Systems biology from simple models to biomarker identifications

Ibrahim Elsemman

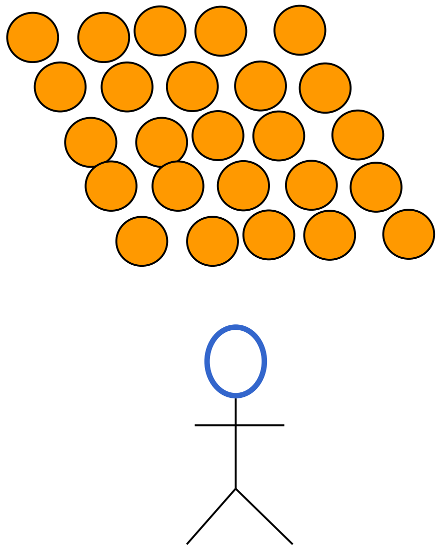
PostDoc,
The Novo Nordisk Foundation Center for Biosustainability,
Technical University of Denmark,
Denmark.

Outlines

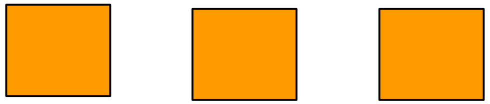
- **What is Systems Biology?**
- **How to cross the bridge from math to biology or from biology to math?**
- **How to build a simple model?**
- **How to analyze RNA-seq data and integrate the results with a simple model?**
- **How to identify new biomarkers?**
- **Can we build a whole-cell model?**

Simple Problem

25 circles



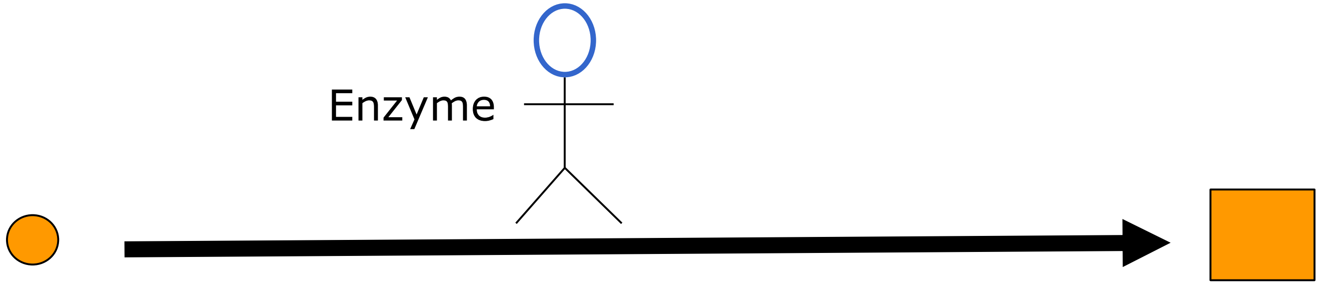
Converts **three circles** to square **per hour**



Q. How many circles remain **after two hours?**

A. This man can convert 6 circles after two hours, so there are 19 circle after two hours.

MATH



Enzyme

Metabolite

Substrate

Kcat = three molecules per hour per one Enzyme

Flux <= kcat [number of enzymes]

Flux <= V_{max}

Metabolite

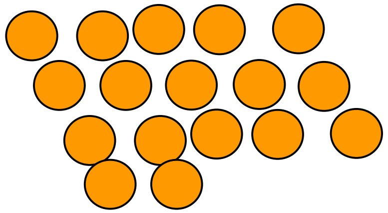
Product

Chemical Reaction

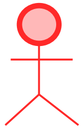
D-glucose 6-phosphate → **D-fructose 6-phosphate**

PGI1

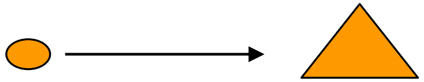
BIO



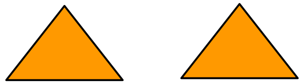
17 circles

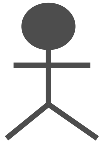


Converts thre 1 circles to triangle per hour

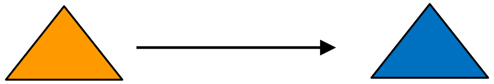


How many triangles can this man moves after two hours?

