

Systems Biology: Metabolic Modeling Laboratory Handout

Instructor: R. Mahadevan, Department of Chemical Engineering and Applied Chemistry, University of Toronto.

Introduction:

The part of this laboratory is to teach how constraint-based models of bacterial metabolism are implemented using state of the art modeling tools. In order to achieve this objective, we will be utilizing the COBRA toolbox that was developed by Prof. Bernhard Palsson's group at the University of California, San Diego. The details of this toolbox and its use are presented in the Nature Protocol paper published in 2007 and is attached to the handout(1). The toolbox has already been installed at your workstations and you can access it via MATLAB and navigating to the COBRA toolbox folder. In the directory C:\Documents and Settings\cbw-guest\Desktop\cbw\krishna\, you should find the SBML file for the genome-scale model of *Escherichia coli*, iJR904, published a few years back(3). For this lab, we will use the glpk software which is access through glpkmex.

Please ensure that the glpkmex, and the SBML toolbox are in the MATLAB path and then initialize the COBRA toolbox by typing

```
>>initCobraToolbox
```

1) The goal of the first part of the laboratory is to load the SBML model of *E. coli* iJR904 and to evaluate its contents. You should be able to use the following commands to access the model (detailed instructions are also present in the accompanying paper).

```
>> model=readCbModel('iJR904');
```

```
>>model
```

- Can you identify the number of genes and metabolites and reactions ?
- What is the rank of the matrix ? (hint: use the full(model.S) with the rank cmd)
- What is the degree of freedom or the extent to which the model is underdetermined ?
- Use the spy command to plot the S matrix ? Can you identify the highly connected metabolites?

2) Solve the *E. coli* model using the FBA approach where the biomass maximization is the objective function. Use the following commands to solve the FBA model.

```
>> solution1=optimizeCbModel(model);  
>> printFluxVector(model,solution1.x,true,true);
```

The growth rate should 0.539 hr^{-1}

Calculate the biomass yield for aerobic growth during glucose oxidation using these results ?

- What would be the biomass yield if glucose uptake was 1 mmol/gdw hr rather than 6 mmol/gdw hr ? Does the biomass yield decrease linearly with glucose uptake ? If so or not, why ?

Use the command `model2=changeRxnBounds(model,'EX_glc(e)',[-1],1)`; to change the bound and to check the model2 bounds: `[model2.rxns,cellstr(num2str(model2.lb))]`

3) Change objective function for the linear programming problem to that of maximizing the rate of ethanol production using the following command and solve the model:

```
model3=changeObjective(model,{'BiomassEcoli','EX_etoH(e)'},[0,1]);
```

a) Calculate the maximum ethanol yield on a molar basis ? Is this under aerobic or anaerobic conditions ? What is the calculate growth rate ?

b) What is the ethanol yield under anaerobic conditions if biomass maximization is the objective ? How does this compare to the maximum in a ? What is aerobic ethanol yield ?

```
model4=changeRxnBounds(model,'EX_o2(e)',[0],1);
```

c) Is ethanol growth coupled under aerobic and anaerobic conditions ?

4) One of the important limitations of the Flux Balance Analysis /CBMs is the presence of alternate optimal solutions which are a result of the redundancy in the metabolic pathways. One way of analyzing this redundancy is to use the Flux Variability Analysis (2) . Using the COBRA toolbox, perform Flux Variability Analysis at the optimal growth rate under aerobic and anaerobic conditions with the following command:

```
[minfluxaerobic,maxfluxaerobic]=fluxVariability(model,100);
```

a) Calculate the sum of ranges for all fluxes for both conditions and comment on the differences or similarities in the ranges.

b) For aerobic growth, calculate sum of the ranges for the fluxes at 50% of the optimal growth rate ? How does this compare with case a ?

c) Does the extent of redundancy as described by the flexibility in pathway flux differ across conditions ?

5) Use the COBRA toolbox to perform genome wide single deletion analysis using the FBA growth maximization objective with the following command.

```
grRatioFBA=singleGeneDeletion(model,'FBA');
```

a) Perform the genome wide deletion analysis for aerobic and anaerobic conditions and compare the number of essential reactions for the two cases ?

6) Take Home part: For both aerobic and anaerobic conditions, plot the absolute values of the growth rate when a reaction is deleted and its flux value under the WT condition ? Do you expect to see a correlation ? You will have to create a new matlab function in a manner similar to the `singleGeneDeletion` function.

Reference List

1. **Becker, S. A., A. M. Feist, M. L. Mo, G. Hannum, B. O. Palsson, and M. J. Herrgard.** 2007. Quantitative prediction of cellular metabolism with constraint-based models: the COBRA Toolbox. *Nature Protocols* **2**:727-738.
2. **Mahadevan, R. and C. H. Schilling.** 2003. The effects of alternate optimal solutions in constraint-based genome-scale metabolic models. *Metab Eng* **5**:264-276.
3. **Reed, J. L., T. D. Vo, C. H. Schilling, and B. Palsson.** 2003. *Escherichia coli* iJR904: An expanded genome-scale model of *E. coli* K-12. *Genome Biol.* **4**:R54.1-R54.12.