

## Practical Bioinformatics

Wagner Section

### Exercise 4.

**In this problem we are going to explore the metabolism of *E. coli* more thoroughly, and study the effects of employing different metabolites as carbon source.**

The analysis will be done with a specific model of *E. coli*'s metabolism called iJO1366. The following paragraphs contain some guidelines of analyses you can do, but we encourage you to come up with questions on your own.

**Exercise 4.1. How many metabolites, genes and reactions does the model of *E. coli* iJO1366 include? Maximize biomass production. How many reactions are active (carry a flux different from zero)? Take a look at the active external reactions. What is *E. coli* consuming and excreting?**

**Exercise 4.2. Find all essential reactions in the *E. coli* metabolic network for growth in a minimal glucose-containing environment. What fractions of the total reactions are essential?**

**Exercise 4.3. Compute the fraction of reactions that are essential for *E. coli* biomass synthesis in an acetate minimal environment. Is it higher or lower than the number of essential reactions for biomass synthesis on glucose? Which reactions are essential in both environments?**

**Exercise 4.4. Modify the python program written for the previous exercise to find the reactions that are essential for growth on acetate, but not on glucose.**

**Exercise 4.5. Use the iJO1366\_reactionInfo.csv file to compute how many essential reactions on glucose occur in each metabolic pathway (or subsystem). Which pathway is the one with the most essential reactions?**

Understanding the metabolic pathway in which essential reactions occur can provide useful biological information, because pathways with many essential reactions are

arguably more important for survival in a specific chemical environment.

In exercise 3.2 we accessed the subsystem information directly from the .xml file. The model we are working with in this exercise, iJO1366.xml, does not include that information. All the information regarding the model is given in a separate file (iJO1366\_reactionInfo.csv). Explore the file. Where is the information about the subsystem given?

The following commands can be used to open the file iJO1366\_reactionInfo.csv, store the information of every line in a list (info\_lines) and close the file.

```
#Open the file with the reactions info
filo = open('iJO1366_reactionInfo.csv')
info_lines = filo.readlines()
filo.close()
```

**Exercise 4.6.** In Exercise 4.2, you identified essential reactions for growth in a glucose minimal environment. Now compute these reactions for the rich environment specified by the file rich\_environment.txt. For a gut bacterium such as *E. coli*, which of these environments may be more realistic? Why? Is the number of essential reactions for growth in these environments different? If so, is it smaller or larger, and why?

**Exercise 4.7.** Search for the metabolites that have the potential to be carbon sources and calculate the maximum growth achievable on each.

Is it the same in aerobic and anaerobic conditions? Why or why not? Can you provide an example of a carbon source for which *coli* is viable only aerobically and a carbon source for which *coli* is viable only anaerobically?

**Exercise 4.8.** Focus on the aerobic condition. Plot the distribution of maximum biomass growth on the different carbon sources. Does biomass growth depend on the total number of carbon atoms in a carbon source? Why or why not? Does the number of active reactions change? Why or why not?

**Exercise 4.9.** Does the number of essential reactions vary across different carbon sources? (Plot a histogram of this number.) Why or why not?