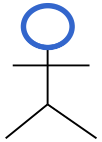




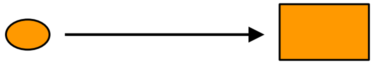
Changes color of 0.5 triangel to blue per hour



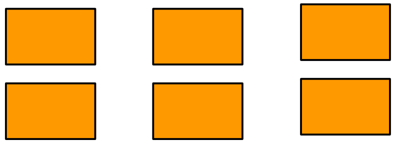
How many triangle can this man change after two hours?

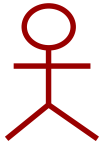


Converts three circles to square per hour

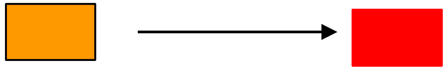


How many squares can this man convert after two hours?





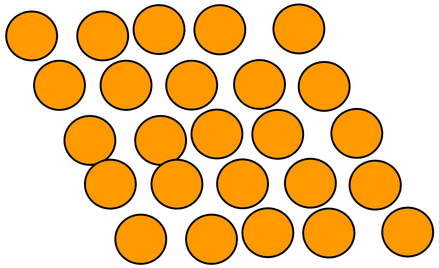
Changes color of two squres to red red per hour



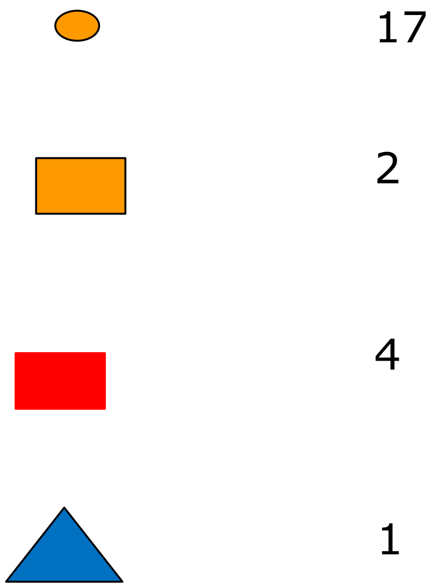
How many red squares can this man convert after two hours?

Mathematical Modeling

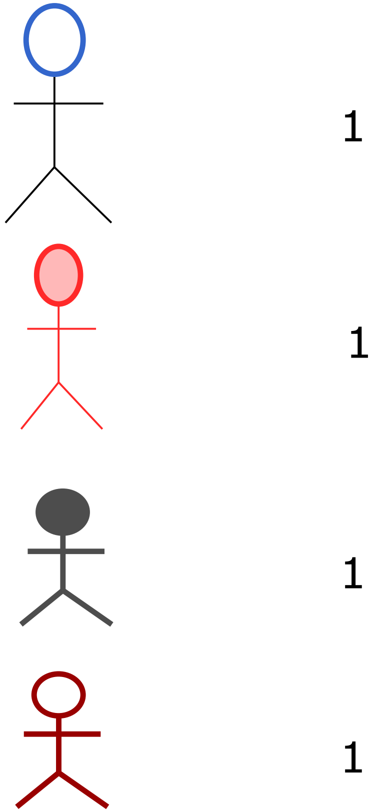
So If we have 25 circles



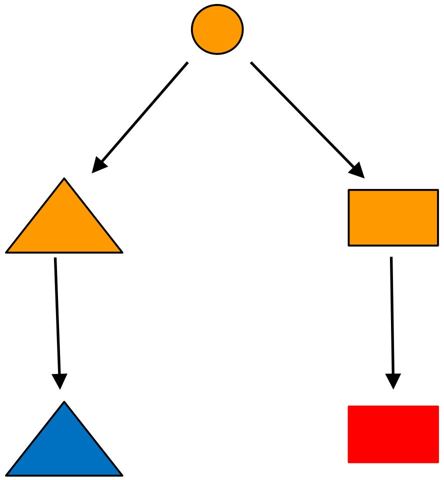
How manys objects do we have after two hours



