Predicting metabolic and proteomic variations in *B. subtilis* as a response to stress and protein overproduction

Juan D. Tibocha-Bonilla, Eammon Riley, Cristal Zuñiga, Colton Lloyd, Kevin Rychel, Kit Pogliano, Karsten Zengler

The 1st International BioDesign Research Conference

UC San Diego

Zengler Lab
at University of California, San Diego

Background

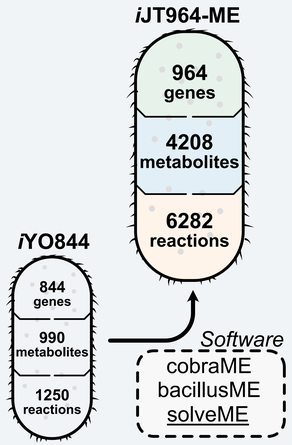
Bacillus subtilis is a well-characterized microorganism and a model for the study of Gram-positive bacteria. The bacterium also exhibits the capability to produce proteins at high densities and yields, which has made it valuable for industrial bioproduction. With a new generation of metabolic models, or ME-models, it is possible to study proteome allocation and its response to biologically and industrially relevant conditions. Here, we describe the reconstruction and validation of the first ME-model of *B. subtilis*, *i*JT964-ME. This model achieved higher performance scores on the prediction of gene essentiality as compared to the M-model *i*YO844. We successfully validated the model by integrating physiological and omics data associated with gene expression responses to ethanol and salt stress. The model further allowed to identify the mechanism by which tryptophan synthesis is upregulated under ethanol stress. Further, we employed *i*JT964-ME to predict amylase production rates at two different growth conditions. We analyzed these flux distributions and identified key metabolic pathways that permitted the increase in amylase production. All in all, we illustrate the potential of *i*JT964-ME in the study of proteomic response to stress and the optimization of protein production.

Model and gene essentiality

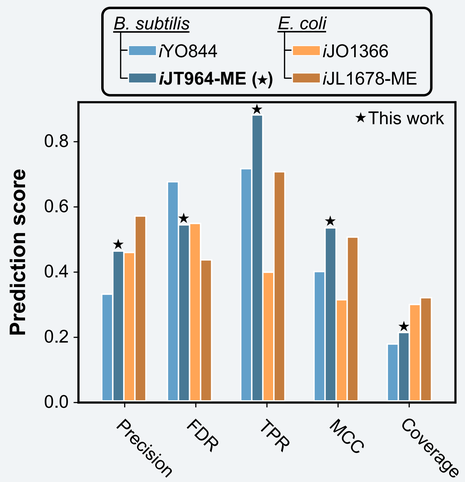
To reconstruct iJT964-ME we adapted the available metabolic modeling packages COBRAme, COBRApy, and ECOLIme. COBRAme that were originally built for E. coli. The resulting B. subtilis-specific pipeline expanded the existing M-model iYO844 with non-metabolic reactions, including translation, transcription, tRNA charging, and post-translational modification. The final ME-model (iJT964-ME) contains 964 genes, 6282 reactions, and 4208 metabolites.

The addition of gene expression reactions into the network of B. subtilis resulted in a 14% increase in genome coverage, with 32% of them being essential as predicted by iJT964-ME. The extensive manual curation performed for iJT964-ME significantly increased the prediction scores of gene essentiality.

Model properties

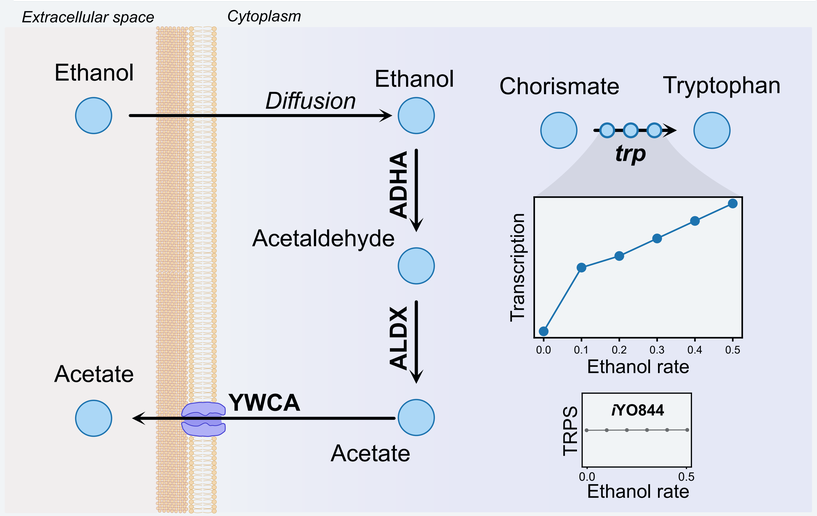


Gene essentiality



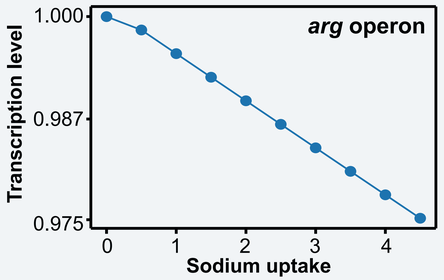
In some cases, prediction scores surpass those predicted for E. coli M- and ME-models. Interestingly, just a 14% increase in gene content allowed iJT964-ME to predict essentiality with increases of 34% in the Matthews Correlation Coefficient (MCC) and 40% in Precision.

