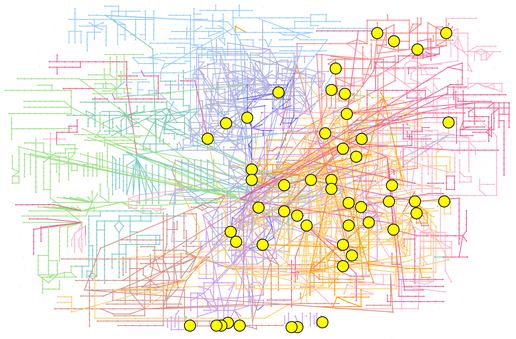
Absolute quantification of key cellular metabolites in bioprocessing samples using machine learning

Luke S. Ferro[1], Ashley Rettberg[2], Jack Howland[1], Gabriel Asher[1],Timothy Kassis[1], Ana S. H. Costa[1], Jeff Pruyne[1], Steven Hooper[1], Joshua D. Lauterbach[1], Mimoun Cadosch Delmar[1], J.M. Geremia[1], Jennifer M. Campbell[1]

1. Matterworks, Inc.
2. Eli Lilly and Company

Introduction

Metabolomics provides insight into biological systems through the quantification of cellular metabolites. The traditional approach to absolute quantification with mass spectrometry is based on supervised peak integration and isotope dilution. However, the need for expert intervention and matched heavy isotopologues limits the scalability of this technique to large sample numbers and metabolite libraries. At ASMS 2022, we presented a novel machine learning model, Pyxis™, which enables absolute quantification across a large diversity of metabolic pathways. Here, we use that model to quantify metabolite concentrations in real-world biological samples. The model demonstrated excellent performance across a diversity of bioprocessing samples, setting a new standard for rapid metabolite quantification. We believe that our model will be broadly useful for applications requiring high-throughput monitoring of intracellular dynamics.



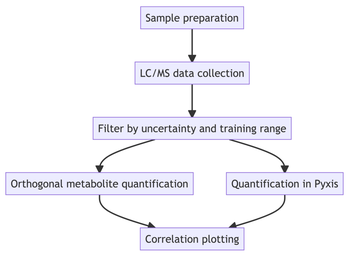
Pyxis provides quantitation across a large diversity of metabolic pathways.

Glycolysis
Amino acids
TCA cycle
Nucleotides
Vitamins

Bioprocessing dataset

CHO cells producing mAB
Goal to optimize titer, maintain CQA

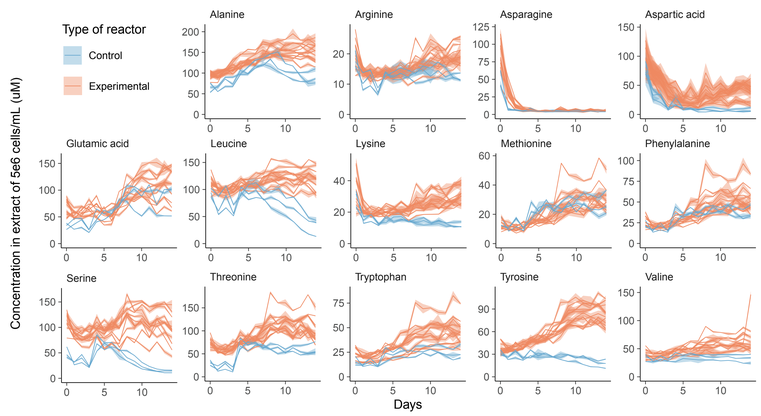
3 control, 11 experimental bioreactors
Altering feed, temperature, DO, etc.



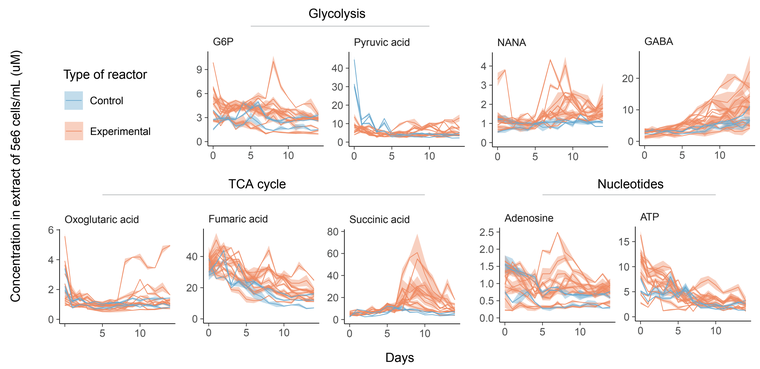
Time series analysis of bioprocessing samples

CHO cells grown in bioreactors were sampled at different time points. Metabolites were extracted from the cells and absolute quantification was performed using Pyxis. Here, 'experimental' reactors are the user trying out different new cell culture conditions.

