Practical Bioinformatics

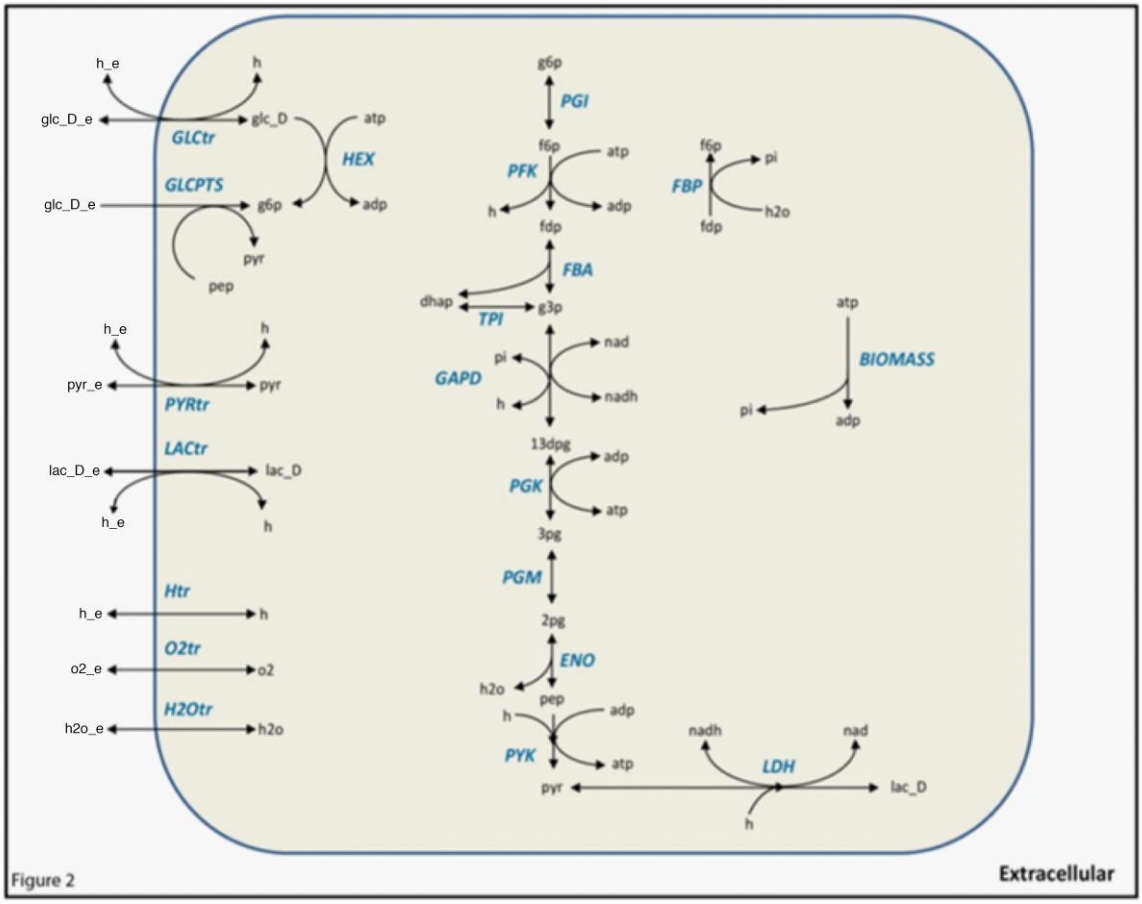
Wagner Section

Exercise Block 2. Analyzing simple biological networks

Metabolic networks – We will use the following two metabolic networks in this section:

1. glycolysis.xml - This metabolic network consists of reactions involved in the glycolysis pathway. The default objective function is the maximization of ATP production (abbreviated as atp in the metabolic model). Figure 2 below serves as a visual representation of the network.

2. e_coli_core.xml - This metabolism contains what is called the core E. coli metabolic network, as proposed by the laboratory of Bernhard Palsson. The files Ecoli_core_figure.pdf and Ecoli_core_figure.jpg are added to this folder as a visual representation of the core E. coli model.



Exercise 2.1. Load the glycolysis model into python. How many metabolites and reactions does it include?

Note that this model uses short-hand reaction identifiers (IDs) that are not the same as the reactions' names in figure 2.

How many metabolites and reactions does the model have? Write a script that lists all reactions and their associated information. It should look something like this:

```
name : id : reaction : lb : ub
GLCtr : R1 : glc_D_e + h_e <=> glc_D + h : -1000.0 : 1000.0
GLCPTS : R2 : glc_D_e + pep --> g6p + pyr : 0.0 : 1000.0
HEX : R3 : atp + glc_D <=> adp + g6p : -1000.0 : 1000.0
```

Provide a list of the metabolites as well. A model's reactions can be found in `model.reactions`, its metabolites in `model.metabolites`. In the COBRApy reference, have a look at the properties of the reaction and metabolite objects, respectively.

Once you have gathered the information about the reactions, perform FBA to maximize ATP production. You do not have to change the `model.objective`, but take a look at it and try to explain why having this reaction as the objective maximizes ATP production. Provide a list of all active reactions. How many are there?

Exercise 2.2. (OPTIONAL) GLCPTS (glucose phosphotransferase system) is a reaction that imports glucose into the cell. Use figure 2 to answer if GLCPTS is an essential or a nonessential reaction? Why or why not? Then use COBRApy to check your answer.