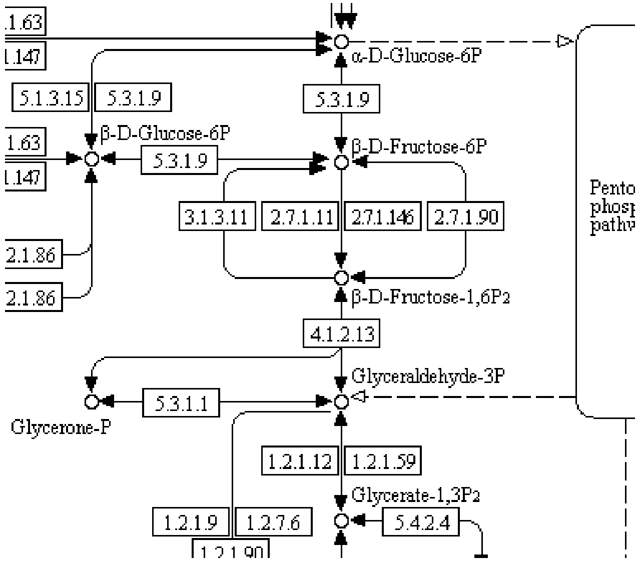
We can also estimate the wait time?



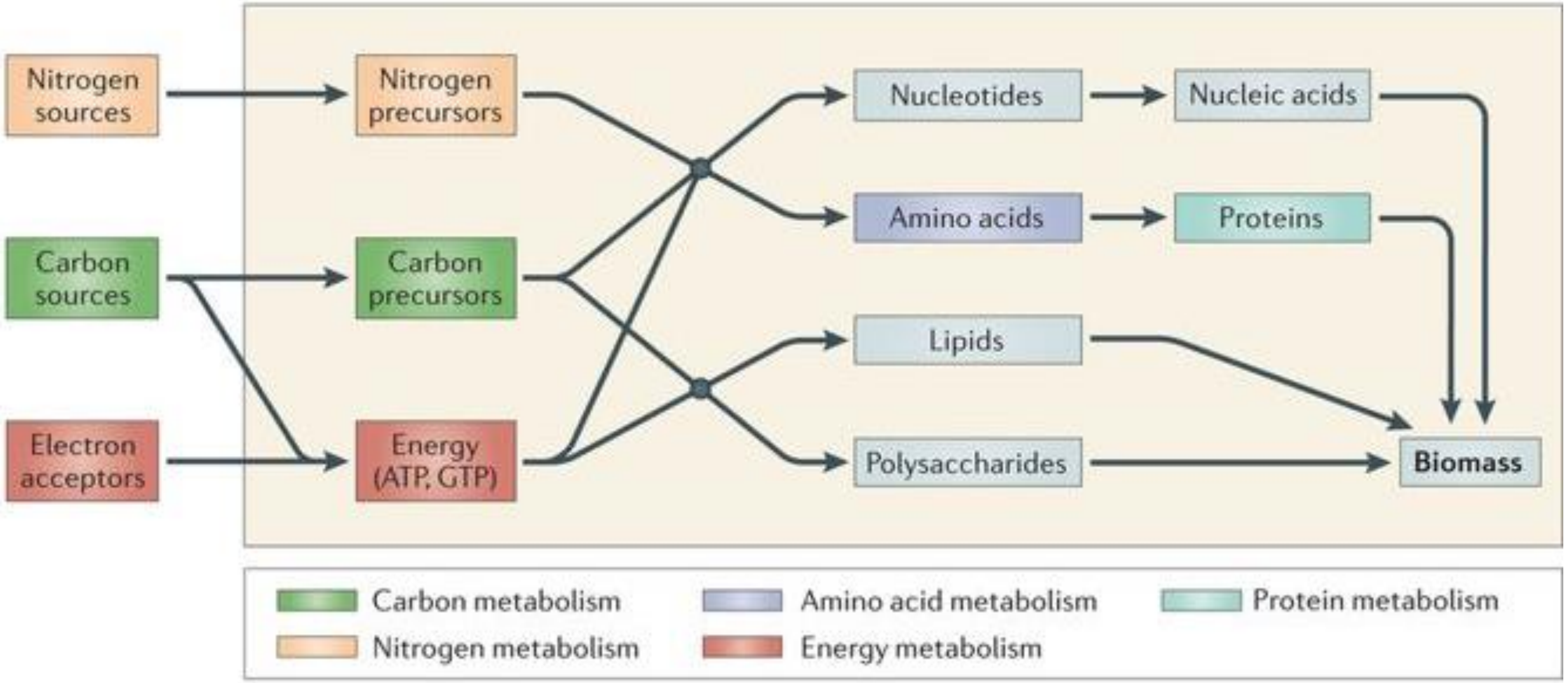
