Anirudh Rao
BE21B004

BT4110 – Computational Biology Lab

Constraint-Based Modelling Assignment

1

Genome-scale metabolic modelling allows us to simulate the cellular metabolism of various organisms. They provide a mathematical relation between genes, metabolites, and reactions that occur in a cell. Creating a genome-scale metabolic model typically begins with a genome annotation to identify genes encoding metabolic enzymes, followed by mapping these genes to known biochemical reactions.

The principle behind genome-scale metabolic modelling involves formulating the metabolic network as a set of reactions, including metabolites, enzymes, and products. These reactions are organized into a stoichiometric matrix.

We define the rates of metabolite production or consumption as metabolic fluxes. We designate the flux of one or more reactions as our objective function to be maximised under steady-state mass balance. We wish to compute the optimal fluxes through all the reactions in the network. However, we can also apply linear constraints on the fluxes of other reactions to restrict the solution space. This is the principle of flux balance analysis (FBA).

The scope of genome-scale metabolic modelling extends across diverse applications, including predicting the effects of gene knockouts, metabolic engineering for improved bioproduction of compounds, and identifying drug targets in pathogens. For instance, genome-scale metabolic models are widely used in biotechnology to optimize microbial strains for biofuel or pharmaceutical production. In medicine, genome-scale metabolic modelling can help elucidate disease mechanisms by identifying metabolic vulnerabilities in cancer cells or pathogenic organisms.

Genome-scale metabolic models are particularly valuable for studying non-model organisms and can be adapted for synthetic biology applications, where synthetic pathways are designed and integrated into organisms. The flexibility of genome-scale metabolic models, coupled with their capability to integrate omics data, allows for the holistic modeling of metabolism and its responses to environmental changes, making it a powerful tool in systems biology.

References

1. Passi, A., et al (2021). Genome-Scale metabolic modeling enables In-Depth understanding of big data. Metabolites, 12(1), 14. https://doi.org/10.3390/metabo12010014

2. Gu, C., et al. (2019). Current status and applications of genome-scale metabolic models. Genome Biology, 20(1). https://doi.org/10.1186/s13059-019-1730-3
3. Raman, K. (2021). An introduction to Computational Systems Biology. In Chapman and Hall/CRC eBooks. https://doi.org/10.1201/9780429486951

2

The stoichiometric matrix is given by:

	V1	V2	V3	V4	V5	V6	V7	R1	R2	R3	R4
Glucose	−1	0	0	0	0	0	0	1	0	0	0
Pyruvate	2	−1	−1	−1	0	0	0	0	0	0	0
Lactate	0	1	0	0	0	0	0	0	−1	0	0
Acetate	0	0	0	0	1	0	0	0	0	0	−1
Acetyl CoA	0	0	1	1	−1	−1	0	0	0	0	0
Formate	0	0	0	1	0	0	0	0	0	0	0
Ethanol	0	0	0	0	0	1	0	0	0	−1	0
ADP	−2	0	0	0	−1	0	1	0	0	0	0
ATP	2	0	0	0	1	0	−1	0	0	0	0
NAD	−2	1	−1	0	0	2	0	0	0	0	0
NADH	2	−1	1	0	0	−2	0	0	0	0	0
CoA	0	0	−1	1	1	1	0	0	0	0	0
H_2O	2	0	0	0	0	0	0	0	0	0	0
CO_2	0	0	1	0	0	0	0	0	0	0	0

The bounds are:

Reaction	Bounds
R1	$(-\infty, 0)$
V1	$(0, \infty)$
V2	$(-\infty, \infty)$
V3	$(0, \infty)$
V4	$(0, \infty)$
V5	$(0, \infty)$
V6	$(0, \infty)$
V7	$(0, \infty)$
R2	$(0, \infty)$
R3	$(0, \infty)$
R4	$(0, \infty)$

For uptake reactions, the bounds are set as $(-\infty, 0)$. For secretion reactions, the bounds are set as $(0, \infty)$. For irreversible reactions, the bounds are set as $(0, \infty)$. For reversible reactions, the bounds are set as $(-\infty, \infty)$. In practice, we set ∞ as a real number like 1000.

