

Assignment 1

Question 1:

The two factors which critically hindered the construction of the high-quality whole-cell models are:

- There is not sufficient biological information about individual molecules and their interactions (with itself and others).
- There is not a single computational method that is capable of modeling or explaining the molecular components and interactions which result in the complex phenotypes.

Question 2:

In this paper, the authors divided the whole-cell model into a number of independent small submodels (on the basis of functional processes) which are comparatively easy to model. This strategy is called as “Divide and Conquer” in the field of computer science. Using this strategy, the authors were able to use the different computational methods which were not useful for modeling the whole-cell. After modeling these sub-models independently, they are integrated together by linking their common inputs and outputs, similar to what is being done in “Divide and Conquer” algorithm.

Question 3:

The *M. genitalium* whole-cell model was validated and tested by *in silico* genome perturbations and by using model independent datasets having a variety of biological functions. Some of the validations results are described below.

- Flux through glycolysis is predicted to have more than 100-fold than through pentose phosphate and lipid biosynthesis pathway which is in agreement with previous studies (Yus et al., 2009).
- The metabolite concentrations predicted by the model were within an order of magnitude of concentration as measured in *E. coli* for two independent studies.
- The prediction of the “burst-like” protein synthesis due to the local effects of intermittent mRNA expression and the global effect of stochastic protein degradation is in agreement with previous reports.
- Distribution of protein and mRNA levels are in concurrent with the previously reported single cell measurements.
- By using *in silico* genome perturbations, the authors carried out a total of 525 single gene disruptions and found out that 284 genes are essential for growth of *M. genitalium* which accounts for 79% accuracy as compared with previous studies.
- The growth rates of 12 single gene disruption strains were measured experimentally and compared with the rates predicted by the model. Two-third of the measured growth rates were consistent with the predicted ones.

Question 4:

In the paper, there is an example in which the authors predicted the DNA binding proteins interactions. In this example, the model predicted the protein occupancy of the DNA binding proteins including the DNA and RNA polymerase accurately. In addition to that, the model also predicted the protein-protein collisions on the chromosome which was previously been reported but only for some specific pairs of proteins. The model predicted that there are approximately 30,000 collisions occurs per cell cycle. Experimentally studying DNA binding protein's collisions at such scale and level is currently not feasible. Therefore, only computational methods can be used to study such mechanisms. The model further characterizes these predicted collisions based on the chromosomal location and found out that these collisions correlate with the DNA-bound protein density. The model also found the most of these collisions are caused by RNA polymerase and DNA polymerase. This is in concordance with what has been proposed in the new central dogma of life. The new central dogma of life says that all the processes (replication, transcription, and translation) occurs simultaneously and due to that there is an accumulation of so many molecules at one place which causes collisions between those molecules. In addition to that, it has also been proposed that these collisions also lead to the removal of the proteins which binds at non-specific regions leading to the destruction of wrong transcripts or peptides.