SEARCHING NUTRIENT PREFERENCE ON NETWORK TOPOLOGY

(not a paper draft)

1 E. COLI NUTRIENT PREFERENCES

It is found that E. coli has preferences at the time to consume nutrients in a complex medium. For instance, even when several carbon sources are available E. coli might consume only one nutrient till its depletion, and only later, start to consume others. The sequence of consumption is reproducible.

Figures 1 and 2 show two examples.

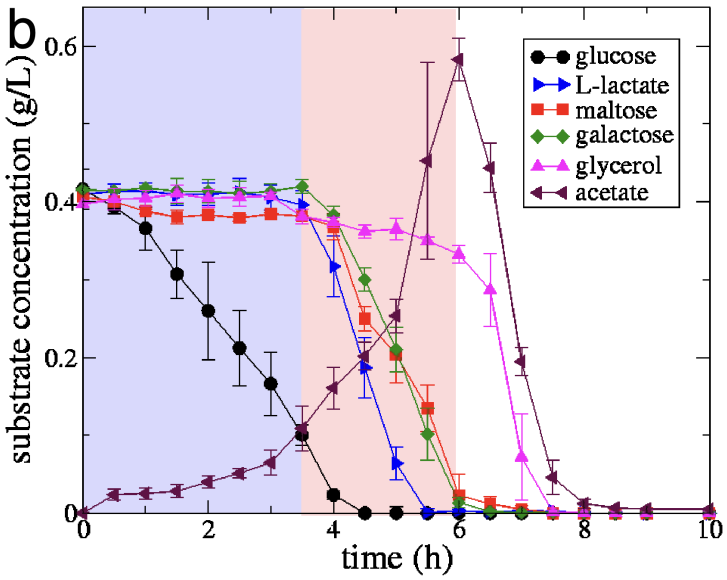

Figure 1: Time evolution of the concentration of several nutrients (plus acetate) present in the medium of an E. coli batch culture. Taken without request from [1].

2 TOPOLOGICAL PREFERENCE

Because we are very brave, we will try to find if the topology of the network do contain biases which are conherent with the nutrient expeimental preferences. The cell can control its metabolic configuration using several regulatory stratergies. We can not explore must of this regulatory space, which includes alosteric, equilibrium, expression and localization regulation, among others. But, we can try to explore one of the consequences of such regulation events, a flux knockout (KO). Given a

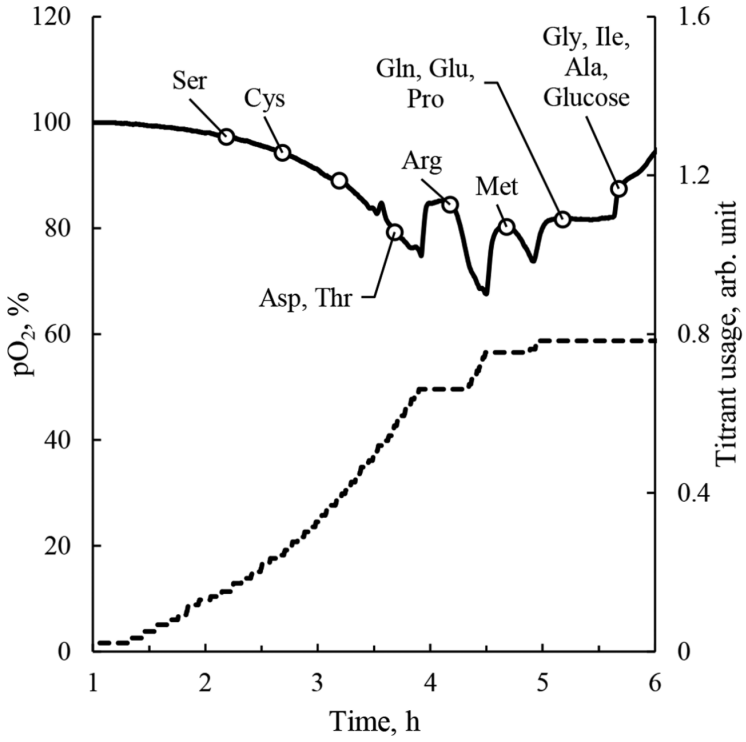

Figure 2: Ignore the lines and the vertical axis. The empty circles mark the moment when the labeled nutrients were depleted in an E. coli batch culture. Taken without request from [2].