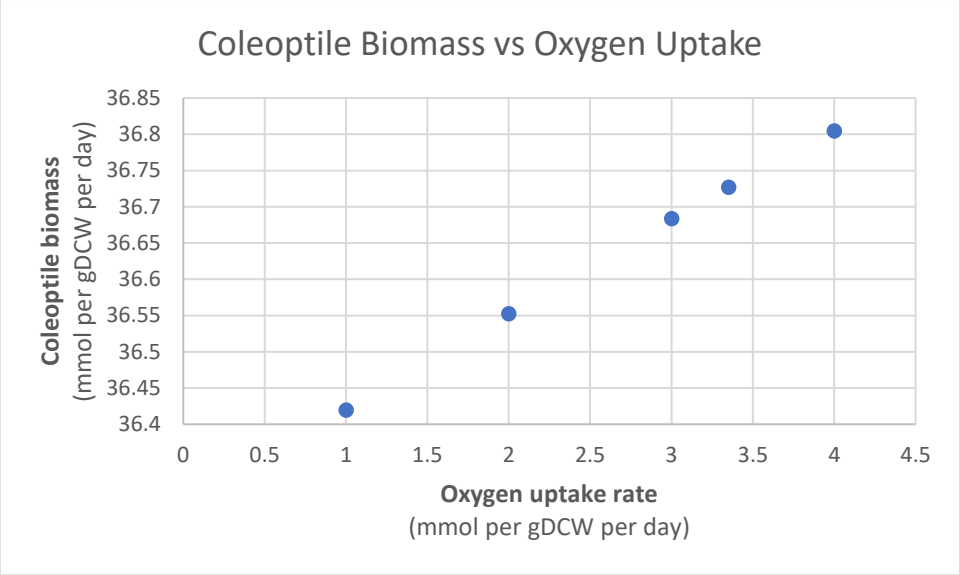
3a

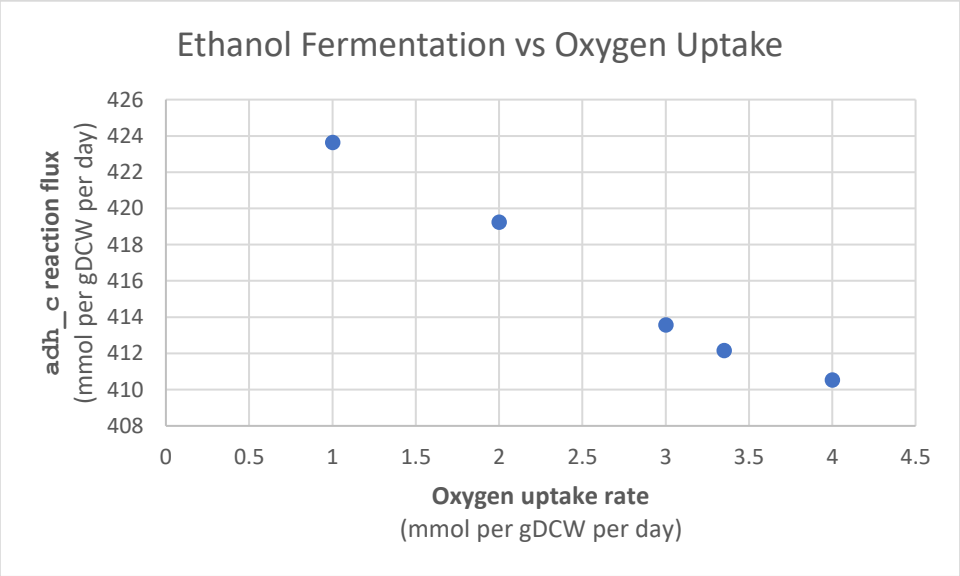
To simulate seed-derived cells in aerobic and anaerobic conditions, the following steps were performed:

- nanoCOBRA was added to the MATLAB path
- The *Oryza sativa* model was loaded
- The objective function was changed to the coleoptile biomass
- The sucrose uptake rate was constrained to 1 mmol per gram DCW per day
- The oxygen uptake rates were iteratively changed to 1, 2, 3, 3.35, and 4 mmol per gram DCW per day by updating the lower bound
- The FBA simulation was performed by optimizing the model
- The flux vectors under various oxygen availabilities were compared

The results observed from the simulation were similar to that in the paper:

- There was a linear decrease in biomass as oxygen uptake reduced.
- There was a linear increase in ethanol fermentation (‘`adh_c`’) with decrease in oxygen uptake.
- When the oxygen uptake increased from 3.35 to 4 mmol per gram DCW per day, the ethanol production did not increase by much.





3b

To simulate seed-derived cells in batch culture with sucrose / glucose under aerobic and anaerobic conditions, the following steps were performed:

