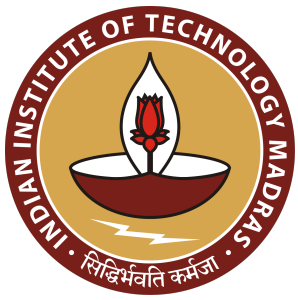
BT5240 Computational Systems Biology

Computational Systems Biology - Assignment 4

N Sowmya Manojna | BE17B007
Department of Biotechnology,
Indian Institute of Technology, Madras



1 Double lethal Genes

Cobra `version 3.1` (Cobra 2020), MATLAB `2020a` and the `gurobi` solver was used in this problem.

1.1 Algorithm

- The `iIT341` model was loaded using the `readCbModel` command and was stored as `model`.
- Exchange reactions of the `model` were analyzed and the reaction number of D-Glucose and Galactose were found.
- The initial lower bound of glucose and galactose uptake was zero in the model and the main source of carbon for the organism was found to be amino acids.
- A copy of the model named `glu_model` was created and the lower bound of glucose exchange reaction was set to -10 i.e. `model.rxns(96) = -10;`.
- Another copy of the model named `gal_model` was created and the lower bound of galactose exchange reaction was set to -10 i.e. `model.rxns(95) = -10;`.
- Similarly the `glu_gal_model`, where there is both glucose and galactose uptake was set to -10, i.e. `model.rxns(95:96) = -10;` was created.
- The Reaction-Gene Matrix, `rxnGeneMat`, was built for all the four models using `buildRxnGeneMat`.
- Double deletion lethal genes were identified using `fastSL_dg`, with a lethality cutoff of 0.05.
- Single lethal genes of WT, glucose based, galactose based and glucose + galactose based were named `sgd, glu_sgd, gal_sgd` and `glu_gal_sgd` respectively and the double deletion lethal pairs of WT, glucose based, galactose based and glucose + galactose based were named `dgd, glu_dgd, gal_dgd` and `glu_gal_dgd` respectively.

1.2 Results

The number of single and double lethal genes are as follows:

Model	Number of single lethal genes	Number of double lethal gene pairs
WT	193	35
Glucose	177	241
Galactose	175	259
Glucose & Galactose	175	102

The exact gene pairs has been saved as `hw4a_data.mat` and can accessed in the `data` folder.

1.3 Biological Significance

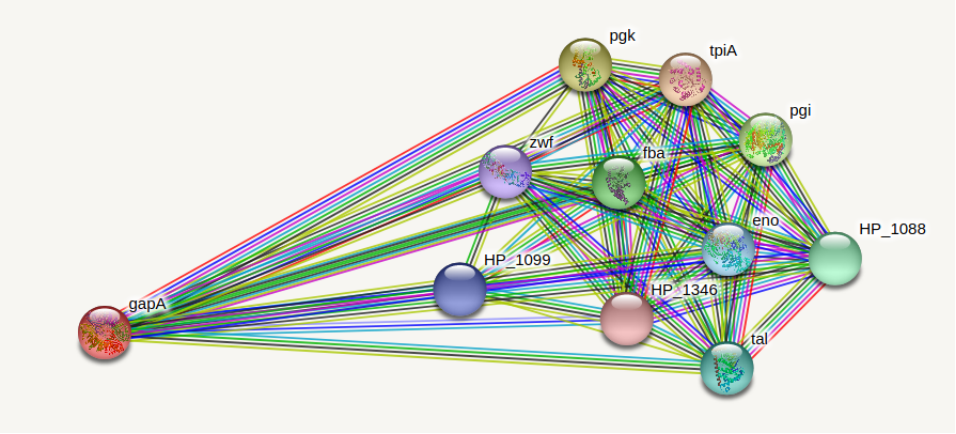
The variations is the single gene deletions and double gene deletions were analyzed and the following results were obtained:

1.3.1 Single Gene Deletions

- The most of the lethal genes that were present in `WT` (and absent in `glu_model`) were found to participate in vitamin production pathways (particularly biotin, dethiobiotin and riboflavin synthesis), fatty acid synthesis pathways and amino acids biosynthesis. All these pathways are known to be very significant pathways in a minimal medium comprising of only amino acids (as in case of `WT`).
- The variations between the `glu_model` and `gal_model` single lethal genes were found to be due to the genes: `HP0646` and `HP0360`. `HP0646` encodes for UTP–glucose-1-phosphate uridylyltransferase, while `HP0360` encodes for UDP-glucose 4-epimerase. This functional specificity of the proteins coded by the two single lethal genes should be the main reason for them to be single gene lethals in `glu_model`.
- The single lethal genes in `glu_gal_model` matches with that of `gal_model` and can be explained as the intersection of all single gene lethal sets considered.

1.3.2 Double gene lethal pairs

- Most double lethal gene pairs were found to interact with each other or have the same function. The gene pair interactions were analyzed using STRING.
- For example, the genes - `HP1346`, `HP0921`(gapA) were found to interact as co-occurring genes that are essential in glyceraldehyde 3- phosphate dehydrogenation.



- And the genes - `HP0875`, `HP1461` that code for Catalase and Cytochrome c551 peroxidase respectively, were found to share the same function of reduction of hydrogen peroxide to water, thereby reducing oxidative damage in the cell. Hence, the simultaneous deletion of both genes, resulted in a lethal phenotype.
- Most of the additional double lethal genes under galactose medium were found to comprise of genes coding for proteins involved in galactose metabolism coupled with proteins that play a key role in redox balancing. Hence, the combined deletion of both genes resulted in double gene deletion lethals in the `gal_model`.
- And as expected, the double lethal gene pairs in the `glu_gal_model` were found to be the intersection of the pairs in `glu_model` and `gal_model`.

2 Minimal Reactome

Cobra `version 3.1` (Cobra 2020), MATLAB `2020a` and the `gurobi` solver was used in this problem.

2.1 Algorithm

- The `iAF1260` model was loaded and was stored as `model`.
- Objective function value was optimized and the solution was stored in `sol`.
- In order to identify blocked reactions, zero flux reactions, MLE reactions and ELE reactions, a `Parsimonious FBA` was run on the `model` and the reaction classes were stored in `RxnClasses`.
- The pruning of reactions was carried out in four stages. They are as follows:
  - **Stage 1:** All blocked reactions, zero flux reactions, MLE reactions and ELE reactions were removed from the model using `removeRxns` function.
  - **Stage 2:** The `pFBAOpt_Rxns` are iterated over and each reaction is removed iteratively from the model. This model was named as `min_model`. The model was solved using `optimizeCbModel` and all reactions carrying zero flux were removed.
  - **Stage 3:** The models with least number of reactions from the previous step are selected and `singleRxnDeletion` was used to identify important genes. Reactions that have `hasEffect` flag value as `0` were removed.
  - **Stage 4:** The models with least number of reactions from the previous step are selected and `singleRxnDeletion` was used to identify all reaction deletions that resulted in a growth rate ratio greater than 5%. These reactions were then iteratively deleted.
- The models with the least number of reactions were then analyzed.

2.2 Result

The results obtained following each step is as follows:

Step	Reaction size during beginning	Reaction size during end
Step 1	2382	410
Step 2	410	402
Step 3	402	400
Step 4	400	399*

* Upon analysis of the minimal reactome obtained, it was found that the last reaction that was deleted was that of `ATPM`. Hence, the size of the minimal reactome obtained is only 400.

The complete workspace of this code is saved as `hw4b_data.mat` in the `data` folder.