network with N independent variables, the space of KO's contain 2^N KO sequences. For a typical genome scale metabolic network, this number is very big ($> 10^{180}$ for E. coli). This makes any computation exploring such space unfeasible. Luckly, we hope, not all KO sequences are interesting. Because the network is well connected, very large KO sequences (a set of many blocked reactions) are expected to affect important network properties like the posibility to growth from the medium. This is our hope, that the feasible KO space is not prohibited big.

Once we can explore the feasible space of KO configurations, we can compare the configurations dependency on all nutrients of interest. For instance, we can characterize the average number of KO required for the network to have glucose as an essential nutrient. Or we can count how many of such networks there are. By comparing such measurements between nutrient, we can argue we are computing the relative bias toward a given nutrinet present in the topology of the network. Note that, for now, we are only taking into consideration the connectivity of the network, not the stoichiometry (eg. shadow prices).

3 WIP, SOME RESULTS: KOMC

At the cluster we are running a Monte Carlos exploring the KO space for the genome scale network of E. coli iJO1366 (metabolites 1110, reactions 1694, free reactions 584) [3]. Right now we are just computing the limits of the feasible space (finding KO sequences which are fatal). Here we are presenting the results for the first 550011 samples for the medium described at Figure 1 reference. I have not completed the scripts for evaluating nutrient essentiality yet. So, I'll present only some statistics of the KO samples.

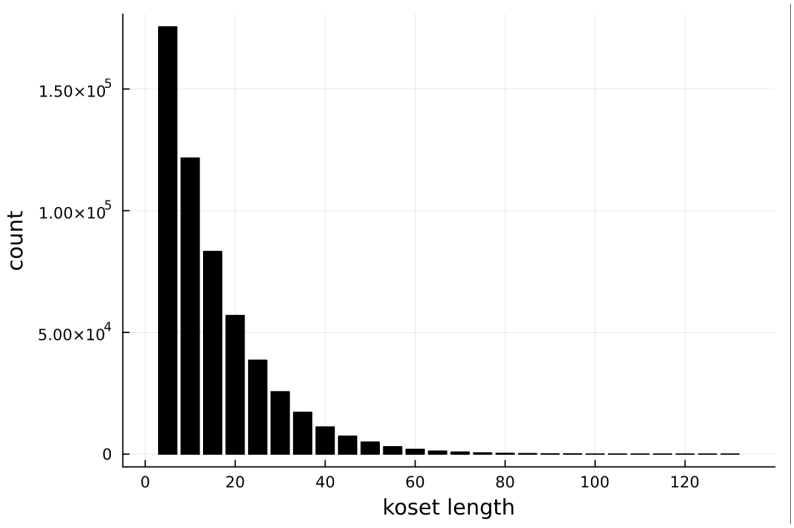

Figure 3: Histogram of the lenghts of KO sequences for 550011 MC samples.

Figure 3 shows the distributions of the lenghts of the KO sequences. A 20 means that 20 reactions were blocked before the network was unable to produce biomass.

Figure 4 shows the histogram for the single reaction combinatory (the count of each reaction on the KO configurations), and the two reaction combinatory (the count of each pair of reactions). In the first panel (top) we can clearly see the fatal single KOs, they are very common at shorter KO sets, and less common at the larger. This is obious, short KO sets can mostly be so it they contain a very fatal limitation, for large KO sets is the oposite, if they contain them it can only be one after many non fatal reactions. Note that the x axis are just an enumeration of the elements (reactions or pair of reactions) and aren't the same for each line (beacuase of plotting considerations). That is, at Figure 4 we are only studying the shape of the histograms.

At the momment, the main questions are: How close to equilibrium is the MC? How much time we should keep running it? It is feasible?

