitioning in Soybean with a Systems Approach for Yield and Oil Content Improvement

Principal Investigator: Chi Zhang

Co-PI's & Institutions: Thomas E Clemente and Bin Yu, University of Nebraska - Lincoln

Project Date (Including Extension): 10/01/2019 **to** 09/30/2021 **(For example: mm/dd/yyyy to mm/dd/yyyy)**

Total Budget for Project: \$ 111,218.00

1. Briefly State the Rational for the Research:

Plants, including soybean, extract carbon dioxide (CO₂) from the atmosphere as they grow, and pathways for carbon capture and partitioning, such as photosynthesis and fatty acid synthesis, directly affect the biomass, seed yields, and oil content. While it is apparent that the carbon capture and partitioning has direct effect on the soybean yield and oil content, the linkage between carbon capture/partitioning and their corresponding phenotypes is poorly understood in transgenic soybeans, carrying some key genes involved in photosynthesis and/or fatty acid synthesis. Therefore, we proposed a study to apply a systems approach by analyzing transcriptomic sequencing data for variant transgenic lines to uncover the intricate interactions among the genes involved in carbon capture and partitioning pathways, and to develop a predictive model by constructing a gene co-expression network to link them to the corresponding phenotypes, such as seed size and oil content. The predictive model allows the identification of key cellular processes that need to be considered to break the embryo's natural restriction to uncontrolled seed lipid increase. The key genes will be used to produce new soybean lines with different oil contents.

2. Research Objectives: (copy from project, but keep in a brief bullet format)

Objective One: Identify the transcriptomic profiles of transgenic soybeans with carbon capture and partitioning pathway modifications.

Objective Two: Systematically discover carbon capture and partitioning pathways, genes, and their interactions by a statistical model for co-expression gene networks and linkage to phenotyping data

Nebraska Soybean Board
FINAL Research Report Form

3. General Approach Used and (if applicable) the Nebraska Test Locations:

We conduct RNA-seq experiments at three time points during seed development for several transgenic soybean lines, carrying AtWri1, AtWri1/AtDGAT1, AtDGAT1/AtKasII, and one wild type soybean as control. The corresponding metabolomic profiles, including photosynthesis and fatty acid content, were measured. AtWRI1 is a master regulator of fatty acid (FA) biosynthesis, and DGAT1 encodes an enzyme catalyzing the final and rate limiting step of triacylglycerides biosynthesis. Soybeans were growing in both fields and greenhouse. The metabolic and transcriptional responses to AtWRI1 and AtDGAT1 expression in soybean seeds were evaluated. Differentially expressed genes were identified, and their enriched pathways were analyzed.

Combined with gene expression profiles and known AtWri1 DNA binding motifs (AW-box), we developed a bioinformatics tool to predict AtWri1 targets in the soybean genome. Novel putative AtWRI1 targets, presenting an AW-box in the upstream region of the genes, were identified by comparison with an event that harbors only AtWRI1.

We also constructed a co-expression gene network based on the transcriptomic data from all soybean lines, and identify the functional modules or pathways related to carbon capture and partitioning as well as their corresponding phenotypes. In the gene co-expression network, each gene is one node, and two genes share a link if their gene expression profiles are correlated. After the identification of modules in the co-expression network, and a statistical model were developed to determine which modules are related to the oil content. The genes identified by this method were directly related to the carbon capture and partitioning pathways as well as the corresponding phenotype.

The Illumina short read RNA-seq experiments were executed by UNMC genomic core facility and metabolomic profiles were measured by our collaborators at the University of North Texas.

4. Describe Deliverables & Significance Attained for Each Research Objective:

The improvement of the yield and seed composition in transgenic soybean lines comes from knowledge about carbon capture and partitioning pathways. Many photosynthesis/fatty acid synthesis genes are discovered, and some of them were transferred into soybeans. A natural next step will be to systematically mine this massive transcriptomic data for the key information to answer the question - how the carbon capturing and partitioning is altered in the transgenic plants, which will provide strategies to manipulate related processes through genetic engineering.

For objective one, we evaluated the metabolic and transcriptional responses to AtWRI1 and AtDGAT1 expression in soybean seeds. The events expressing these genes in the embryo did not show an increase in total FA content, but they responded with changes in the oil and carbohydrate composition. Transcriptomic studies revealed a down-regulation of genes putatively encoding for oil body packaging proteins, and a strong induction of genes annotated as lipases and FA biosynthesis inhibitors.