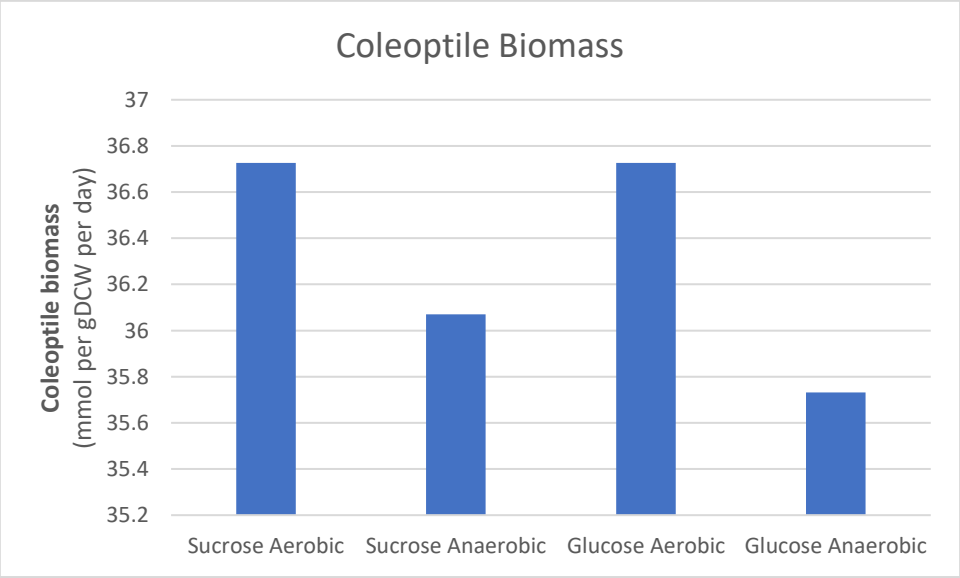
- The objective function was set to the coleoptile biomass
- The sugar uptake rates were constrained based on the experimental results. These rates are tabulated below
- The oxygen uptake rates was set to 3.35 or 0 mmol per gram DCW per day by updating the lower bound for aerobic and anaerobic conditions respectively
- The FBA simulation was performed by optimizing the model
- The flux vectors under the 4 conditions were compared