4 SUMMARY

Computations are probably unfeasible, questions aren't clear, data non available or noisy, rewards not very promising. Unfortunately, just like 7 months ago.

REFERENCES

[1] Q. K. Beg, A. Vazquez, J. Ernst, M. A. de Menezes, Z. Bar-Joseph, A.-L. Barabási, and Z. N. Oltvai. Intracellular crowding defines the mode and sequence of substrate uptake by Escherichia coli and constrains its metabolic activity. *Proceedings of the National Academy of Sciences*, 104(31):12663–12668, July 2007.

[2] Andres Maser, Karl Peebo, and Ranno Nahku. Avoiding amino acid depletion in a complex medium results in improved Escherichia coli BW25113 growth. *Microbiology*, 165(1):37–46, 2019.

[3] Jeffrey D Orth, Tom M Conrad, Jessica Na, Joshua A Lerman, Hojung Nam, Adam M Feist, and Bernhard Ø Palsson. A comprehensive genome-scale reconstruction of Escherichia coli metabolism—2011. *Molecular Systems Biology*, 7(1):535, January 2011.

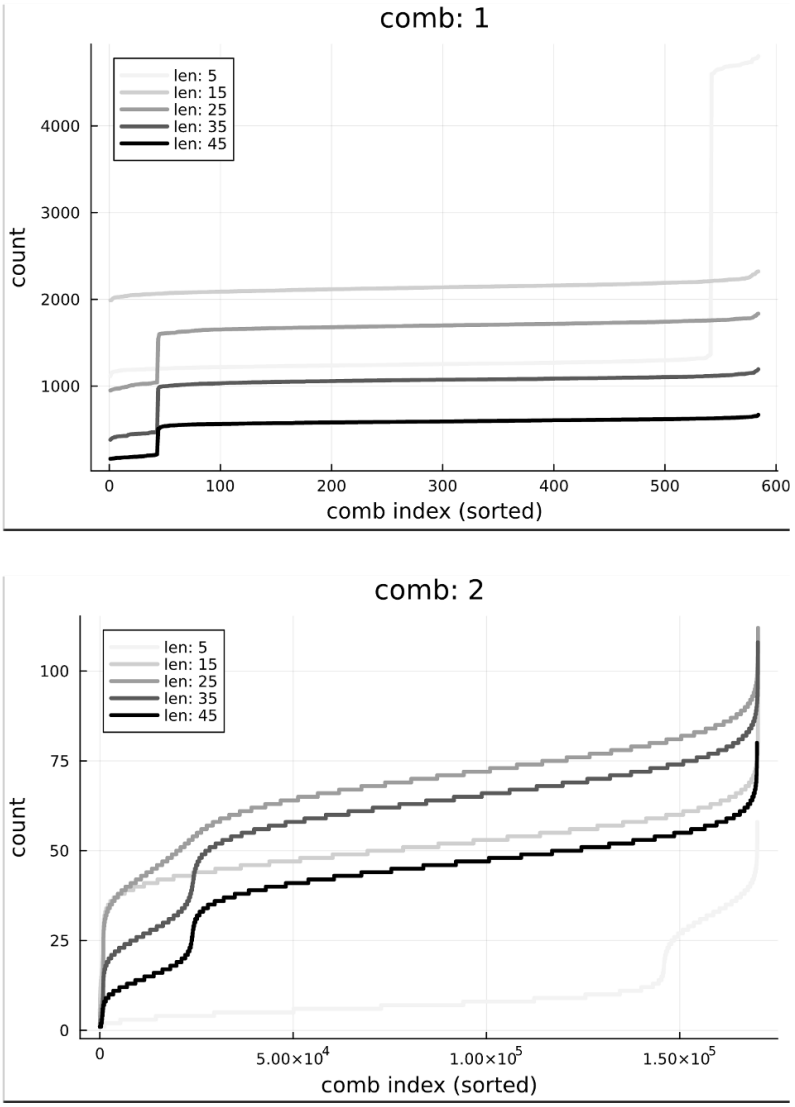

Figure 4: Histogram of the count of each reactions and pair of reactions for 550011 unfeasible KO samples.