

170. Estimating Intracellular Fluxes Using Relative Expression and Phenotype Datasets (REPPS)

Matthew Long^{1*} (mrlong2@wisc.edu) and Jennifer L. Reed¹

¹Dept. of Chemical and Biological Engineering, Univ. of Wisconsin-Madison

Project Goals: Cyanobacteria offer a promising route for directly converting solar energy and CO₂ into biofuels. The objectives of this research are to integrate modeling and experimental approaches to guide development of a butanol producing cyanobacterium, *Synechococcus* sp. PCC 7002. New computational approaches will be developed to facilitate these efforts which will (1) design experiments and analyze their results, and (2) identify genetic engineering strategies for improving butanol production in *S. 7002*. Experiments will subsequently be performed to construct and analyze *Synechococcus* 7002 strains engineered for butanol production. The developed approaches will be systematically applied to suggest genetic engineering strategies for improving production of a variety of biofuels in five other microorganisms. This research will support the U.S. Department of Energy's mission for developing renewable ways of producing advanced biofuels.

Renewable sources of transportation fuels are needed to reduce the amount of oil used to satisfy transportation energy needs in the U.S. and to alleviate our dependence on foreign sources of oil. Microbes can be used to produce a wide variety of liquid biofuels including: ethanol, butanol, isobutanol, isoprene, hydrogen, and alkanes. Cyanobacteria offer an alternative route for converting solar energy and CO₂ into biofuels, without the need for using lignocellulosic biomass as an intermediate. The biofuel production capabilities of microbes can be improved through metabolic engineering, where metabolic and regulatory processes are adjusted using targeted genetic manipulations. Traditionally, metabolic engineering strategies are found through manual inspection of metabolic pathways, where enzymes involved in biosynthesis are overexpressed or added, competing pathways are eliminated, and the performance of resulting strains are evaluated. However, such approaches cannot predict the effects that these changes will have on other parts of metabolism, and generally will not suggest alterations to more distant pathways.

Computational models of cellular metabolic and regulatory networks can be used to guide and accelerate these metabolic engineering efforts by integrating and analyzing experimental data, and identifying genetic manipulations that would increase product yields. In the process of developing metabolically engineered strains, genetic manipulations proposed by computational strain design algorithms depend on the metabolic state of parental strains. One such algorithm is RELATCH (for relative change), which has been shown to accurately predict the effects of gene deletions and environmental shifts on metabolic fluxes [1]. However, RELATCH requires knowledge of both gene expression and fluxes in the parental strain, including intracellular flux measurements (e.g., ¹³C metabolic flux analysis), to predict fluxes in knockout mutants. While gene expression is easily measured, intracellular flux measurements are harder to generate and are not widely available, particularly for cyanobacteria. As such, alternative methods for obtaining knowledge of fluxes through metabolism are needed to evaluate and improve engineered strains.

A number of experimental measurements can be made to evaluate the metabolic state of a cell, such as enzyme activity, gene expression, metabolite concentrations, protein concentrations, and cellular uptake and secretion rates. The integrated analysis of these various datasets can be used to help estimate metabolic fluxes in cells and identify potential bottlenecks in biofuel production. Here we have developed a novel constraint-based modeling method for calculating the flux distribution and enzyme contributions in a parental strain using experimental data from multiple gene deletion strains. This method, Relative

Expression and Phenotypes for Parental Strain estimation (REPPS), incorporates multiple modules which can utilize growth rates, extracellular fluxes, and gene expression data from multiple knockout reference strains and the parental strain to predict intracellular fluxes. We have further evaluated the importance of both the abundance and type of data used by REPPS on the resulting intracellular flux estimates for both *Escherichia coli* and *Saccharomyces cerevisiae*. By integrating multiple datasets, we are able to more accurately estimate the parental strain flux distribution, yielding as much as a ~44% improvement compared to existing approaches (e.g., pFBA). The improved parental strain flux prediction from REPPS has then been used with RELATCH to accurately predict fluxes in new mutant strains with greater accuracy. Future work will be to validate and apply these methods to the identification and alteration of fluxes in cyanobacterial strains (e.g., *Synechococcus* 7002 and *Synechocystis* sp. PCC 6803) engineered for biofuel production.

References:

1. Kim, J. and J.L. Reed, RELATCH: relative optimality in metabolic networks explains robust metabolic and regulatory responses to perturbations. *Genome Biol*, 2012. 13(9): p. R78.

This work was supported by the Office of Science (BER), U.S. Department of Energy (DE- SC0008103).