You can find the effects of performing single reaction deletions in the following way:

```
from cobra.flux_analysis import single_reaction_deletion
del_results = single_reaction_deletion(model)
```

(see COBRApy reference for further description of the function)

Exercise 2.3. Use figure 2 shown above and the model glycolysis.xml to identify the role of the reaction catalyzed by lactate dehydrogenase (LDH in the model), beyond its enzymatic function.

Hint: Is LDH an essential reaction? Why or why not? Which cofactors are involved in the LDH reaction? How many other reactions use these cofactors? After performing FBA, look at `model.metabolites.nadh.summary()`, which may help you answer the question.

Exercise 2.4. (OPTIONAL) Flux Variability Analysis at different fractions of the maximum rate of biomass production.

The following command will perform FVA for all reactions in a model when the biomass growth flux is constrained to its maximum value:

```
fva_result = cobra.flux_analysis.flux_variability_analysis(model,
fraction_of_optimum=1)
```

The argument `fraction_of_optimum` has the default value of `1`, which means that FVA is performed subject to the constraint that the objective function (e.g., maximal biomass growth) attains the maximal (optimal) value. This constraint set to the objective function can be loosened by setting `fraction_of_optimum` to any number in the range of 0 to 1.

Consider the network's flux variability when constraining biomass growth to different fractions of the optimum. A simple measure of a ***network's flux variability*** *is the sum of the flux variabilities of all reactions in the network. A reaction's flux variability is the range of possible values the flux through said reaction can take.*
(Note that this measure is highly dependent on the number of reactions in a network, the network itself, and the reaction bounds, and can therefore only be used to compare FVA results at different fractions of optimum.)

Does the network's flux variability change when relaxing the constraint of optimal growth? Why or why not?

Exercise 2.5. In this and in the following exercises, we will work with the core metabolism of *E. coli.* Before performing any calculations, it is always good to take a look at the model. Open the model and list the number of metabolites, genes and reactions. How many external reactions does this model include? Which metabolites can be imported into the cell?

To find the external reactions in the list `model.reactions`, it may be useful to know that in this model the reaction IDs of all external reactions start with 'EX_'. For example, the identifier for the external reaction of oxygen is `EX_o2_e`. Alternatively, external (or exchange) reactions are also stored in the list `model.exchanges`. Try both to see if your results are the same. Also, keep in mind that if a metabolite is taken up by the cell, its associated external reaction will show a negative flux.

Exercise 2.6. Compute the biomass production of the network for different uptake rates of glucose. Plot biomass against glucose uptake rates. Explain the relationship between the two quantities that you observe.

An important concept for this section is that of a ***(growth-)limiting nutrient****. A nutrient is limiting if a reduction in its uptake rate leads to a reduction in growth.*

Start by computing FBA with the default value (10 mmol/gDW/hr of glucose).
Look at the results, then compute the rates with different glucose uptake values ranging from 0 to 50. To perform this task, you will need to change the flux boundaries.
Remember this can be done with:

```
model.reactions.get_by_id('REACTION ID').lower_bound = ANY VALUE
```

Look at the details of the model's exchanges to find the correct reaction ID. Make a plot to visualize your results. Describe what you see. Is glucose growth-limiting over the entire concentration range? If not, which other nutrient(s) might be growth-limiting?

Here is an example of how to plot in python using matplotlib:

```
import matplotlib.pyplot as plt
import numpy as np

x = np.linspace(0, 20, 21)
# x is an array of 21 numbers between 0 and 20
y = x**2
plt.plot(x, y, marker='o', linestyle='none')
plt.xlabel("values of x")
plt.ylabel("values of y")
plt.title("A Simple Plot")
plt.show()
```

Exercise 2.7. Let's take a closer look to understand what happens when the glucose uptake changes. Write a script that will perform FBA for different glucose uptake values between 0 and 50 mmol/gDW/hr. Store the flux value of the external reactions with the IDs EX_o2_e, EX_co2_e, EX_ac_e and EX_etoh_e. Plot these fluxes against the glucose uptake rate and try to explain the pattern you observe.

You can add more than one set of data to the same plot:

```
x = np.linspace(0, 20, 21)
y = x**2
z = x*10
plt.plot(x, y, label='y values', marker='o', linestyle='none')
plt.plot(x, z, label='z values', marker='o', linestyle='none')
plt.legend()
plt.show()
```

Exercise 2.8. Compute the biomass growth rate for acetate as the sole carbon source with an acetate uptake flux of 10 mmol/gDW/hr. Is it different from the biomass growth rate on glucose with the same uptake rate? Why or why not?

Exercise 2.9. Are there any other metabolites that could be used as carbon sources for the core metabolism of *E. coli*?

For a metabolite to be a potential carbon source, it must be able to be imported into the cell and must possess at least one carbon atom. The first condition is met by all metabolites with an associated external reaction. Checking if a given metabolite possesses at least one carbon atom can be a little more troublesome. Luckily, each `metabolite` in this model has a property `metabolite.elements`, which stores the number of atoms of each element in the metabolite in a dictionary, and a property `metabolite.formula`, which stores its chemical formula as a string. To find the metabolites involved in a reaction, you can use the reaction property `reaction.metabolites`.

Since this model is small enough, you may print the `elements` or the `formula` of all external metabolites to find the ones that fulfill the conditions for being a carbon source by hand. If you feel confident, write a python script to do this automatically.