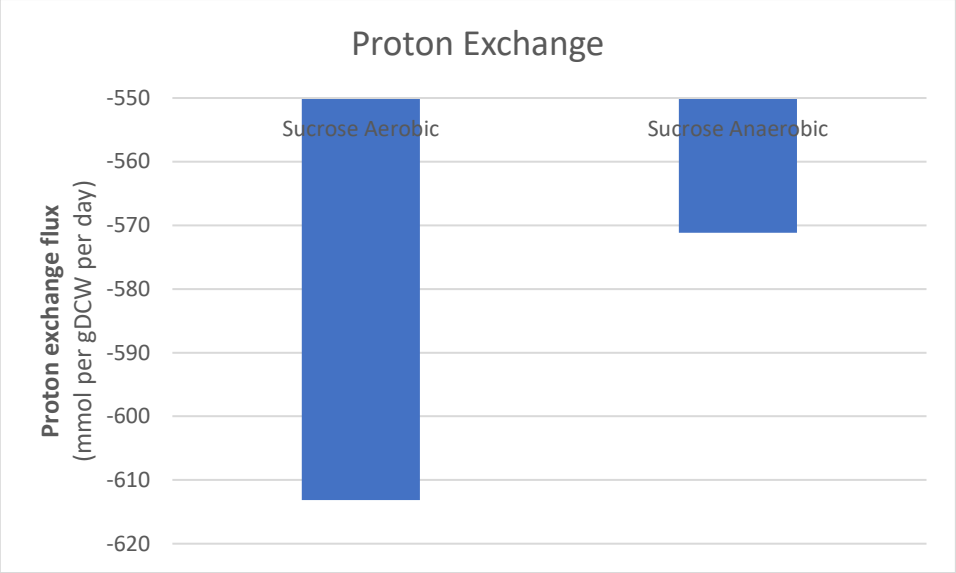
Constraints Used

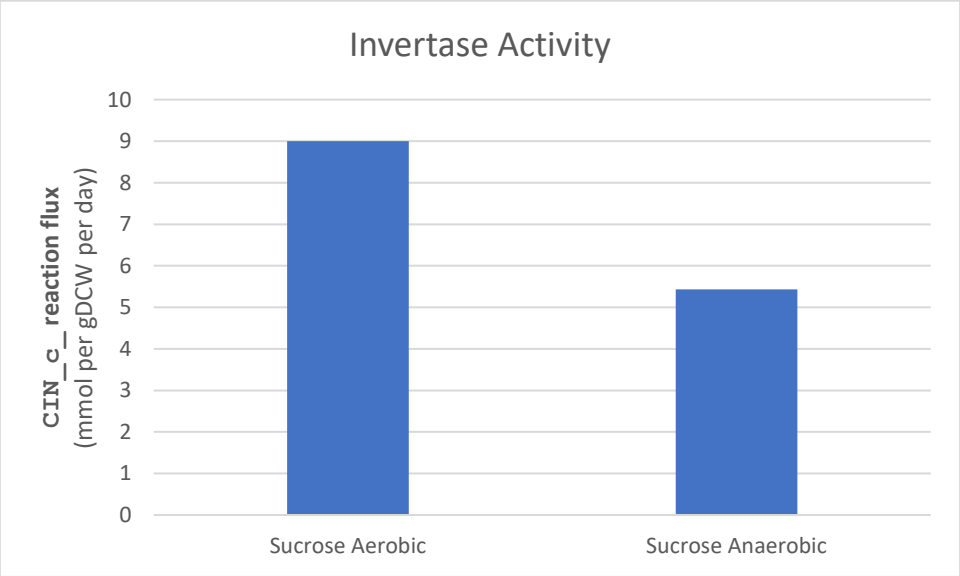
Simulation	Sugar Exchange Rate (mmol per gDCW per day)	Oxygen Exchange Rate (mmol per gDCW per day)
Sucrose + Aerobic	Suc: $\frac{7.5-30}{(8-4)\times 2}\times\frac{10^3}{342}=-8.224$ Glc: $\frac{5-0}{(8-4)\times 2}\times\frac{10^3}{180}=3.472$ Fru: $\frac{6-0}{(8-4)\times 2}\times\frac{10^3}{180}=4.167$	3.35
Sucrose + Anaerobic	Suc: $\frac{17.5-24}{(5-4)\times 2}\times\frac{10^3}{342}=-9.5$ Glc: $\frac{0-0}{(5-4)\times 2}\times\frac{10^3}{180}=0$ Fru: $\frac{2-0}{(5-4)\times 2}\times\frac{10^3}{180}=5.556$	0
Glucose + Aerobic	Glc: $\frac{28-31}{(5-2)\times 2}\times\frac{10^3}{180}=-2.778$	3.35
Glucose + Anaerobic	Glc: $\frac{25-32}{(4-3)\times 2}\times\frac{10^3}{180}=-19.444$	0

The results observed from the simulation were similar to that in the paper:

- The growth rate was slower in anaerobic conditions for both sucrose and glucose.
- When sucrose was used as the carbon source, the proton exchange was higher in the aerobic case. This results in a more acidic pH in the culture medium.
- When sucrose was used as the carbon source, the cell wall-associated invertase ('`CIN_c_`') activity was higher in aerobic conditions.





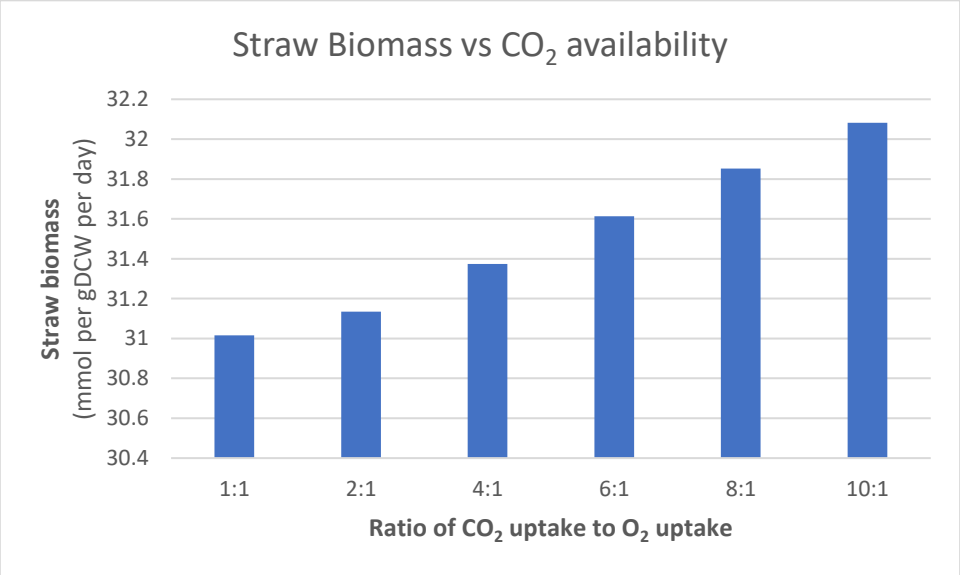


3c

To simulate photorespiring leaf cells under different oxygen and carbon dioxide conditions, the following steps were performed:

- The objective function was set to the straw biomass
- The photon uptake rate was constrained to 100 mmol per gram DCW per day, as suggested by the paper
- The oxygen and carbon dioxide uptake rates were iteratively changed to the values given in the question by updating the upper bound
- The FBA simulation was performed by optimizing the model
- The flux vectors under the 7 conditions were compared

The growth rate was lower when carbon dioxide was in lower availability. This is similar to the result obtained in the paper.



4

To perform flux variability analysis under the given conditions, the following steps were performed:

- The objective function was set to the coleoptile biomass for seed cells and straw biomass for leaf cells
- The oxygen uptake and growth rates were constrained based on the question
- FVA was performed with the `optPercentage` parameter set to 90%

The script and results of all simulations and FVA are attached with this report.