SFU

CMPT441/711
Bioinformatics Algorithms

Project D
Understanding Metabolic Network Models

Members: Hooman Zabeti(301348564), Yiji Wang(301286922)
Peiyu Cui(301345033), Haihong Tang(301268397)

Supervised by: Prof. Leonid Chindelevitch

Outline

- Introduction to metabolic network
- Analysis of the models
- Conclusion
- Q&A

Introduction

- Metabolic network
- Constraint-based models
- Flux Balance Analysis
- Blocked reactions

Metabolic Network

Representation of the biochemical reactions within the cell:

- Metabolite: an organic substance
- Biochemical reaction: the process in which two or more molecules (reactants) interact

Modelling the network:

- Kinetic models
- Approximated kinetic models
- Constraint-base models
- Conventional models
- Topological analysis

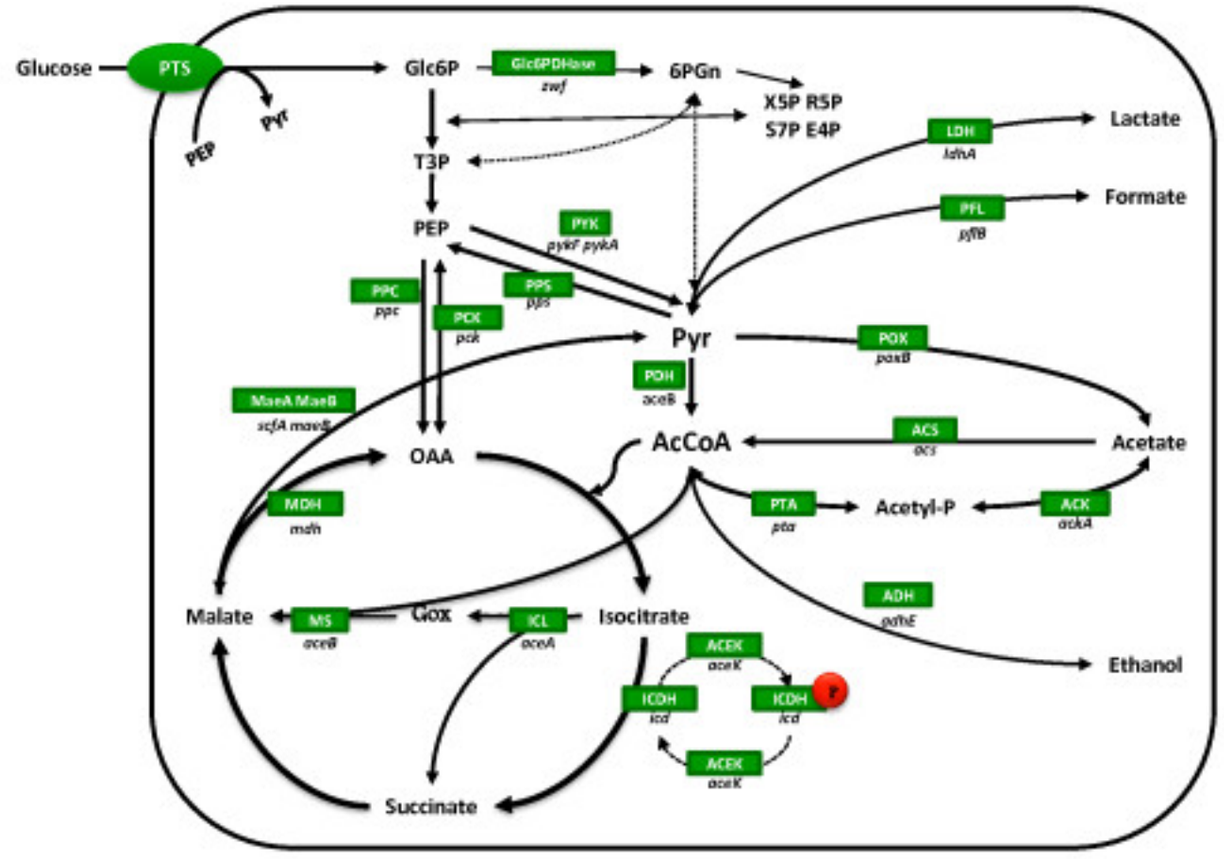


Figure 1: Simplified model for the central metabolic network of E. coli metabolism. [1]

Constraint-based Models

Mathematically encoding a metabolic network reconstruction

Stoichiometric Matrix:

- Rows = represents a unique **Metabolite**
- Columns = represents a **Biochemical reactions**
- Entries = **Stoichiometric coefficients**

