

## A media optimization pipeline for improving TRY with machine learning

Matthew Incha<sup>1,2,3\*</sup> (incha@lbl.gov), Tijana Radivojević<sup>2,3</sup>, Vincent Blay Roger<sup>2,3</sup>, Matthias Schmidt<sup>2,3,4</sup>, Hector Garcia Martin<sup>2,3</sup>, & Jay D. Keasling<sup>1,2,3,5</sup>

<sup>1</sup>Department of Plant and Microbial Biology, University of California, Berkeley, CA 94720, 14 USA; <sup>2</sup>Joint BioEnergy Institute, 5885 Hollis Street, Emeryville, CA 94608, USA; <sup>3</sup>Biological Systems & Engineering Division, Lawrence Berkeley National Laboratory, Berkeley, 10 CA 94720, USA; <sup>4</sup>Institute of Applied Microbiology (iAMB), Aachen Biology and Biotechnology (ABBt), RWTH 12 Aachen University, Aachen, Germany; <sup>5</sup>Department of Chemical and Biomolecular Engineering, University of California, Berkeley, CA 20 94720, USA

### Project Goal:

Improve the biosynthetic performance of an engineered microorganism via machine-learning guided optimization of media components. Using a robotics platform, many unique media formulations can be tested. Titers achieved in each formulation will inform a Bayesian model over 5-10 iterations which will then design new improved media formulations.

### Abstract:

Bioprocess optimization is often nonintuitive and requires many iterations of the canonical design, build, test, learn (DBTL) cycle. One method for improving biosynthetic performance is by altering the media conditions with design of experiment (DoE) tools. Using robotics and colorimetric screening, these processes can be conducted quickly and effectively. To this end, we chose the red compound, flaviolin, as our first product for this pipeline. A genome integrated production strain with a constitutive promoter driving flaviolin biosynthesis (*rppA*) was developed. This strain demonstrated highly reproducible titers in minimal media and was used as the reporter for media performance. A robotics platform was further developed to rapidly iterate over diverse media formulations and assay production via simple spectrophotometry. Following platform development 22 media formulations based on MOPS minimal medium were designed using Latin Hypercube Sampling (LHS)<sup>1</sup>. Using the LHS results as training data, a machine learning model developed in JBEI, the Automated Recommendation Tool (ART)<sup>2</sup>, designed new formulations. Following 5 DBTL cycles, ART consistently designed media with 2x titer improvement over MOPS. Following strain construction, these improved media formulations were developed in less than 2 months, illustrating the cost & time effectiveness of this machine learning media optimization pipeline.

### References

- <sup>1</sup>McKay, M. D., R. J. Beckman, and W. J. Conover. "A Comparison of Three Methods for Selecting Values of Input Variables in the Analysis of Output from a Computer Code." *Technometrics* 21, no. 2 (1979): 239–45. <https://doi.org/10.2307/1268522>.
- <sup>2</sup>Radivojević, T., Costello, Z., Workman, K. et al. A machine learning Automated Recommendation Tool for synthetic biology. *Nat Commun* 11, 4879 (2020). <https://doi.org/10.1038/s41467-020-18008-4>

### Funding Statement

Office of Science, Office of Biological and Environmental Research, of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231.