Biological Design of *Lemnaceae* Aquatic Plants for Biodiesel Production

Evan Ernst*,[1,2] James Birchler,[3] Eric Lam,[4] Jorg Schwender,[5] John Shanklin,[5] and **Robert A. Martienssen (martiens@cshl.edu)**[1,2]

[1]Cold Spring Harbor Laboratory, Cold Spring Harbor, NY; [2]Howard Hughes Medical Institute, Cold Spring Harbor, NY; [3]University of Missouri, Columbia, MO; [4]Rutgers University, New Brunswick, NJ; and [5]Brookhaven National Laboratory, Brookhaven, NY

Project Goals:

1. **Leveraging our transformation methods, we will develop a comprehensive toolset for genetic manipulation of *Lemnaceae*. We will establish CRISPR/Cas9 genome editing to complement our artificial miRNA silencing methods. We will construct artificial chromosomes in *Lemna minor* to potentiate whole pathway engineering.**
2. **Resting and over-wintering fronds have higher starch content than corn kernels, but the energy density of oil is more than twice that of starch. We will use regulatory network and metabolic flux modeling to re-engineer the carbon allocation pathways to optimize triacylglyceride (TAG).**
3. **We will use comparative genomics of multiple *Lemnaceae* genome sequences, an extensive living collection of global accessions, and systems network analysis to characterize gene expression networks underpinning developmental and environmental responses to maximize bioenergy products while preserving rapid biomass accumulation. Nutrient deprivation and CO_2 irrigation will be used to enhance yield.**

Lemnaceae species (commonly called duckweeds) are the world's smallest aquatic flowering plants. Under optimal conditions, their rapid clonal growth rate can double the number of fronds in 30 hours and produce 64 grams of biomass per gram starting weight in a week, which is far beyond that of terrestrial crops such as corn (2.3 g/g /week), and unencumbered by secondary products such as lignin. *Lemnaceae* offer an attractive alternative to algae as biofuel feedstocks because of their robust growth in open ponds and the relative ease of harvesting dry material. Convenient metabolic labeling in culture makes Lemna a good system for pathway modeling and engineering, as nutrients are taken up from liquid growth media, and non-responsive stomata can utilize very high levels of atmospheric CO_2. Our goal is to divert substantial accumulated carbon from starch to oil metabolism in *Lemnaceae*, using resting fronds as the storage tissue.

The Martienssen and Lam labs have produced eight new reference quality *Lemnaceae* genome assemblies of *W. australiana*, *L. gibba*, *L. minor*, *L. turionifera*, and three allotriploid hybrid *L. japonica* (*L. minor* x *L. turionifera*) clones using Oxford Nanopore long reads and Hi-C contact maps. Comparisons of the chromosome-scale assemblies demonstrate a high degree of synteny across all 21 chromosomes within the *Lemna* genus, with only a single translocation consistently identified specifically in *L. turionifera*, while *W. australiana* has 20 chromosomes like *S. polyrhiza*, yet with significant architectural differences. In the Birchler Lab, antibodies against centromeric histone H3 were raised for four species to determine centromere organization and

identify centromeric repeat sequences across the duckweeds via CUT&RUN. The sequenced species encode between 14,000 and 20,000 genes, significantly fewer than terrestrial monocots such as rice and *Brachypodium*, and comparable to the algae *Chlamydomonas reinhardtii*. Methylome and small RNA sequencing revealed dramatic differences between the three genera consistent with known pathways of RNA directed DNA methylation. Analysis of missing and diverged orthogroups across the *Lemnaceae*, 10 other monocots and 7 non-monocots revealed variations that likely account for these and other traits including: reduced morphology, aquatic habitat, clonal reproduction, dormancy, high photosynthetic rate, and lipid production.

Key experiments in the Lam Lab have confirmed that natural genetic variation in *S. polyrhiza* leads to variable turion production. RNA-sequencing of two genotypes at the extremes of turion yield have identified turion-specific genes associated with dormancy, starch biosynthesis, and putative transcription factors that may be involved in the developmental transition. In addition, turion-specific expression of genes involved in lipid metabolism and oil biosynthesis were found in both *S. polyrhiza* as well as *L. turionifera*. Current work comparing transcript induction kinetics between different genotypes is underway to filter out the most promising candidate genes for functional validation. In collaboration with the Shanklin lab, a four- to six-fold increase in total TAG levels were found in turions of both duckweed species, consistent with predictions from RNA-seq, providing novel leads to target genes for directed modification of lipid content.

In previous work from the Shanklin and Schwender Labs, we expressed an Arabidopsis WRINKLED1, (WRI1) the master transcriptional activator of fatty acid synthesis in *Lemna japonica* line Lj8627. This resulted in <1% of DW of TAG along with large reduction in growth rate along and significant developmental abnormalities. Next, we constructed a CFP-N terminally tagged version of the Arabidopsis WRI1 under the control of inducible estradiol inducible XVE promoter, co-expressed with a sesame Oleosin 1 gene variant, (ROGUE Biosystems Design) in which its degradation signals had been minimized to optimize its TAG protective function, along with a very strong mammalian DGAT2 (CABBI Energy Center funding) to create Lj8627-33 (ODW) transgenics. Growth of OWD transgenics cultured in the presence of 100uM estradiol for four days resulted in the accumulation of 16.4% total fatty acid by DW compared to 5.2% in the parental line and 8.7% TAG per DW compared to 0.07 (124-fold increase). Subsequent detailed analyses confirmed strong TAG production and accumulation of very long chain fatty acids in TAG, but revealed that estradiol induction of WRI1 in ODW lines results in an allocation tradeoff of starch to TAG and a reduction in growth rate. Work is underway to evaluate the extent of lipid futile cycling and to model biomass synthesis using physiological, biochemical, and transcriptomic data to understand the growth rate reduction and inform strategies for mitigation.

Funding for this project is provided by the DOE Office of Biological & Environmental Research (DE-SC0018244).