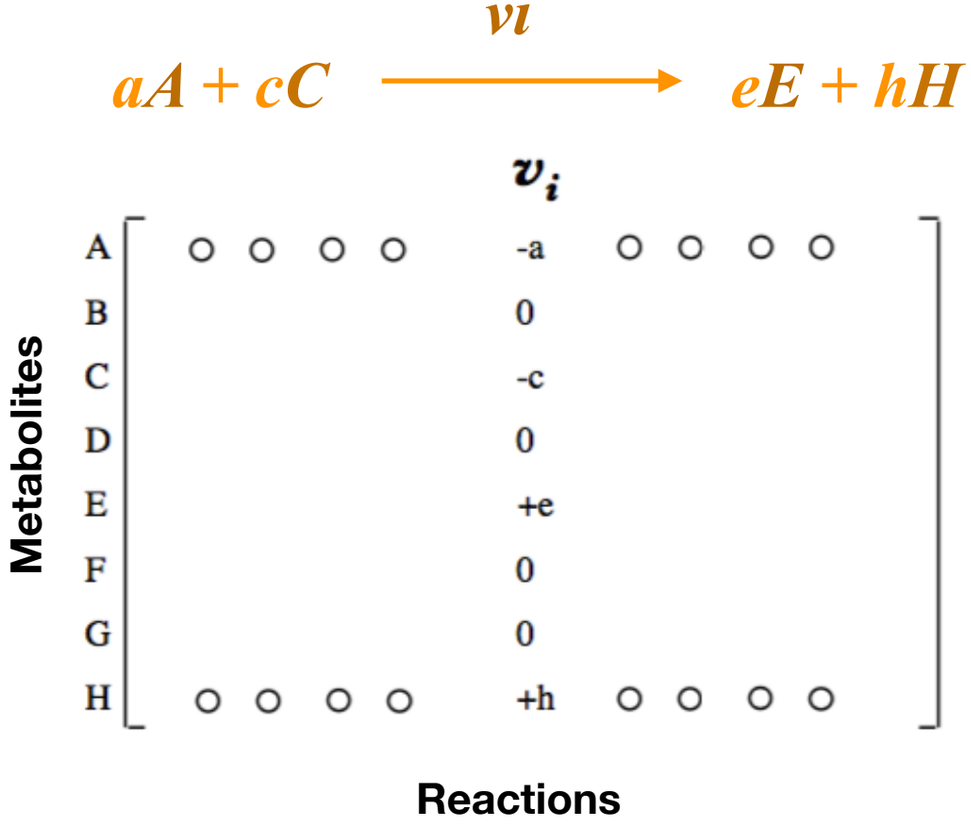
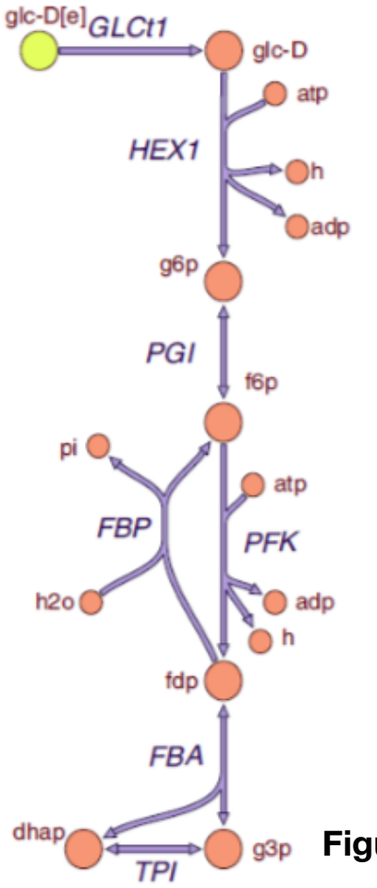


Figure 2: Stoichiometric Matrix [2]



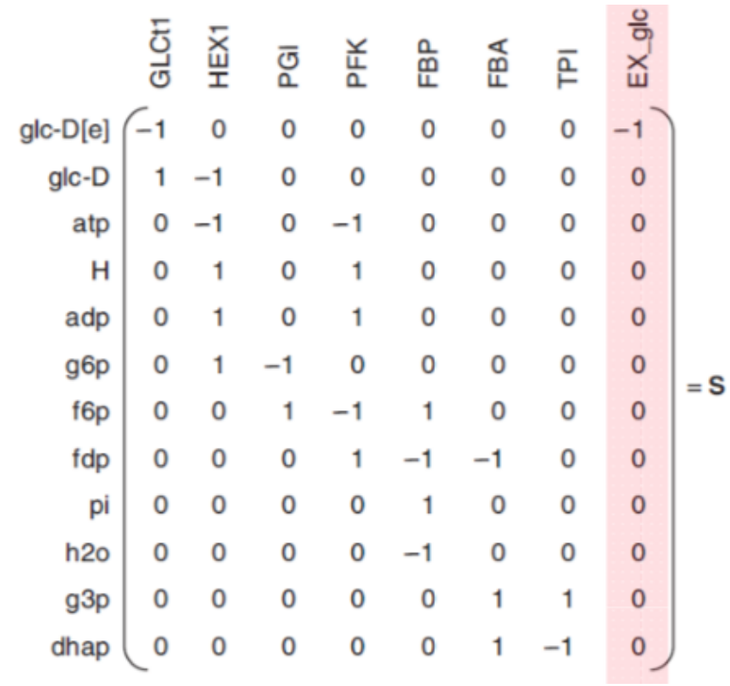


Figure 3: Metabolic pathway and Stoichiometric Matrix [3]

