

Exploring the Mechanism of Soybean Oil Deposition and Increasing Soybean Oil Content Through Genetic Engineering

Submitted by Yun Lin, Lila O. Vodkin, and Jack M. Widholm,
University of Illinois

Lin: yunlin@uiuc.edu
Vodkin: l-vodkin@uiuc.edu
Widholm: widholm@uiuc.edu

IMBA Project 2007-5

Project Period: 1/1/07-12/31/08

Progress Report (covers period of 1/1/07-12/31/07)

The rising cost of petroleum and vegetable oils has generated high interest in developing soybean with greater oil content. This research addresses that issue. Earlier studies in *Arabidopsis* suggested that several master developmental factors determine the architecture of the transcriptional network in seeds, and that the fatty acid biosynthetic pathway is integrated into the network through the master regulator *FUS3*. To date, research under this project showed the feasibility of *Arabidopsis* mesophyll protoplasts in discovering further detailed structures of seed transcriptional network within a short turnover time. Similar methods could likely be applied to soybean. In addition, transgenic soybean lines expressing the *Arabidopsis FUS3* driven by *LECTIN* promoter were generated. Although the original goal was to increase fatty acid and oil synthesis through over-expression of *FUS3*, the transgene appeared to interfere with soybean reproductive development. Comparison of the cDNA and the genomic DNA sequences of soybean *FUS3* (*GmFUS3*) confirmed a stop codon that truncates *GmFUS3*. Whether this difference explains the low oil content of soybean relative to *Arabidopsis*, and the unexpected deleterious effect of *AtFUS3* in transgenic soybean awaits future research. It is also unknown whether the soybean genome contains additional *FUS3* genes.

Specific Objectives

Objective 1. Testing the ability of *GmFUS3* to induce oil biosynthetic genes.

- (a) The *GmFUS3* cDNA sequence (Gm-r1089-1051 as 5' sequence and Gm-c1036-6074 as 3' sequence) and the corresponding genomic sequence were compared. An open reading frame that lacks the N-terminal 106 amino acids in the *Arabidopsis FUS3* (*AtFUS3*) was confirmed. Thus, the available *GmFUS3* cDNA clone represents a truncated *FUS3* gene encoded by the soybean genome. This truncation also deletes the N-terminal 21 amino acids from the B3 DNA binding domain (B3 in *AtFUS3* is 106-amino acids long, whereas B3 in *GmFUS3* is 85-amino acids long). Whether the soybean genome encodes more than one *FUS3* gene, and if so, whether they differ in length requires screening of soybean cDNA libraries and/or completion of the genomic sequence. The soybean genome that is

scheduled for completion in the next few months from JGI sequencing will be scanned for other FUSCA sequences.

- (b) Because it was unclear whether the soybean genome contains full length *FUS3* genes, and there also lacked full length cDNA clones for additional soybean transcriptional factors, we first tested Arabidopsis seed transcriptional factors in Arabidopsis mesophyll protoplasts. After transiently expressing *AtLEC1*, *AtLEC2*, *AtFUS3*, and *AtABI3* in protoplasts, the expression of 35 seed storage genes and seed transcriptional factors were measured by real-time RT-PCR. Because the master regulators function at the top of the seed transcriptional network to organize downstream branches, further structural details of the transcriptional network were revealed by the responses of downstream genes to the master regulators. Here, the analysis of a small number of genes demonstrated the feasibility of this approach to resolve genome-wide network structures using high-throughput methods. It is very likely that similar methodology can be applied to soybean.

Objective 2. Increasing soybean oil content by modification of *GmFUS3* expression in transgenic soybean.

Due to the lack of full length *GmFUS3* clones, we first tried to express the *AtFUS3* in soybean seeds. The *AtFUS3* gene was cloned downstream of the soybean *LECTIN* promoter and transgenic lines expressing the construct were generated. The transgenic soybean (T0 generation) showed severely reduced fertility, suggesting interference of the Arabidopsis gene with soybean seed or reproductive development. Different possibilities exist regarding the deleterious effects of *AtFUS3* in soybean. In one, the seed specific *LECTIN* promoter is too strong so that the expression level of the transgene is much higher than the endogenous *GmFUS3*. Second, the effect of the transgene reflects an intrinsic difference between *AtFUS3* and *GmFUS3* proteins. This hypothesis would be consistent with the possibility that the known truncated *GmFUS3* is the only *FUS3* gene in the soybean genome. If this truncation attenuates the function of *GmFUS3* in soybean, it may also explain why soybean oil content is lower than Arabidopsis. It is still possible that the soybean genome contains both full length and truncated *FUS3*, and their relative expression levels modulate seed development and storage pathways.

Further Plans to Complete the Project

1. Promoter analysis will be conducted as another approach to understand the transcriptional network in Arabidopsis seeds. Success in the combined experimental and bioinformatic studies will guide future studies of soybean seed transcriptional network.
2. A construct containing *LECTIN* promoter driving the truncated *GmFUS3* will be introduced into transgenic soybean and its influence on seed development and oil deposition will be evaluated.