

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}

¹Cold Spring Harbor Laboratory, Cold Spring Harbor, NY; ²Howard Hughes Medical Institute, Cold Spring Harbor, NY; ³University of Missouri, Columbia, MO; ⁴Rutgers University, New Brunswick, NJ; and ⁵Brookhaven 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 CO₂ 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 CO₂. 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 three new reference quality *Lemnaceae* genome assemblies of diploid *L. gibba*, diploid *W. australiana*, and allopolyploid *L. minor* clones using single molecule long-reads and Hi-C contact maps. Comparisons of the chromosome-scale assemblies reveal that the 21 chromosomes of *L. gibba* are highly colinear with the two subgenomes of *L. minor*, while *W. australiana* has 20 chromosomes like *S. polyrhiza*, yet with significant architectural differences. *S. polyrhiza*, *L. gibba*, and the B subgenome of *L. minor* all encode around 18,000 genes – significantly fewer than terrestrial monocots such as rice and *Brachypodium*, and comparable to the *Chlamydomonas reinhardtii*. Strikingly, *W. australiana*

has undergone a reduction to only about 14,000 coding genes. Methylome and small RNA sequencing revealed dramatic differences between the three genera consistent with known pathways of RNA directed DNA methylation. Orthologous gene analysis across the *Lemnaceae*, 10 other monocots and 7 non-monocots, revealed variations that likely account for these contrasts, as well as for reduced morphology, clonal reproduction, and aquatic habit.

The Birchler Lab has completed the construction of a transgene stacking system with alternating transformation vectors that enable iterative recombination into a locus determined by a previously integrated landing pad. This novel design is compatible with consecutive transformation of *Lemnaceae* under strictly clonal propagation. Lines bearing the landing pad have been regenerated and are being screened for single insertions. In addition, antibodies against centromeric histone H3 have been raised for four species to visualize centromere organization and identify centromeric repeat sequences across the duckweeds via CUT&RUN.

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 minor* line Lm8627. This resulted in <1% of DW of TAG along with large reduction in growth rate along and significant developmental abnormalities. We revisited our TAG producing strategy and constructed a CFP-N terminally tagged version of the Arabidopsis WRI1 under the control of inducible estradiol inducible XVE promoter. This was co-expressed with a Sesame Oleosin 1 gene variant, (ROGUE Biostems Design) in which its degradation signals had been minimized to optimize its TAG protective function and a very strong mammalian DGAT2 (CABBI Energy Center funding) with both genes under strong constitutive promoters to create Lm8627-33 transgenics. Growth of Lm8627-33 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. Thus, our inducible WRI1 strategy resulted in up to 124-fold increase in TAG in line Lm8627-33-6 four days after induction with little to no reduction in growth rate and no developmental abnormalities observed.

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