Flux Balance Analysis

- Mathematical approach for analyzing the flow of metabolites through a metabolic network
- Predict the growth rate of an organism
- Stoichiometric matrix S is a linear transformation of flux vector v to derivatives of concentration vector x
- $dx/dt = Sv$
- Steady-state assumption: $Sv=0$

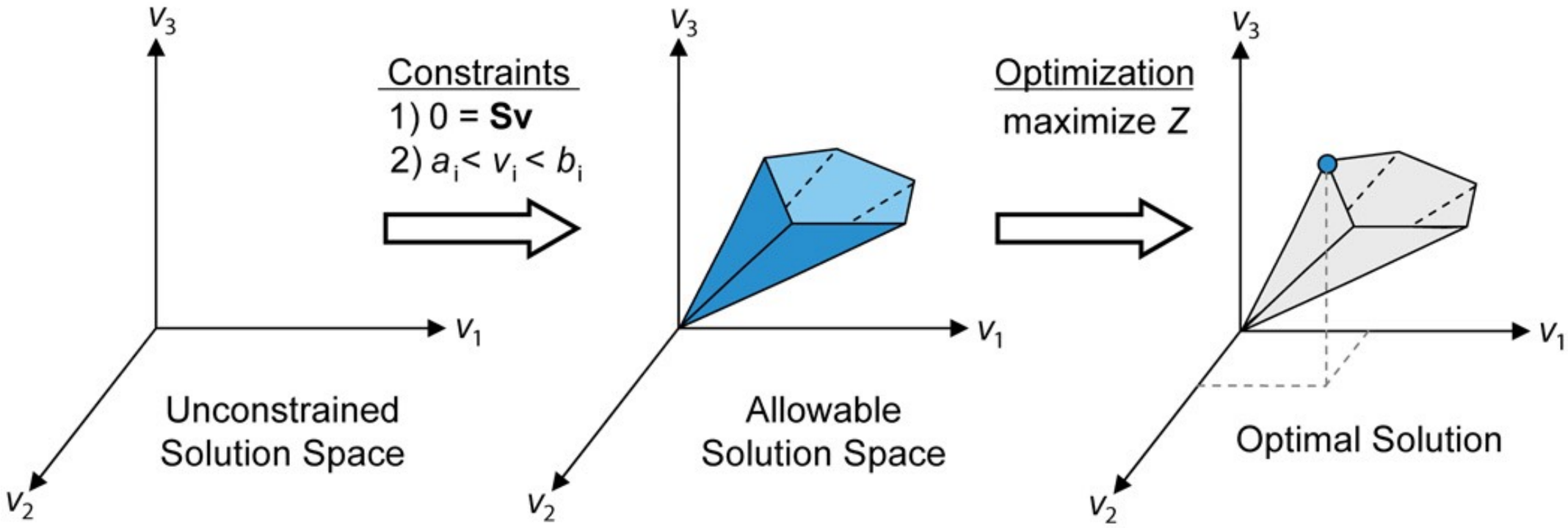


Figure 4: The conceptual basis of constraint-based modelling. [4]

Flux Balance Analysis

- **Flux Optimization:**

Max (Min):

$$c^T v$$

Subject to:

$$Sv = 0$$

$$\alpha_i \leq v_i \leq \beta_i$$

c = **vector of weights**

α_i = **lower bound of** v_i

β_i = **upper bound of** v_i

Formula 1: Flux Optimization functions [5]

- **Formulation of FBA**

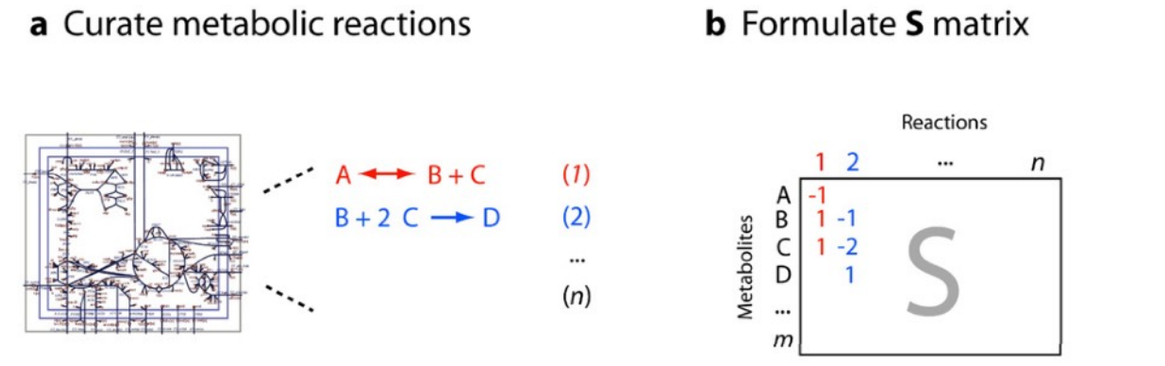


Figure 5: Formulation of an FBA problem. [4]

Blocked Reactions

- Reaction i is called **blocked reaction** if:

Steady-state: $Sv = 0$

Non-negative flux: $v_i \geq 0 \quad \forall i \in I$

$$\Rightarrow \quad v_i = 0$$

Formula2: Blocked reaction definition [6][7]