MATH



Pathway



Bio



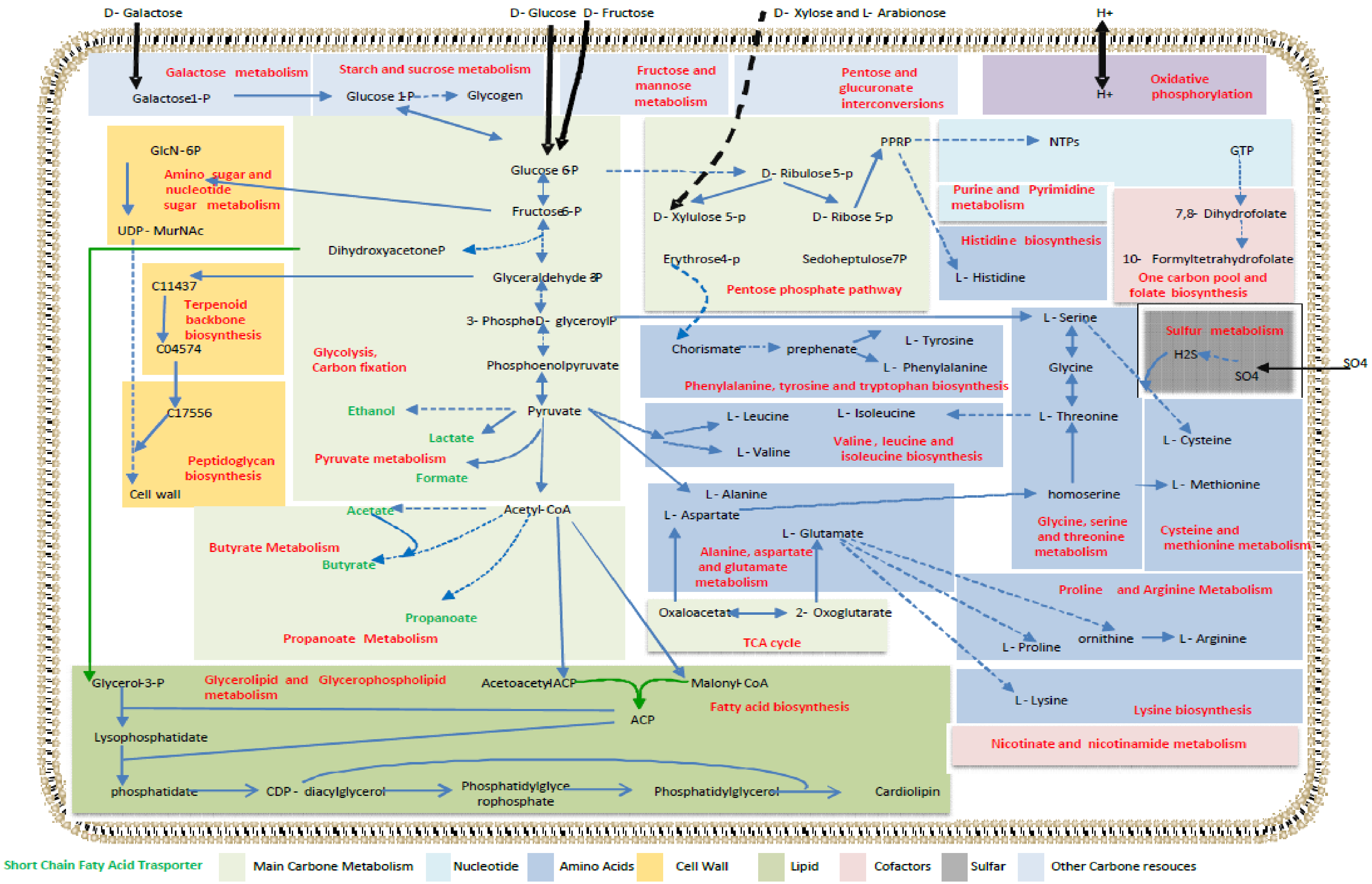
Nat Rev Microbiol. 2014 May;12(5):327-40.

