

BT5420 - COMPUTATIONAL SYSTEMS BIOLOGY

Assignment 4

Submitted by Sahana (BE17B038)

Note – The spreadsheet of Synthetic double lethal gene pairs for question 1 and the codes for solutions and fastSL_dg.m are attached in the zip folder.

Question 1

Identify synthetic double lethals(genes) for the given metabolic model of *Helicobacter pylori* in the following carbon sources:

1. Only glucose as carbon source
2. Only galactose as the carbon source
3. Both glucose and galactose as the carbon source.

Is there a change in the synthetic lethal (SL) with change in the carbon source? Comment on the same with biological significance.

Solution –

Helicobacter Pylori's default nutrition source are amino acids. Minimal medium required by the model to generate required biomass constituents was calculated, indicating the requirement of eight amino acids, six of which correspond to essential human amino acids.

The wild-type biomass flux of the model is 0.6928, and the number of synthetic lethal pairs is 35. This was identified by the algorithm – “fastSL_dg.m” with a 5% as a cut-off percentage for lethality.

Case 1: When Glucose is (additionally) provided as a carbon source

Here, the amount of nutrition to the model has been increased. Therefore, we'd expect a higher growth rate/ biomass flux since the food source has been increased.

As expected, the model showed as increased biomass flux of 0.9351.

In this case, 241 synthetic lethal pairs were observed. The rise in the number of SL pairs is due to the following reason –

There is more room for the model to survive in this situation, i.e., we can easily delete 241 genes and still expect the model to survive (have a biomass flux value $\geq 95\%$ of the wild-type flux). Whereas, this would not have been possible in the wildtype case. What are SL pairs in this scenario, can as well be single lethal genes (Genes that when knocked out displays a lethal phenotype). Hence, the model has an increased chance of survival with glucose as carbon source.

Case 2: When Galactose is (additionally) provided as a carbon source

Similar to the previous scenario, since, the carbon source in the growth medium has increased, the biomass flux increased to 0.9351.

In this case, 259 synthetic lethal pairs were observed. The rise in the number of synthetic lethals from wildtype to this case holds the same reason as the previous scenario.

There is more room for the model to survive in this situation, i.e., we can easily delete 259 genes and still expect the model to survive (have a biomass flux value $\geq 95\%$ of the wild-type flux). Whereas, this would not have been possible in the wildtype case. What are SL pairs in this scenario, can as well be single lethal genes (Genes that when knocked out displays a lethal phenotype). Hence, the model has an increased chance of survival with glucose as carbon source.

However, we can also observe that the number of synthetic lethals has increased from glucose as a carbon source to galactose as a carbon source. This is possibly due to environment specific factors that alter various pathways, thereby requiring certain genes for important pathways in the former case, while they can be held as functional redundancies (synthetic lethals) in the latter case. Certain cellular redundancies are unique to specific environments while some occur irrespective of the growth medium.

Case 3: When both Glucose and Galactose are provided as carbon source

Since the amount of carbon source for the model has increased, it clearly means that our bacterium is now getting a lot more nutrition and hence the biomass flux further increases to a value of 1.0554.

The number of synthetic lethals in this case has reduced from the previous two cases to 102 pairs.

This is because, when both glucose and galactose are added in the medium, the possible synthetic lethals would be a part of the intersection of the above two scenarios, since the model has to survive in both the environments simultaneously while also maintaining flux through vital pathways.

Conclusion –

Condition	Number of synthetic double lethal pairs
1. Wild-type	35
2. Glucose	241
3. Galactose	259
Common between case 2 and case 3	225
Glucose media specific double lethal	16

Galactose media specific double lethal	34
4. Glucose and Galactose (Is a subset of the intersection of “Common between 2 and 3”)	102

In general, it has been observed that, in the wild type, where the bacterium has to prepare every other metabolite from the scratch, i.e., from the amino acids, multiple pathways need to be active simultaneously in order for survival of the bacterium. Therefore, the number of synthetic lethal pairs is lesser. When additional nutrition source was added, specific to its environment, there was a larger room for the model to grow and survive, hence the number of synthetic lethal pairs increased. When two media were combined to form the third carbon source, the bacteria had to maintain those pathways that were active in both the media individually. Hence, the number of synthetic lethal pairs decreased from the previous 2 scenarios, but was still greater than the wild-type case.

Question 2

Identify a minimal ‘reactome’ for *E. coli* iAF1260. A minimal reactome must be such that the removal of ANY reaction from the network results in a growth rate that is lower than 5% of wild-type growth rate. How can you make your code more efficient?