- Type of blocked reactions (Mongoose):
  - Topologically blocked
  - Stoichiometrically blocked
  - Irreversibility blocked

Analysis

Toolbox	**Fullname**	**Platform**	**Year**	**Homepage**
Cobra	**CO**nstraint-**B**ased **R**econstruction and **A**nalysis Toolbox	Matlab	2007	https://opencobra.github.io/cobratoolbox/stable/
Mongoose	**M**etab**O**lic **N**etwork **G**r**O**wth **O**ptimization **S**olved **E**xactly Toolbox	Python	2014	http://cb.csail.mit.edu/cb/mongoose/

Cobra VS Mongoose

Use Yeast Model (iMM904 & iND750) as gold standard data for comparison

http://gcrg.ucsd.edu/InSilicoOrganisms/Yeast

- Less "Bias" for different tools
- Reliable and Stable
- Many studies on yeast models
- Moderate size for analysis

Yeast Model	# Metabolites	# Reactions	# Genes
iMM904	1266	1577	905
iND750	1059	1266	750

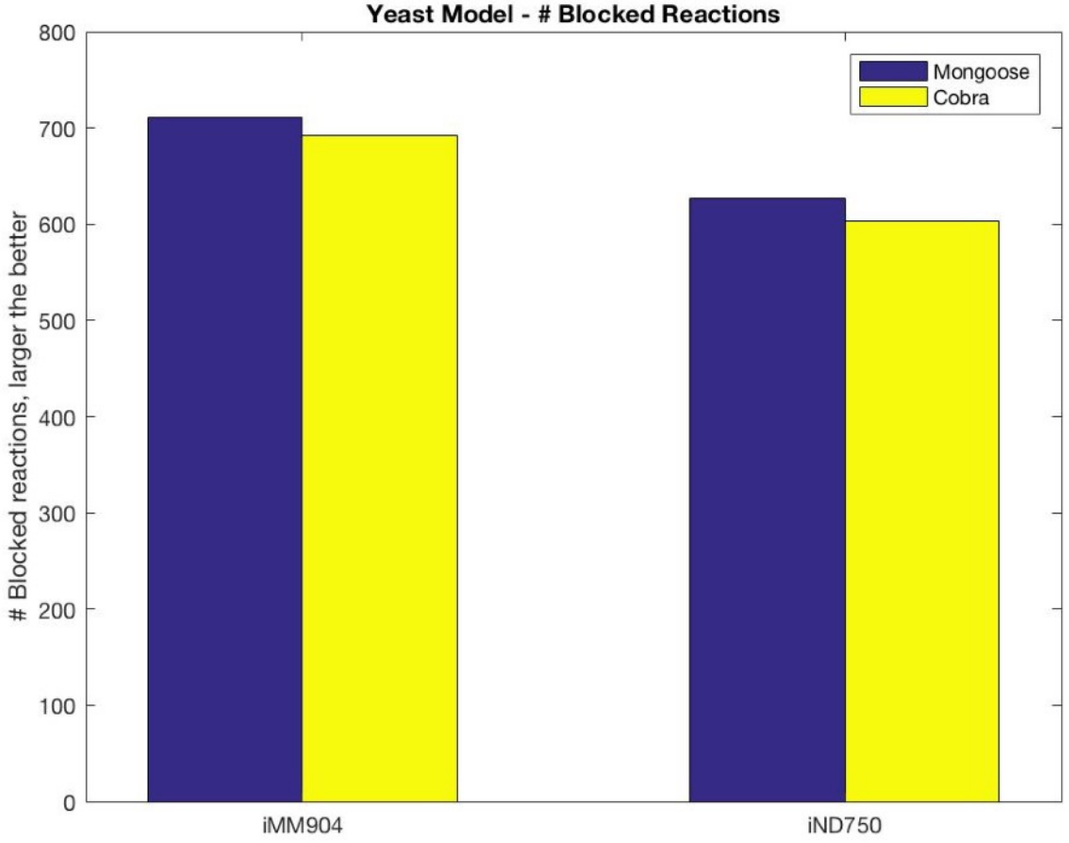


Figure 6: # of blocked reactions for yeast models

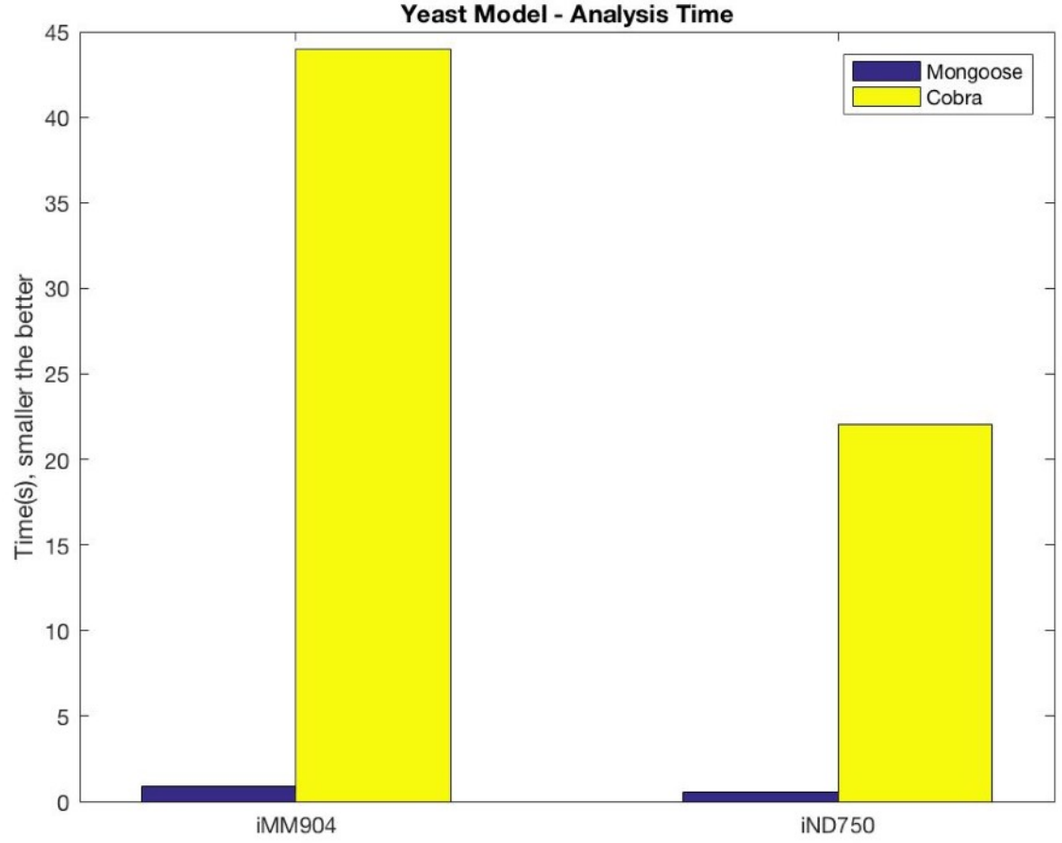


Figure 7: Analysis time for yeast models

