**Summery**

Plants extract carbon dioxide (CO2) from the atmosphere to synthesize carbohydrates, proteins, and lipids. As the plants grow, the pathways for carbon capture and partitioning among different types of macromolecules directly affect the yields, and seed composition. For soybeans, possessing an efficient carbon capture pathway, such as photosynthesis, during seed development can increase the soybean yield, and redirecting flux to fatty acids by manipulating carbon partitioning pathways, including fatty acid synthesis, can increase the oil content in soybean seeds by more than 20%. While it is apparent that the carbon capture and partitioning has a significant effect on the soybean yield and seed composition, the link between carbon capture/partitioning pathway regulation and their corresponding phenotypes is poorly understood in transgenic soybeans, carrying some key genes involved in photosynthesis and/or fatty acid synthesis. Therefore, uncovering the intricate interactions among the genes involved in carbon capture and partitioning pathways is important since it is a key step towards generating optimally-transgenic soybeans with enhanced capture and partitioning ability for seed development to further increase the yield and the oil content.

**Description of Research Project**

Since carbon capture and partitioning are important, many studies have been conducted. Especially, our understanding of the pathways involved in carbon capture and carbon flux to fatty acids in plants has been increased greatly in the past couple decades. Many key genes for photosynthesis and fatty acid synthesis were discovered. In photosynthesis pathways, the gene *ictB* in *Synechococcus* was discovered and found involved in HCO3-accumulation, and the transgenic plants, including *Arabidopsis thaliana* (Arabidopsis) and *Nicotiana tabacum,* overexpressing *ictB,* exhibited significantly faster photosynthetic rates than their wild types. Another enzyme involved in photosynthesis is edoheptulose-1,7-bisphosphatase (SBPase), which is an important enzyme involved in photosynthetic carbon fixation in the Calvin cycle. Transgenic lines with increased SBPase activity accumulated up to 50% more sucrose and starch than the wild-type. In fatty acid synthesis, the enzyme, AtDGAT1 in Arabidopsis, is required for triacylglycerol accumulation in seeds, and the enzyme, KasI, β-Ketoacyl-[acyl carrier protein] synthase I, catalyzes the elongation of de novo fatty acid (FA) synthesis. The expression of these genes is regulated by the transcription factor WRINKLED1 (Wri1) protein, which is crucial for oil accumulation in maturing Arabidopsis seeds. These genes were transferred into soybeans in the Co-PI Clemente’s Lab to improve the photosynthesis and oil content in soybean seeds. Indeed, photosynthesis measurements revealed significant increases in carbon capture in soybean events carrying the transgenes ictB and SBPase at late reproduction stages relative to the corresponding controls. Above ground biomass estimates showed increases in means of the transgenic events relative to controls. To improve the oil content in seeds, three genes, DGAT1, KasII, and Wri1 in Arabidopsis were individually and combinedly transferred into soybean by Co-PI Clemente’s Lab. Phenotyping data (oil/protein over canopy level, 100 seed weight, and plot weight) were measured and are being ascertained on the harvest collected from plots sown with transgenic events carrying the transgene stacks of AtDGAT1/AtKasII, AtWri1/AtDGAT1 along with a four transgene stack, in a lanceolate leaf genetic background, consisting of ictB/SBPase, AtWri1, and AtDGAT1.

In spite of the progress in manipulation of carbon capture and partitioning pathways in transgenic soybeans carrying different key genes, a number of questions remain to be resolved, particularly those relating to what kinds of interactions and coordination among key genes exist in a given pathway and/or from different pathways and how carbon capture and flux to fatty acid synthesis are controlled to generate corresponding phenotypes. Investigations of these questions can guide us to further develop transgenic soybeans that possesses quality phenotypes, such as high seed yield and oil content, besides increased biomass. One promising method to investigate these hypotheses is applying systems approach to analyze transcriptomic sequencing data for variant transgenic lines carrying different functional genes, which have altered metabolism but not-fully-understood phenotypes, and to genome-wide seek soybean genes that show different patterns of expression between different development stages and/or between different transgenic lines. An additional value of investigating correlated gene expression profiles by construction of a gene co-expression network is to identify regulatory genes to those encoding the enzymes of biosynthetic pathways, and hence, link the gene interactions to phenotypes. Therefore, this project will explore the differentially expressed genes in several transgenic soybean lines and identify the linkage between gene expression patterns and phenotyping results, with two specific objectives:

**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*.*

**G. Statement of novelty and significance**

Our understanding on carbon capture and partitioning in soybeans is limited, and corresponding phenotype is not yet fully predictive. To improve research in this field, a systems approach is proposed. If successful, the new project will lead to a complete picture of carbon capture and partitioning pathways in transgenic soybean lines and a predictive model to link genotype to phenotype, which is novel in this field.

The improvement of the yield and seed composition in transgenic soybean lines comes from the 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. Our timely research will work on photosynthesis/fatty acid synthesis in soybeans and directly unveil interactions among genes involved in carbon capture and partitioning pathways, which can guide us to the development of high yield and high oil content soybeans. This can substantially impact to 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. In addition, despite the importance, the systems biology research on the mechanism of carbon capture and partitioning is still in its infancy, and this project will 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.