I started by reading through the not gate code that Bryan wrote and trying to understand it. I think I understand it works except for the IPTG input portion. I will look more into this later. I plan to try to use this template as a stepping stone to make a two input nor circuit.

Bryan suggested using spyder for python to make graph visualization a bit easier. I don’t want to reinstall python as anaconda python because I am not sure if it will still have Tkinter, so I will see if there is a way to just get spyder IDE for my current python.

It turns out the python I already had came with Spyder. I was able to get it running and will now continue with trying to make a NOR circuit. The two inputs will be x and y, and the output will be z.

I am shifting focus to what Jing is working on. After discussing with Bryan and Jing, we determined that one goal is to find how often ribosomes are present at each codon.

I learned how to plot multiple reads on one genome in mochiview so I can plot all 4 rdm files and get a more whollistic view of things.

We still need to look into normalizing the scores for number of mrna transcripts especially for the minimal media which doesn’t have mrna-seq data.

For now, though we will try to get what we can from the data. I will first look into how often translation occurs without AUG in the e.coli genes. Jing will also look into which codons we would be expecting. Then we will use the wig files to get the scores for each codon and how often the codon appeared in genes as a normalization factor. We may need to use the mrna wig file data to get a sense of the total number of transcripts which will allow us to better normalize the counts.

Cool aside note that was discussed: If cotranslational folding is a factor that determines overall protein three dimensional shape, then it may not be entirely true that it is possible to predict a proteins structure based solely on the amino acid sequence since different organisms can use rarer codons to allow the protein slightly more time to fold.

{'ATG': 3753, 'ATT': 2, 'CTG': 2, 'TTG': 70, 'GTG': 318}

That is the dictionary of the different start codons and how frequently they appeared.

I also found this on Wikipedia which seems to roughly confirm the numbers.

“E. coli uses 83% AUG (3542/4284), 14% (612) GUG, 3% (103) UUG and one or two others (e.g., an AUU and possibly a CUG).”

The next step is to go through and find the score for each codon and how often each one appears.

prfB, I noticed doesn’t have a multiple of 3 codons. After some research and speaking to Jing, Jing pointed out that one of the papers, mentioned that this gene along with dnaX go through what is called a translational frameshift. Partway through translation, the ribosome shifts reading frames. For dnaX it does this only 50% of the time and makes two different proteins. The paper mentions a lot of special cases that they had to account for like selenocysteine.

I made finished the program that gets the score and total number for each codon, and it is functional.

We compared it to the paper and we got stuck trying to figure out how the paper did its analysis. In the paper in the table, the ribosome occupancy seems to be centered around 1, but for us it is more centered around 8. Their codon usage scale also seems a bit weird. I will need to think about it more overnight and hopefully will have some ideas for tomorrow.