Cobra VS Mongoose

An exact arithmetic toolbox for a consistent and reproducible structural analysis of metabolic network models.

Leonid Chindelevitch, Jason Trigg, Aviv Regev, Bonnie Berger.
Nature Communications, 5:4893 (2014).

Keywords: Exact Arithmetic VS Floating Arithmetic

Mongoose > Cobra ?

Models

1 to 84 (84)

BiGG ID	Organism	Metabolites	Reactions	Genes
e_coli_core	Escherichia coli str. K-12 substr. MG1655	72	95	137
iAB_RBC_283	Homo sapiens	342	469	346
iAF1260	Escherichia coli str. K-12 substr. MG1655	1668	2382	1261
iAF1260b	Escherichia coli str. K-12 substr. MG1655	1668	2388	1261
iAF692	Methanosarcina barkeri str. Fusaro	628	690	692
iAF987	Geobacter metallireducens GS-15	1109	1285	987
iAPECO1_1312	Escherichia coli APEC O1	1942	2735	1313
iAT_PLT_636	Homo sapiens	738	1008	636
iB21_1397	Escherichia coli BL21(DE3)	1943	2741	1337
iBWG_1329	Escherichia coli BW2952	1949	2741	1329
ic_1306	Escherichia coli CFT073	1936	2726	1307

Figure 8: BIGG models (http://bigg.ucsd.edu/models)