Predicting response to stress



Strategies adopted by living organisms to overcome stress are widely studied but not fully understood. In this section, we studied the predictive capability of iJT964-ME on salt and ethanol stress. Previous reports highlighted a group of significantly co-regulated genes associated with an outstanding upregulation of tryptophan synthesis under ethanol stress (4% v/v). We used these findings to assess the capability of iJT964-ME of identifying metabolic shifts of B. subtilis under ethanol stress.

First, we modeled ethanol stress by constraining experimentally observed uptake ethanol rates in the model. Our simulations show that the higher tryptophan demand can be caused solely by an increase in the demand for ethanol processing and acetate secretion enzymes. According to the resulting flux distributions, ethanol was converted to acetate through alcohol dehydrogenase (adhA) and aldehyde dehydrogenase (aldX), which was then secreted through a sodium-dependent acetate symporter (ywcA). The translation of ywcA caused approximately 65% of the total increment in tryptophan synthesis.



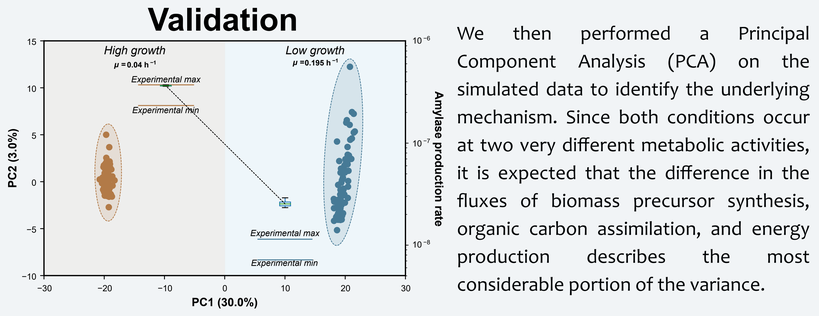
Second, we focused on the understanding of possible metabolic mechanisms to overcome osmotic stress. We used transcriptomics data of B. subtilis growing under salt stress to evaluate the predicted flux distributions at the genome-scale. Our model accurately predicts the arg operon's decrease when the excess sodium enters the cell, as reported previously.

Interestingly, the regulation of 60% of differentially expressed genes in salt stress was accurately captured by iJT964-ME. The model accurately captures the response of genes associated with main metabolic pathways, such as amino acid synthesis, ribosome formation, and nucleotide synthesis. These pathways are accurately captured since their activity is related to the organism's core metabolic response to stress.

Analyzing amylase overproduction

First, we tested whether the new features of iJT964-ME allowed it to capture amylase secretion shifts at two different growth conditions. While under low growth, the model slightly overestimates the amylase secretion rate, while iJT964-ME predicts secretion rates within the reported experimental ranges at a high growth rate .

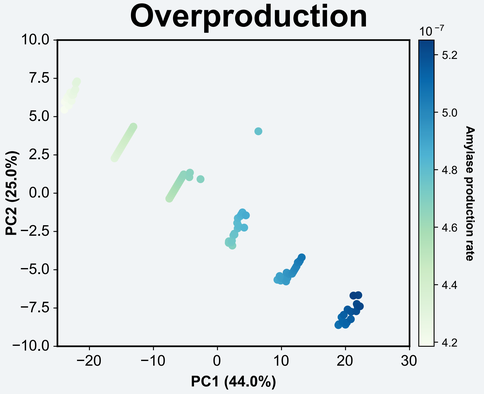
Validation



We then performed a Principal Component Analysis (PCA) on the simulated data to identify the underlying mechanism. Since both conditions occur at two very different metabolic activities, it is expected that the difference in the fluxes of biomass precursor synthesis, organic carbon assimilation, and energy production describes the most considerable portion of the variance.

Overproduction

We tested what groups of reactions would significantly drive the overexpression of amyE by fixing the growth rate at the lowest of the high growth conditions and forcing amyE overproduction until reaching the highest secretion rate. The two largest components described a strikingly higher portion of the variance (69%). Both components mostly consist of amino acid synthesis reactions.



Conclusion

The iJT964-ME model represents a significant advancement in the metabolic modeling of B.subtilis. Its broad scope and ability to capture expression changes have improved gene essentiality predictions, shed light on recent hypotheses relating amino acid metabolism and stress, and explored the capacity to secrete industrially relevant proteins. This model can serve as the basis for unraveling further questions about metabolism and has the potential to be the foundation on which to optimize heterologous protein expression in this important model organism and cell factory.