Solution

Based on the question, it was necessary to not let the biomass flux of the model get lower than 5% of wild-type objective value at any point of time.

Therefore, the wild-type solution was first identified. To this model, the lower bound of the biomass-flux-associated reaction was set to 95% of wild-type objective value.

After this was done, to identify the set of reaction classes, “pFBA” was run for the model. pFBA classifies genes and reactions for finding an optimal solution.

The outputs of pFBA contains a structure called “RxnClasses” which states the categories of reactions of this model. The respective categories are –

- Essential reactions – Those that are very important for the survival of the model and removal of even a single reaction from this class would result in lethality phenotype.
- Blocked reactions – Those reactions that do not carry any flux under the given conditions.
- Zero-flux reactions – Those reactions that carry zero flux under given conditions.
- MLE reactions – Metabolically less efficient reactions which might be responsible for reducing the biomass flux, by diverting it to the production of secondary metabolites.
- ELE reactions – Enzymatically less efficient reactions which are a part of long alternate pathways that thereby increase the number of enzymes involved.

- pFBA optimal reactions – Other reactions that help in attaining the optimal solution of the model.

From this set, Blocked, Zero-flux, MLE and ELE reactions are all redundant, i.e., their absence will still keep the cell away from lethality condition. Hence, for obtaining a minimal reactome, we can delete these reactions.

Deletion - Deletion of a reaction was done by setting the lower bound and the upper bound of a reaction to zero.

This resulted in 405 reactions with the same biomass flux as that observed in wild type conditions.

How can it be improved?

It was noted that upon running the obtained minimal model in `singleRxnDeletion.m`, certain reactions when deleted did not give a lethal phenotype, implying that the model at hand is not fully minimal yet. But the only class of reactions that can be checked is `pFBAOpt_Rxns`, as Essential reactions cannot be tampered with because it might lead to lethality.

In an iterative fashion, the result of deletion of each reaction in the class of `pFBAOpt_Rxns` were carefully studied. Once a reaction is deleted, `optimizeCbModel(minmodel)` was run to check if the biomass flux was still greater than 95% of wild-type objective value. The solver status in standardised form was also checked.

The solver status tells us the kind of solution that the model offers. For example –

stat - Solver status in standardized form:

- 0 - Infeasible problem
- 1 - Optimal solution
- 2 - Unbounded solution
- 3 - Almost optimal solution
- -1 - Some other problem (time limit, numerical problem etc)

Hence, upon deleting reactions from `pFBAOpt_Rxns`, it was important to make sure the solution was still 'Optimal'. If the biomass flux condition and solver status condition was satisfied, the reaction deletion was considered feasible and was hence deleted from the minimum model obtained earlier.

This way, 5 more reactions were deleted, and the **size of the minimal reactome thus became 400.**

Verification-

The obtained minimum model was run on `singleRxnDeletion.m` and it was observed that the deletion of any of these 400 reactions, each resulted in lethality phenotype. This proves that the final model constitutes a minimal reactome.

Output –

```
>> hw4_2

Wild type flux of biomass production is 0.736701
Single gene deletion analysis in progress ...
100% [.....]
Single reaction deletion analysis in progress ...
100% [.....]

pFBA done for identifying various types of reactions in the model

Results -
There are 223 blocked reactions.
Results -
There are 627 Zero-flux reactions.
Results -
There are 75 enzymatically less-efficient reactions.
Results -
There are 1047 metabolically less-efficient reactions.

After deleting blocked reactions, the biomass flux is 0.736701
.
After deleting blocked and zero-flux reactions, the biomass flux is 0.736701
.
After deleting blocked, zero-flux and MLE reactions, the biomass flux is 0.736701
.
After deleting blocked, zero-flux, MLE and ELE reactions, the biomass flux is 0.736701
```

```
.
Identified minimal reactome has 400 reactions

The minimal model is -
  mets: {1668x1 cell}
  metNames: {1668x1 cell}
  metFormulas: {1668x1 cell}
  metCharge: {1668x1 double}
  genes: {1261x1 cell}
  rxnGeneMat: {2382x1261 double}
  grRules: {2382x1 cell}
  rxns: {2382x1 cell}
  rxnNames: {2382x1 cell}
  subSystems: {2382x1 cell}
    S: {1668x2382 double}
    lb: {2382x1 double}
    ub: {2382x1 double}
    b: {1668x1 double}
    c: {2382x1 double}
    rev: {2382x1 int64}
  description: 'iAF1260'
```