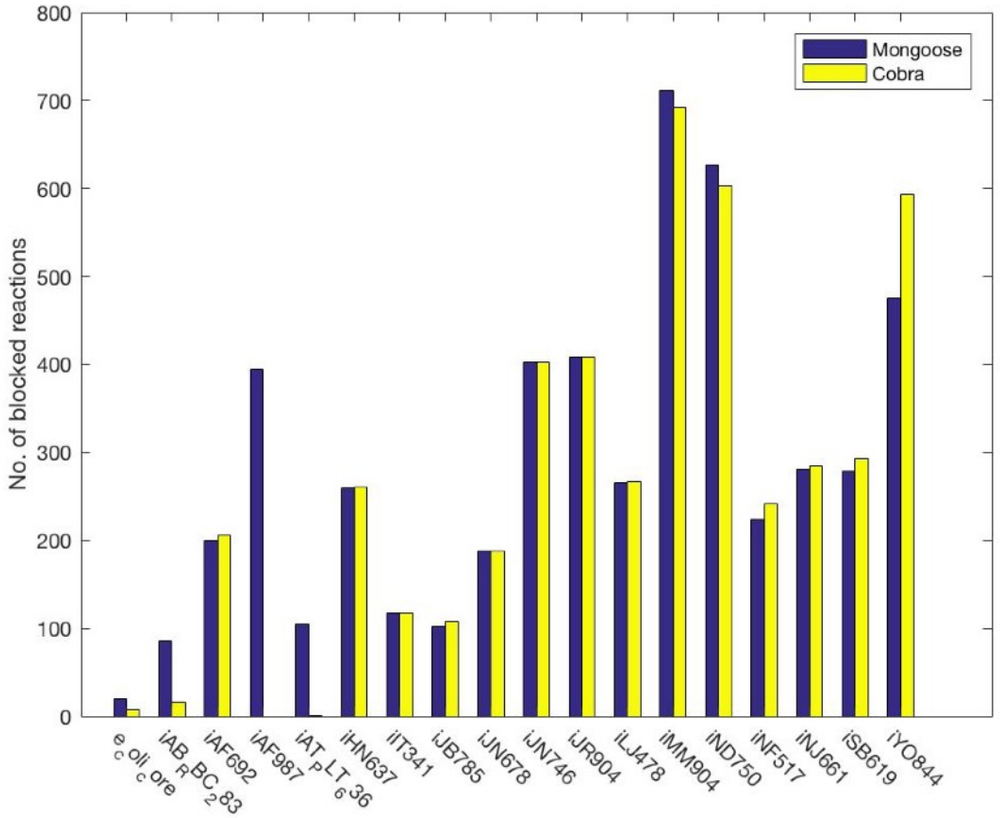


Figure 9: # of blocked reactions for 18 small models

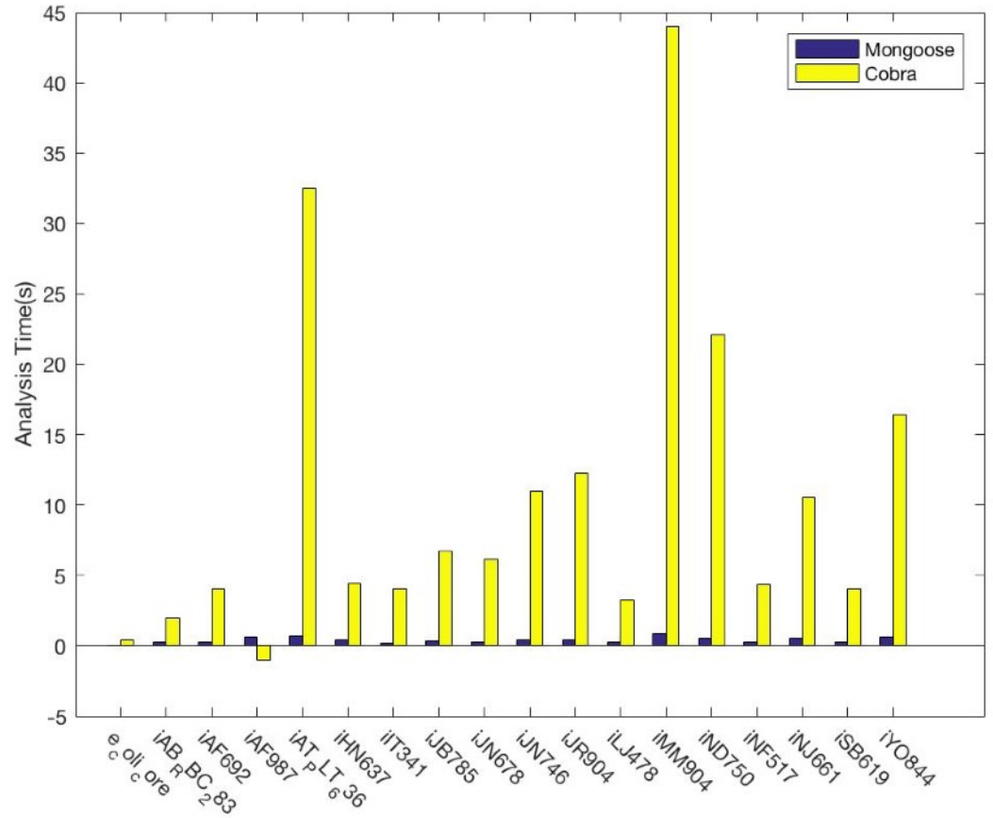


Figure 10: Analysis time for 18 small models

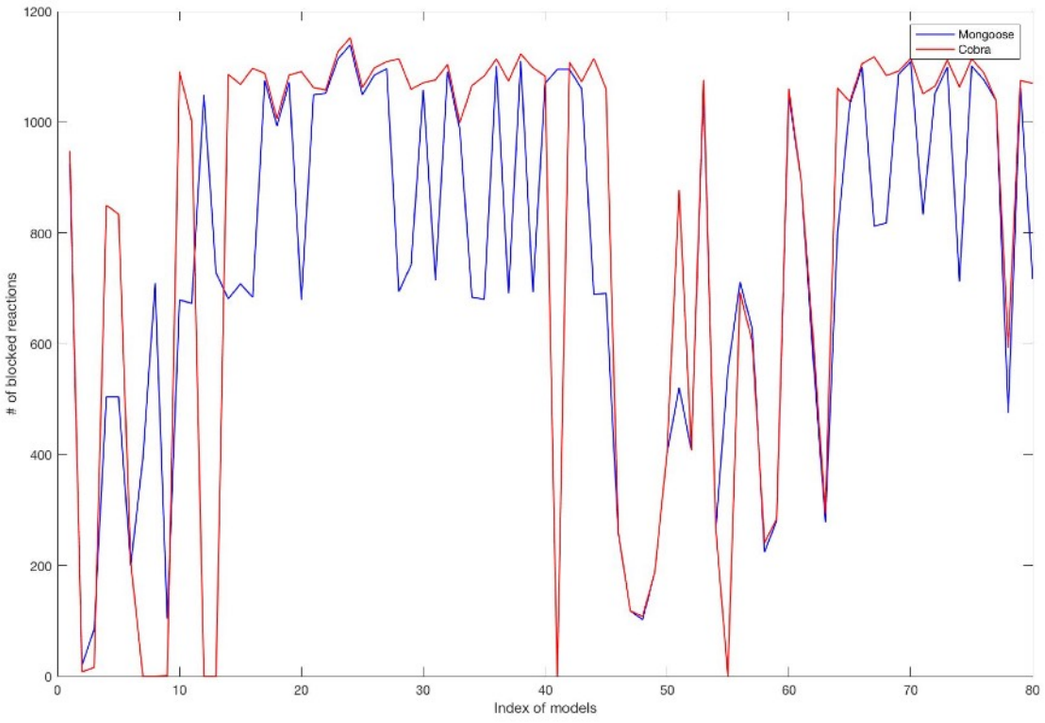


Figure 11: # of blocked reactions for all models

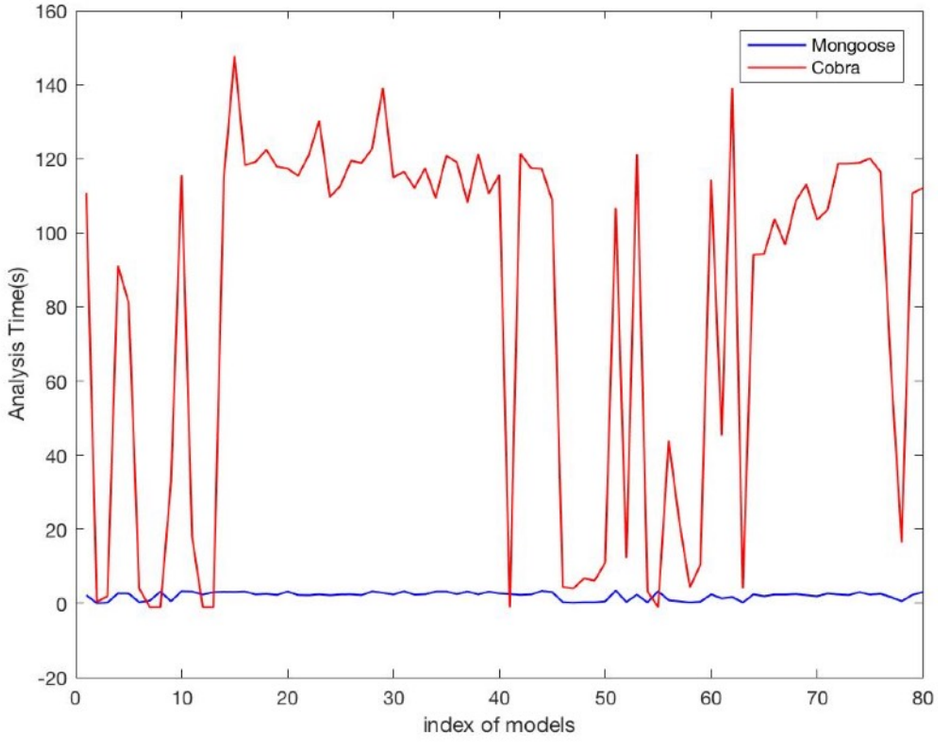


Figure 12: Analysis time for all models

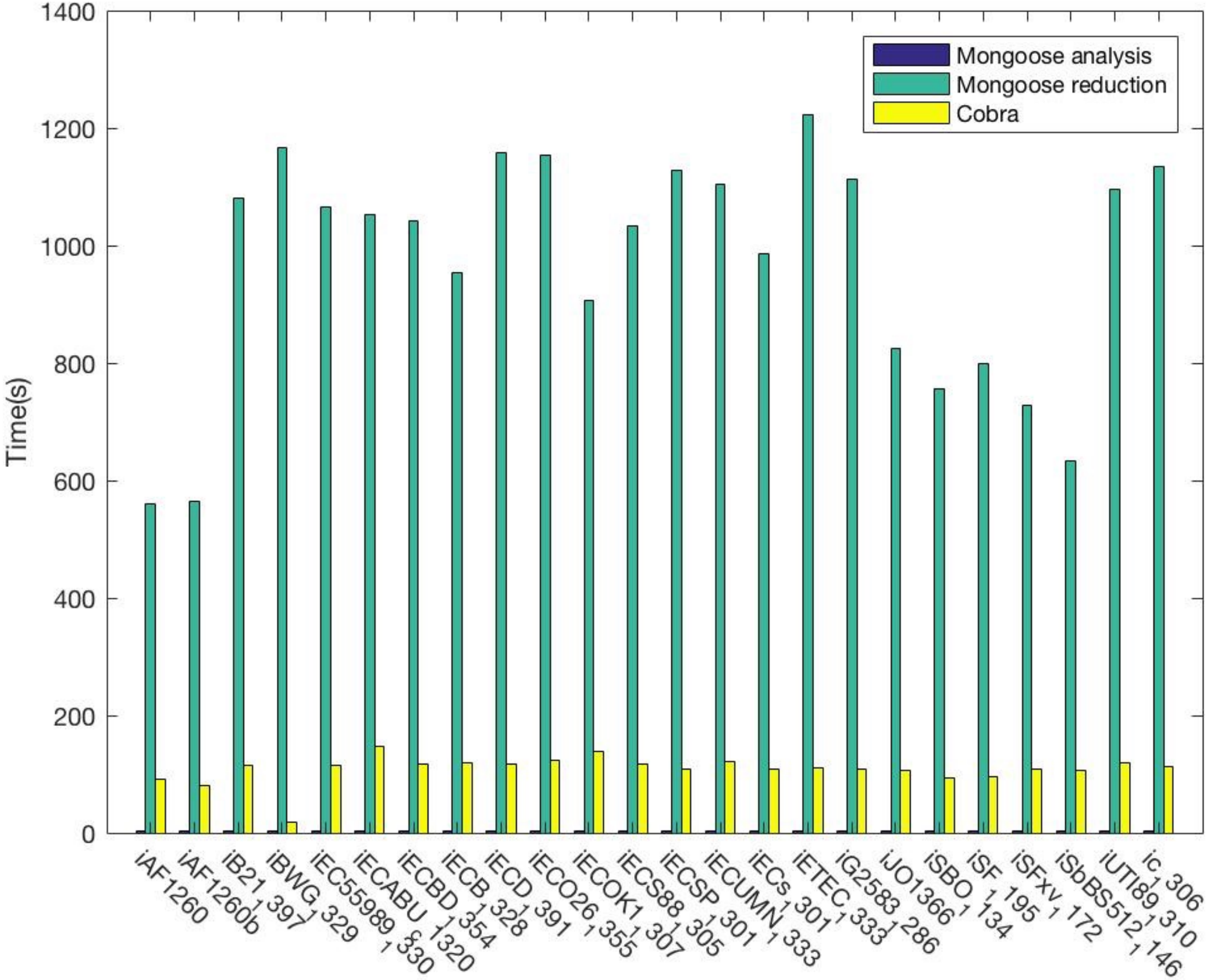


Figure 13: Time cost differences between Cobra and Mongoose

Conclusion

Easy Goal:

- Obtained basic knowledge of metabolic network models
- Went through Cobra tutorial: https://bioinformatics.ca/workshops/2008/systems-network-biology#outline

Hard Goal:

- Installed and learned how to analyze models using Mongoose
- Observed different results provided by Cobra and Mongoose using Yeast Model as gold standard

Stretch Goal:

- Compared Cobra and Mongoose based on results from all 80 BIGG models
- Analyzed the differences between Cobra and Mongoose in terms of accuracy and efficiency

