COMP150CSB: Computational Systems Biology
Tufts University, Department of Computer Science
Prof. Soha Hassoun
Rev1

Homework 3
Analyzing Metabolic Networks

Assignment Overview

Constraint-based modeling offers a powerful way for analyzing metabolic networks. In this homework, you will utilize the COBRA Toolbox v3 [1] to analyze the *i*JO1366 *E. coli* metabolic network [2]. The objectives of this assignment are:

- Understand the representation of metabolic networks using the standard metabolic model
- Use COBRA toolbox for model and constraint checking
- Use COBRA toolbox to perform FBA for a specific objective function
- Use COBRA toolbox to perform FVA
- Use COBRA toolbox to determine gene essentiality

Assignment Details

Setup the COBRA toolbox

Required software – see [1] page 13
Installation steps – LP and MILP solver glpk is installed by default, as well as DQQ, MINOS, PDCO, and QPNG. You might need to install CPLEX to speed FVA.

Run through the COBRA Tutorial, using reference [1]

- Importing a reconstruction model, step 4
- Set simulation constraints, step 18
- Model consistency checking, steps 16, 17, 19, 20, 25-27
- Flux Balance Analysis (FBA), step 21-23
- Integrating metabolic data, steps 31-32 (new research area, contrast with step 39 for transcriptomic and proteomic data)
- Adding biological constraints to a flux balance model, steps 40-45
- Flux variability analysis (FVA), steps 52-53
- (optional) Thermodynamically constrain a metabolic model, steps 78-81

Familiarize yourself with the *i*JO1366 *E. coli* model, reference [2]

- Read section on "Updating the biomass composition and growth requirements" in the pdf supplementary material
- Read section on "Prediction of Gene Essentiality"
- Read the model into MATLAB and examine its various components

Model Analysis

Create a script, analyzeModel.m, that performs the following, suppressing all outputs except the ones shown below. The output of your script should be saved in a file, HW3output.txt. Do not add any blank lines between the required outputs described below.

- Importing a reconstruction model and reporting on various aspects of the model

Load the iJO1366.mat model using `readCbModel`
Write fprintf statements to report on the number of reactions and number of metabolites, and other aspects of the model, generating a summary as follows (without tabs at the beginning of the lines; cut and paste from here to allow us to use automatic grading), where XXX indicates the results obtained from querying the model:

> Summary: Model
> Reactions = XXX, Metabolites = XXX
> Index of wild-type biomass reaction (Biomass 'BIOMASS_Ec_iJO1366_WT_53p95M) = XXX
> Index of glucose uptake (D-Glucose exchange) = XXX
> Number of reversible reactions based on both upper and lower bounds of non-zero = XXX
> Number of exchange reactions (with exchange as part of their reaction name) = XXX

- Compare maximum growth in aerobic vs anaerobic conditions

Set the maximum glucose uptake rate to 8 mmol gDW^{-1} hr^{-1} (millimoles per gram dry cell weight per hour). For aerobic conditions, set the maximum uptake rate to 18.5 mmol gDW^{-1} hr^{-1}. Perform FBA to report the maximum growth rate with and without oxygen. Do not change the ATP maintenance requirement. As described in class, uptake into the cell should be negative!
Write an fprtinf statement that summarizes your results:

> Summary: Aerobic vs anaerobic growth rate in per hr units
> Assume 8 mmol/gDW.hr of glucose uptake
> Assume aerobic conditions with 18.5 mmol/gDW.hr of oxygen uptake
> WT biomass growth rate aerobic vs anaerobic = XXX and XXX

- Compare maximum growth based on glucose vs succinate vs L-lactate

Evaluate the growth rate in aerobic conditions for glucose uptake rate to 8 mmol gDW^{-1} hr^{-1} against uptake rates of succinate, and of L-lactate, each in turn being set to an uptake rate of 16 mmol gDW^{-1} hr^{-1}
Write an fprtinf statement that summarizes your results:

> Summary: Compare maximum growth rates in per hr units using alternative mediums
> Compare glucose vs succinate vs L-lactate, respectively:
> WT biomass growth rate = XXX and XXX and XXX

- Identify essential genes in *E. coli*.

A gene is considered essential if its deletion (suppression) reduces optimal growth rate to less than 5% of that of the wild type strain. Using this definition, identify the essential genes in the iJO1366 model. Assume the cell is growing in aerobic conditions and using glucose as a growth medium. The output file EssentialGenesOutput should list the model genes (as listed in the model), followed by TRUE or FALSE indicating essentiality or lack of it, respectively. This file should be text, and tab delimited.

> Summary: Identifying gene essentiality
> Setting minimum growth rate = XXX
> Number of essential genes = XXX
> Alphabetical listing of number of genes are saved in file EssentialGenesOutput

- Flux variability analysis (FVA)

Perform FVA assuming aerobic conditions, with maximum glucose uptake rate to 8 mmol gDW^{-1} hr^{-1}
Your output should consist of a tab delimited file, where each line has the name of the rxnName, the minimum flux, and the maximum flux, separated by tabs. The order of the listing should follow that in the model rxnNames file.

> Summary: Flux Variability Analysis
> Number of blocked reactions (with zero upper and lower fluxes) = XXX
> Listing of flux ranges are saved in file FluxVariabilityOutput

- Comparison with ref [2]

In a file, compareWithPublishedModel.txt, comment on your results as compared to those reported for the original model. For each of the analysis steps above, comment how your results matched or did not match the published model, and explain why.

Files To Turn in

analyzeModel.m
HW3output.txt
FluxVariabilityOutput
EssentialGenesOutput
compareWithPublishedModel.txt

References

[1] Heirendt, L., Arreckx, S., Pfau, T., Mendoza, S. N., Richelle, A., Heinken, A., ... & Magnusdottir, S. (2017). Creation and analysis of biochemical constraint-based models: the COBRA Toolbox v3. 0. *arXiv preprint arXiv:1710.04038*.

[2] Orth, J. D., Conrad, T. M., Na, J., Lerman, J. A., Nam, H., Feist, A. M., & Palsson, B. Ø. (2011). A comprehensive genome-scale reconstruction of Escherichia coli metabolism—2011. *Molecular systems biology*, *7*(1), 535.