Title: Optimal Experimental Design (OED) of Biological Systems

Authors: Puhua Niu (niupuhua.123@tamu.edu),[3*] Maria J. Soto[2], Byung-Jun Yoon[1,3], Edward R. Dougherty[3,] **Francis J. Alexander**[1,] Ian Blaby[2], and Xiaoning Qian[1,3],

Institutions: [1]Brookhaven National Laboratory, Upton, NY; [2]Lawrence Berkeley National Laboratory, Berkeley, CA; [3]Texas A&M University, College Station, TX.

Project Goals: The project's overall goal is to develop optimal experimental campaigns to achieve a particular objective, namely metabolite yield alteration. The optimal experiments will be designed by quantifying the cost of uncertainty in the current predictive model—a transcriptional regulatory network (TRN) model that regulates metabolism—and selecting the experiments that are expected to maximally reduce the model uncertainty that affects the attainment of the aforementioned objective. This approach will serve as a proof of principle, demonstrating the significant potential of computationally guided biology in areas directly relevant to BER's missions.

Abstract Text: There has been extensive research on *in silico* modeling and prediction of genome-scale metabolic behavior, mostly focusing on mutant strain design with metabolic reaction network modeling [1]. To further model transcriptional regulations, these metabolic network models are also integrated with genetic regulatory relationships involving transcription factors (TFs) that may regulate metabolic reactions. Transcription regulation is often integrated via "transcriptional regulatory constraints" with various heuristics for flux-balance analysis (FBA) of metabolic networks. However, many of these computational tools were often only validated for selected model organisms with curated data and network models, despite access to high-throughput technologies allowing genome-scale engineering.

A major achievement in our project during the past year is the development of TRIMER (Transcription Regulation Integrated with MEtabolic Regulation), a genome-scale computational pipeline for integrative modeling and analysis of TF-regulated metabolism. Specifically, a Bayesian network (BN) is employed in TRIMER, instead of local TF-gene conditional probabilities or transcriptional regulatory constraints, thereby aiming at effectively capturing the global transcriptional regulatory relationships that may affect metabolism. Through this BN, the influence of transcription regulation (and its changes) on metabolic behavior under different conditions can be predicted more accurately via more flexible conditional probability inference, by linking transcription regulation to metabolism based on prior knowledge on TF-gene-reaction interactions. Details of our TRIMER package and comprehensive results can be found in [2, 3].

To experimentally validate the predictions generated by TRIMER, we assayed 163 TF deletants vs. wildtype for biomass and metabolite levels, which served to confirm the performance of the tool.

While *E. coli* is a well-understood model organism, with extensive prior knowledge on both gene regulation and metabolic reaction pathways as well as significant amount of accumulated expression data, one critical challenge to overcome when applying TRIMER to less-studied organisms—especially, in the context of optimal experimental design (OED)—is to understand the model uncertainty or sensitivity due to incomplete knowledge and/or noisy data. Building upon TRIMER, we further investigated how model uncertainty in the TRN may affect the metabolic flux prediction and the TF KO experiment design. With a quantified uncertainty, we aim to further optimize the outcome for OED under uncertainty in an efficient manner (i.e., fewer experiments and less guesswork), whereby optimization is achieved by optimally (most favorably) improving the model or the microbial system represented by the model relevant to the objective (i.e., maximizing the metabolite yield). We here present the recent research progress on model uncertainty analyses as well as the planned research directions based on our analyses.

References/Publications

[1] Amit Varma and Bernhard Ø Palsson. Metabolic flux balancing: Basic concepts, scientific and practical use.Bio/technology, 12(10):994–998,1994.

[2] Puhua Niu, Maria J. Soto, Byung-Jun Yoon, Edward R. Dougherty, Francis J. Alexander, Ian Blaby, Xiaoning Qian. TRIMER: Transcription Regulation Integrated with MEtabolic Regulation, *iScience*, 24(11):103218, 2021.

[3] Puhua Niu, Maria J. Soto, Byung-Jun Yoon, Edward R. Dougherty, Francis J. Alexander, Ian Blaby, Xiaoning Qian. Predicting condition-dependent metabolite yield with TRIMER, *STAR Protocols*, under review (minor revision).

Funding Statement: The materials presented in this paper are based upon the work supported by the U.S. Department of Energy, Office of Science, Office of Biological and Environmental Research under contract number DE-SC0012704