Targeted metabolomics analysis showed that carbon from sugar phosphates could be used for FA competing pathways, such as starch and cell wall polysaccharides, contributing to the restriction in oil accumulation. These results allowed the identification of key cellular processes that need to be considered to break the embryo's natural restriction to uncontrolled seed lipid increase

Novel putative AtWRI1 targets, presenting an AW-box in the upstream region of the genes, were identified by comparison with an event that harbors only AtWRI1.

Despite the importance, the systems biology research on the mechanism of carbon capture and partitioning is still in its infancy. The Objective two can spur the research interest in this field, and the complete pathways for carbon capture and partitioning can subsequently guide functional studies from genome to phenome.

**Nebraska Soybean Board
FINAL Research Report Form**

4. Describe Deliverables & Significance Attained for Each Research Objective (continued)

Our understanding of carbon capture and partitioning in soybeans is limited, and the corresponding phenotype is not yet fully predictive. The study we conducted improved the research in this field, because we developed a systems approach including transcriptome analysis, network analysis, and genome sequence identification. Our study revealed the capture and partitioning pathways in transgenic soybean lines and the link between genotype and the soybean seed oil content. The key genes that we identified will be used to develop new soybean lines for variant oil contents. The discovered genotype and phenotype interactions can guide us to the development of high yield and high oil content soybeans. This can substantially impact the soybean industry in the US and the entire world, because currently about 30 million tons of soybean oil is produced worldwide, constituting about half of worldwide edible vegetable oil production.

5. List where the Project Research Results/Findings were Publicized:

- (1) H. Yu, Q. Du, M. Campbell, B. Yu, H. Walia, Chi Zhang*. Genome-Wide Discovery of Natural Variation in Pre-mRNA Splicing and Prioritizing Causal Alternative Splicing to Salt Stress Response in Rice. *New Phytologist* (2021); 230: 1273-1287. DOI: 10.1111/nph.17189.
- (2) Li, M., Yu, H., Liu, K., Yang, W., Zhou, B., Gan, L., Chi Zhang, Yu, B. Serrate-Associated Protein 1, a splicing-related protein, promotes miRNA biogenesis in Arabidopsis. *New Phytologist* (2021); 32: 1959-1973.
- (3) C.L. Arias, T. Quach, T. Huynh, H. Nguyen, A. Moretti, Y. Shi, M. Guo, A. Rasoul, K. Van, L. McHale, T.E. Clemente, A. Paula Alonso, Chi Zhang*. Expression of AtWRI1 and AtDGAT1 during soybean embryo development influences oil and carbohydrate metabolism. (2021) (Submitted)
- (4) L. Busta., I. Dweikat, S.J. Sato, H. Qu, Y. Xue, B. Zhou, L. Gan, B. Yu, T.E. Clemente, E.B. Cahoon*, Chi Zhang*. Chemical and genetic variation in feral hemp populations across the Nebraska climate gradient. *Phytochemistry* (2021); (Accepted)
- (5) G. Moiseyev\$, K. Park, A. Cui\$, D. Freitas\$, D. Rajagopal\$, A.R. Konda, M. Martin-Olenski\$, M. Mcham\$, K. Liu\$, Q. Du\$, J. C. Schnable, E.N. Moriyama, E. B. Cahoon, Chi Zhang. RGPDB: Database of root-associated genes and promoters in maize, soybean, and sorghum. *DATABASE* (2020); baaa038.
- (6) Li S, Li M, Liu K, Zhang H, Zhang S, Chi Zhang, Yu B. MAC5, an RNA-binding protein, protects pri-miRNAs from SERRATE-dependent exonuclease activities. *Proc Natl Acad Sci USA* (2020) doi: 10.1073/pnas.2008283117. (Effort: 15%).
- (7) S. Jia, A. Yobi, M. Naldrett, S. Alvarez, R. Angelovici, Chi Zhang, D. Holding. Deletion of maize RDM4 disrupts endosperm maturation, vegetative growth and the transcriptomic response to cold stress. *Journal of Experimental Botany* (2020); eraa325.

--- Conference presentations: PAG XXVIII Jan 2020 and ASPB 2021 Worldwide Summit
--- A web page to show relevant information: <http://sysbio.unl.edu/resources>

Note: The above boxes will automatically accommodate for your text inputs; HOWEVER, the Final Report comprised of the above listed items must be kept to THREE PAGES. A Technical Report of no more than TEN PAGES (preferably fewer) can be appended to this report.

Submit both reports as a single PDF with this file name format: #XXX > FINAL > Project Title > PI last name

Please email this completed form to the Agriculture Research Division (jmcmahon10@unl.edu) based on the reporting schedule given to you. If you have any questions, please call Jen McMahon at the ARD at 2-7082.