Developing the yeast *Kluyveromyces marxianus* as a thermotolerant bioproduction host

Ian Wheeldon,[1*] (wheeldon@ucr.edu), Xuye Lang[1], Mengwan Li[1], Pamela Besada-Lombana[2], Danielle Bever-Sneary[2], Tami McTaggart[2], Nancy Da Silva[2], Dipankar Baisya[1], Stefano Lenoardi[1]

[1]University of California, Riverside; and [2]University of California, Irvine

Project Goals: This systems and synthetic biology project seeks to understand and engineer the native stress tolerance phenotypes of the yeast *Kluyveromyces marxianus* with the goal of developing a new synthetic biology chassis for fuel and chemical production.

The non-conventional yeast *Kluyveromyces marxianus* is one of the fastest growing eukaryotes, is thermotolerant to temperatures upward of 50°C, and has the capacity to assimilate a wide range of C_5 and C_6 sugars. These traits make *K. marxianus* an attractive host for the industrial production of biochemicals. However, in comparison to the common yeast synthetic biology chassis, *Saccharomyces cerevisiae*, there is a clear lack of functional genomic and synthetic biology tools. In the initial stages of this project, we were successful in creating a new synthetic biology toolbox for pathway engineering, including standardized promoters with variable expression [1], CRISPR-mediated markerless integration [2], and strategies for rapid pathway refactoring. We have also demonstrated the utility of these tools by engineering enhanced production of triacetic acid lactone (TAL) and 2-phenolethanol. Here we report on the success of our pathway engineering efforts as well as describe our current efforts in creating a genome-wide CRISPR-Cas9 system for functional genomics.

With respect to pathway engineering, we have enhanced production of the native metabolite, 2-phenolethanol, by refactoring the Shikimate pathway [2]. A 27-member strain library that varied the expression of Shikimate pathway genes ARO4, ARO7, and PHA2 demonstrated that flux to 2-phenolethanol was limited by ARO4 expression. Alleviating tyrosine inhibition to ARO4 and increasing expression level resulted in significant improvements in titer. Building on the refactored Shikimate pathway, a titer of nearly 2 g/L of 2-phenolethanol was achieved by eliminating acetyltransferase activity to limit conversion to 2-phenolethyl acetate, overexpressing the key Ehrlich pathway gene ARO10, and by culturing in a small-scale bioreactor with fed-batch operation. After initial studies demonstrated the promise of *K. marxianus* for high-level TAL synthesis [3], we engineered the central carbon pathways using our high-efficiency PolII-based CRISPR system. We redirected carbon flux via native gene knockouts and heterologous gene integrations, demonstrated the importance of metal ion availability, and evaluated enzyme colocalization in the mitochondria to increase efficiency. The introduction of a heterologous carbon saving pathway increased TAL synthesis by 2-4 fold, reaching titers in excess of 1.3 g/L.

In collaboration with the DNA Synthesis team at JGI, we have created a genome-wide CRISPR-Cas9 functional genomics screening tool for *K. marxianus*. We first designed eight single guide RNAs (sgRNAs) for every gene in the genome; four guides targeted the coding region, while another four guides targeted the promoter region. In total, more than 35,000 guides were designed, synthesized, and cloned into a vector for sgRNA expression. Building on our past work in the non-conventional yeast *Yarrowia lipolytica* [4], we are conducting a genome-wide

cutting efficiency experiment to quantify the ability of each guide in the library to make a double stranded break in the genome. To do so, we have eliminated the ability of *K. marxianus* to repair DNA through non-homologous end-joining, the dominant repair mechanism in this and other non-conventional yeast. These genome-wide cutting efficiency experiments are currently ongoing, but we have developed the analysis workflow to successfully translate the outcomes of these experiments into a machine learning-based algorithm to predict guide cutting efficiency. Our method of quantifying Cas9 cutting efficiency across the genome generates tens of thousands of data points relating CRISPR guide sequence, genome structure (*i.e.*, chromatin structure via ATAC-seq and nucleosome occupancy via MNase-seq), and Cas9 activity. From the Yarrowia data set, our deep-learning convolutional neural network (CNN) can predict CRISPR guide sequences that result in high Cas9 activity at a desired locus. Validation experiments are on-going, but initial results suggest that we can predict sgRNAs for Cas9 that lead to a strong cutting efficiency 95% of the time (based on 19 successful guides out of 20 tested). Current work is extending this analysis to *K. marxianus* Cas9 experiments and to Cpf1-based cutting in Yarrowia.

References

1. Lang, X., et al., *Developing a broad-range promoter set for metabolic engineering in the thermotolerant yeast Kluyveromyces marxianus.* Metabolic Engineering Communications, 2020. **11**: p. e00145.
2. Li, M., et al., *CRISPR-mediated multigene integration enables Shikimate pathway refactoring for enhanced 2-phenylethanol biosynthesis in Kluyveromyces marxianus.* Biotechnology for Biofuels, 2021. **14**(1): p. 3.
3. McTaggart, T.L., et al., *Synthesis of polyketides from low cost substrates by the thermotolerant yeast Kluyveromyces marxianus.* Biotechnology and Bioengineering, 2019. **116**(7): p. 1721-1730.
4. Schwartz, C., et al., *Validating genome-wide CRISPR-Cas9 function improves screening in the oleaginous yeast Yarrowia lipolytica.* Metabolic Engineering, 2019. **55**: p. 102-110.

This work was supported by DOE DE-SC0019093