

175. Development of the Oleaginous Yeast *Rhodospiridium toruloides* as a New Model Organism for Systems-Level Analysis of Lipid Productivity

Dominic Pinel^{1,2}, Gina Geiselman³, Samuel T. Coradetti^{3,4}, Adam P. Arkin², Rachel B. Brem^{*3,4} (rbrem@buckinstitute.org) and Jeffrey M. Skerker^{*1,3}(skerker@berkeley.edu)

¹Energy Biosciences Institute, Berkeley, CA; ²Lawrence Berkeley National Laboratory, Berkeley, CA; ³University of California, Berkeley; ⁴Buck Institute for Research on Aging, Novato, CA

http://www.energybiosciencesinstitute.org/program_project/development-oleaginous-yeast-model-system

Project Goals: The most urgent need in bioenergy research is for commercial-scale production of advanced biofuels that can serve as "drop-in" replacements for petroleum-derived fuels. Toward this end, oleaginous yeasts hold great promise owing to their naturally evolved pathways for high fatty acid flux. We are carrying out massively parallel mapping of genotype to phenotype in the oleaginous yeast *Rhodospiridium toruloides*. Our focus is on the genes that underlie lipid production, plant feedstock hydrolysate tolerance, low-oxygen metabolism, and co-utilization of sugars in plant material. Objective 1 is to survey these attributes across wild *R. toruloides* strains. Objective 2 is to identify genes that underlie biofuel-relevant traits using insertional mutagenesis. Objective 3 is to map genes that underlie differences between wild *R. toruloides* isolates. Together, our findings will advance the understanding of lipid metabolism in *R. toruloides*, and enable engineering of an optimized biofuel production host.

The *R. toruloides* system. Upon nitrogen starvation, oleaginous yeast species accumulate large amounts of lipids, primarily triacylglycerols (TAGs), in an organelle called the lipid body; lipid content can accumulate in excess of 70% of total dry biomass. Harnessing TAG production directly for transesterification into fatty acid methyl esters (i.e., biodiesel), or ultimately, diversion of fatty acid flux in these species to synthesis of fatty acid esters and alkanes, is thus a promising strategy for large-scale production of biofuels. However, without a systems-level understanding of lipid metabolism in these systems, their use in the biofuel industry has been limited. We are focusing on *Rhodospiridium toruloides*, an oleaginous yeast that can natively metabolize sugars that are abundant in plant hydrolysates (glucose, xylose, arabinose, and cellobiose); has high de novo lipid productivity from glucose; grows and produces lipids in sugar concentrations equivalent to those in cellulosic plant hydrolysates; and does not itself consume alkanes, which are target biofuels. Although a full genome and annotation of one *R. toruloides* strain was recently reported, in only a handful of studies have *R. toruloides* genes been functionally characterized. To meet the challenges of strain engineering and development in industrial settings, we are carrying out a systems-level effort to identify the genes that govern *R. toruloides* growth and metabolic behaviors.

Lipid production and natural variation in *R. toruloides*. We collected six haploid *R. toruloides* isolates (three of each mating type) from diverse geographic locations and environments to sample genotypic and phenotypic diversity within this species. Sole-carbon-source tests revealed that all strains were able to metabolize glucose, arabinose, and cellobiose, and likewise all isolates except one grew on xylose. Thus, metabolism of sugars present in plant hydrolysates is thus far the rule rather than the exception among wild *R. toruloides* isolates, in striking contrast to *Saccharomycetes*. We used staining with the lipid dye BODIPY to qualitatively assess lipid production in these *R. toruloides* strains. After culturing in nitrogen starvation conditions, an environmental treatment known to induce TAG biosynthesis, each of our six wild *R. toruloides* isolates produced large amounts of lipid de novo from glucose. For one isolate, CBS 6570, we also confirmed lipid production after culturing on arabinose, xylose, and cellobiose. Detection of BODIPY staining with both a plate reader and a flow cytometer revealed a significant difference in

lipid accumulation between the *R. toruloides* strains CBS 349 and CBS 14. Thus, our data have already established that genetic background influences lipid productivity in *R. toruloides*.

Functional genomics in *R. toruloides*. To identify genes underlying lipid productivity and growth traits in *R. toruloides*, we have pioneered tools for transposon mutagenesis followed by competition of transposon mutants and sequencing (Tn-seq) in this system. We first established a protocol for *Agrobacterium tumefaciens* mediated transformation, to mutagenize the genomes of *R. toruloides* strain CBS 14 and CBS 349 with a T-DNA plasmid that confers resistance to the drug nourseothricin. We co-cultured *A. tumefaciens* strain EHA105 containing the T-DNA plasmid with *R. toruloides*, and selected random genomic insertions on 100 µg/mL nourseothricin and 300 µg/mL cefotaxime. We have achieved scale-up of this procedure to amass ~40,000 colonies from a given mutagenesis.

Next we established methods to map the genomic location of T-DNA insertions via Tn-seq. Here genomic DNA of a mixed pool of mutagenized *R. toruloides* is isolated and fragmented; adapters are ligated onto fragment ends; and T-DNA inserts are amplified with a primer recognizing the T-DNA sequence at one end and a primer recognizing the adapter on the other. Each amplicon thus contains a partial fragment of the T-DNA and a fragment of the genomic region into which it has inserted in a given clone, to be sequenced on the Illumina platform (150-bp single-end reads). Proof of concept sequencing experiments have yielded, from 1 million reads per library of DNA fragments from a pool of ~16,000 mutants cultured in rich medium, ~3200 genes in which we ascertained >20 Tn-seq reads per insertion site, i.e. 37% of the ~8600-gene *R. toruloides* genome. Read counts per insertion site agreed between replicates at $R^2 > 0.95$.

Future directions. In the coming year we will work to scale up construction and sequencing of large pools of barcoded *R. toruloides* mutants. We will develop selection schemes for oil production and plant hydrolysate tolerance, in which high-fitness mutants rise to high frequency in a mixed pool while low-fitness alleles drop out, with fitnesses measured by mutant abundance in barcoded Tn-seq data. As these experiments identify genes that modulate biofuel-relevant traits, we will develop a pipeline to validate their role in focused gene disruption experiments.

This work was supported by The Office of Biological and Environmental Research in the DOE Office of Science, and by the Energy Biosciences Institute at the University of California, Berkeley.