Amino acid metabolism

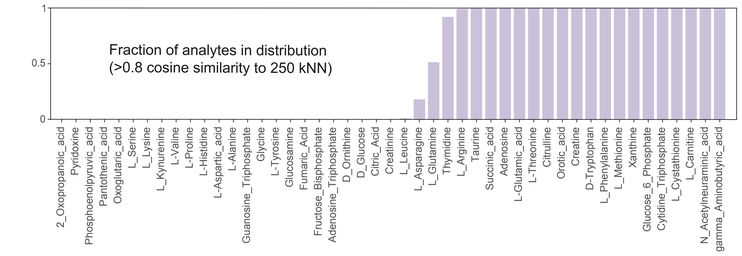


Glycolysis, TCA cycle, nucleotide metabolism

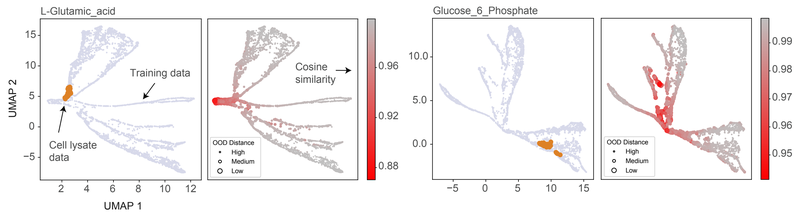


Quantifying model uncertainty on real samples

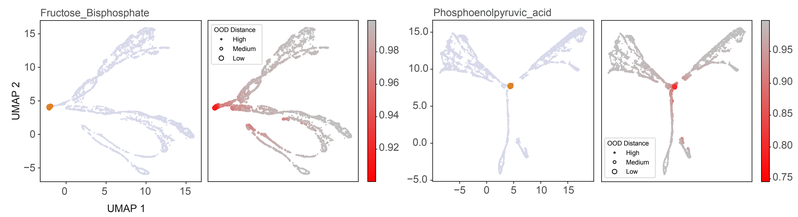
Model uncertainty was quantified by calculating an average cosine similarity score between the embeddings of the sample and the embeddings of the nearest 250 neighbors from the training data. By defining a threshold distance value, this enabled us to estimate which analytes were reported with the most confidence.



Analytes in the distribution of the training data

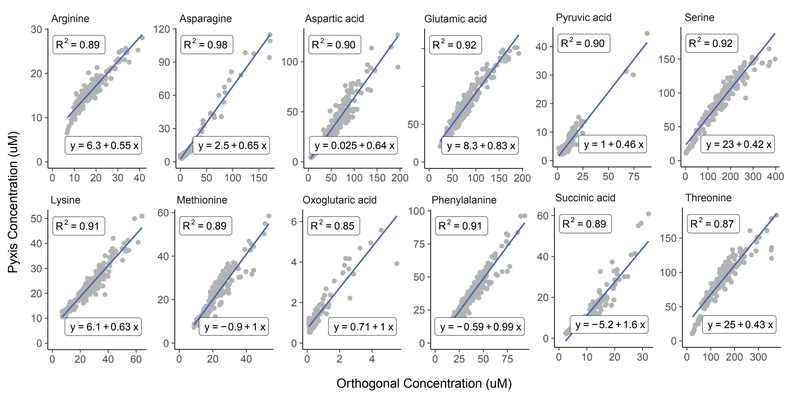


Analytes out-of-distribution of the training data



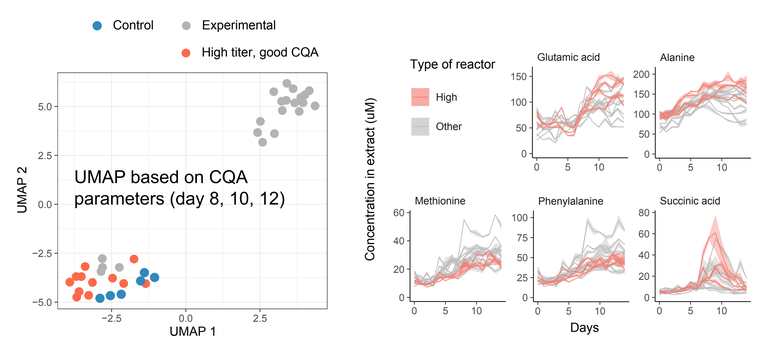
Correlation plots show the accuracy of Pyxis™

Accuracy of Pyxis data from cell lysates was determined by comparing to orthogonal data quantified using matched heavy isotoplogues.



Candidates for additional follow-up

The goal of the experiment was to identify culture conditions which increased titer while maintaining other critical quality attributes of the mAB. We identified which reactors met those criteria and looked for metabolites with unique trends in those reactors.



Roadmap for future product development

2022 Pyxis prototype

- 48 metabolites
- 2 orders of magnitude
- Analyzed diverse customer data sets

Pyxis

- 160 metabolites
- 3 orders of magnitude
- 10x larger training set
- Improved model architecture
- Outlier detection
- Advanced visualizations

Learn more at matterworksbio.com

Matterworks™

All authors are employees and shareholders of Matterworks, with the exception of Dr. Rettberg who is an employee of Eli Lilly and Company