Cobra	• Offers various functions for data integration, model design, analysis, reconstruction and visualization • Provides flexibility in choosing solvers for analysis (solver type: 'LP', 'MILP', 'QP', 'MIQP', for each solver type, provides different solvers e.g 'CPLEX', 'GLPK', 'DQQ') • Supports parallel computing in MATLAB (e.g. fluxVariability) • Has many documentations and supports
Mongoose	• Identifies novel structural features (3 kinds of blocked reactions & 2 kinds of semi-blocked reactions) • Gives better consistency and reproducibility • Diagnoses blockages and proposes ways to remedy them • Significantly reduces the size of the networks • Identifies small reaction cut-sets and minimal media.

To be continued…

Source of image: http://cdn.al-masdar.net/uploads/2014/11/000_Mvd1204547-618x358.jpg

References

[1] Castaño-Cerezo, Sara & Pastor, José & Renilla, Sergio & Bernal, Vicente & Iborra, Jose & Cánovas, Manuel. (2009). "**An insight into the role of phosphotransacetylase (PTA) and the acetate/acetyl-CoA node in Escherichia coli. Microbial cell factories."** 8. 54. 10.1186/1475-2859-8-54.

[2] R. Mahadevan. (2008). "**Constraint-based Models of Metabolism.**" Canadian Bioinformatics Workshops. Module 4: Metabolic modelling. https://bioinformatics.ca/systemsday1-module4pdf

[3] Becker, S. A., A. M. Feist, et al. (2007). "**Quantitative prediction of cellular metabolism with constraint−based models: the COBRA Toolbox.**" Nature protocols 2(3): 727−738.

[4] Orth, J. D., Thiele, I., & Palsson, B. Ø. (2010). "**What is flux balance analysis?**" Nature Biotechnology, 28(3), 245–248. http://doi.org/10.1038/nbt.1614

[5] H. Scott Hinton. (2015). "**Flux Balance Analysis Overview.**" https://systemsbiology.usu.edu/documents/Flux%20Balance%20Analysis%20Overview.pdf

[6] Leonid Chindelevitch, Jason Trigg, Aviv Regev & Bonnie Berger. (2014). "**An exact arithmetic toolbox for a consistent and reproducible structural analysis of metabolic network models.**" Nature Communications, 5:4893

[7] Leonid Chindelevitch. (2006). "**Extracting Information from Biological Networks**". DSpace@MIT. https://dspace.mit.edu/handle/1721.1/64607

Thank you!

Q&A