We can estimate how many copies of nucleotides, amino acids, lipid species, and carbohydrates per cell

↓

Biomass Reaction

Genome-Scale Metabolic Models



If we know the uptake rate of glucose,

Can we estimate the growth rate?

For example,

yeast can consume glucose with the rate 13 * 6.023 x 10^8 molecules per cell per hour,

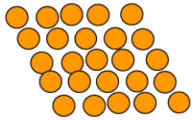
We know the copies numbers of nucleotides, amino acids, lipid species, and carbohydrates per yeast cell

The growth rate is 0.1 per hour

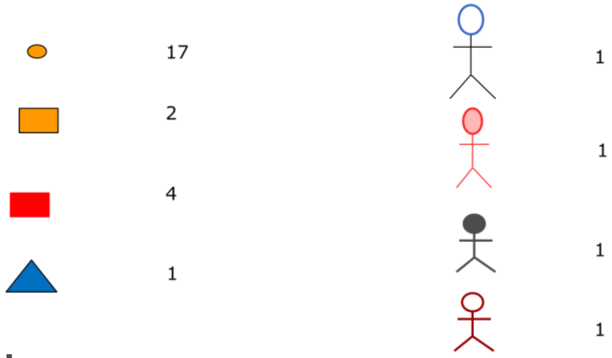
Quantitative description (Proteomics and Metabolomics)

Condition1 (normal or WT)

So If we have 25 circles



How manys objects do we have after two hours

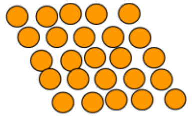


We can also estimate the wait time?

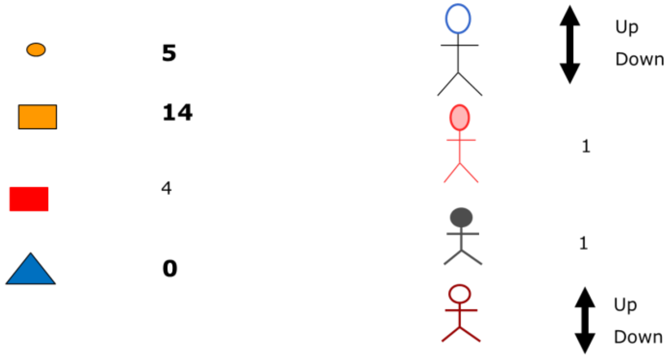
5

Condition2 (Disease or mutant)

If we have 25 circles

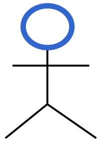


How manys objects do we have after two hours



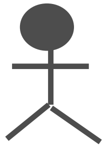
We can also estimate the wait time?

Please note the change in number of Circles and Squares.



Increase or decrease

Increase (up-regulated)



Increase or decrease

Decrease (Down-regulated)

Formal Definition of Systems Biology

- **to obtain new insight into the molecular mechanisms occurring in living cells or sub-systems of living cells, through the combination of mathematical modeling and experimental biology and**
- **to obtain a quantitative description of biological systems in the form of mathematical models that can be used for predictive analysis.**

Mathematical Modeling

Math:

Linear programming

Phyisics:

Thermodynamics,

Experimental biology

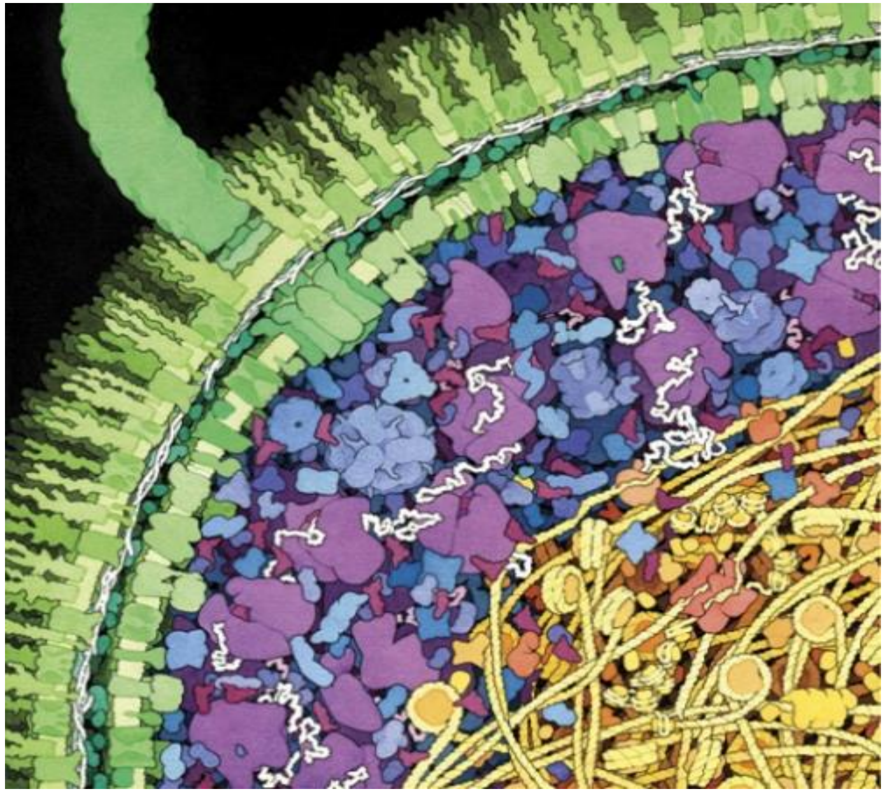
Fermentation
Enzyme Kinetics
13C flux

Quantitative description

Genomics,
RNA-seq,
Proteomics
Metabolomics

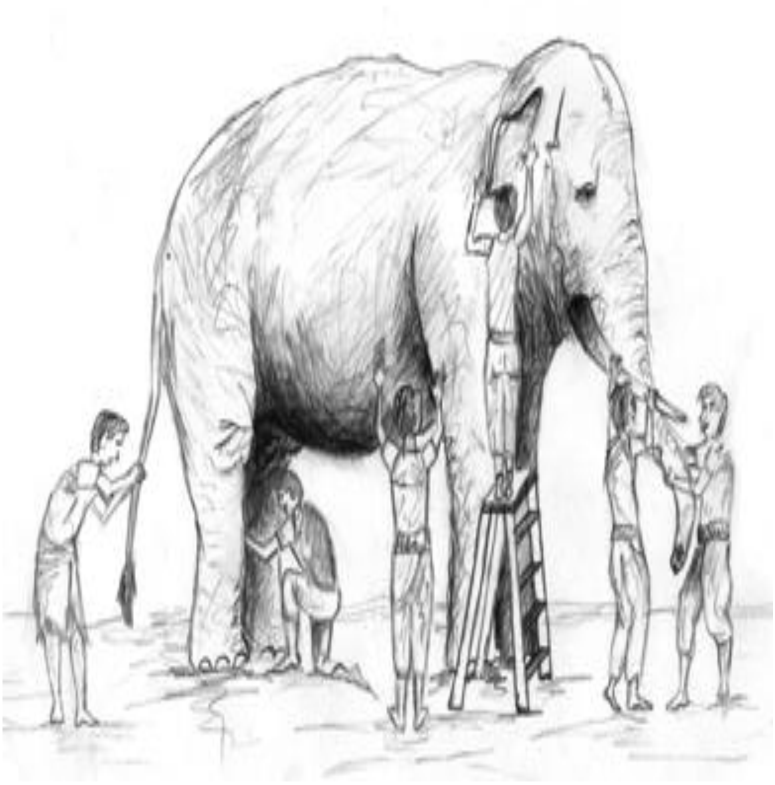
FEBS Letters Volume 583, Issue 24, 17

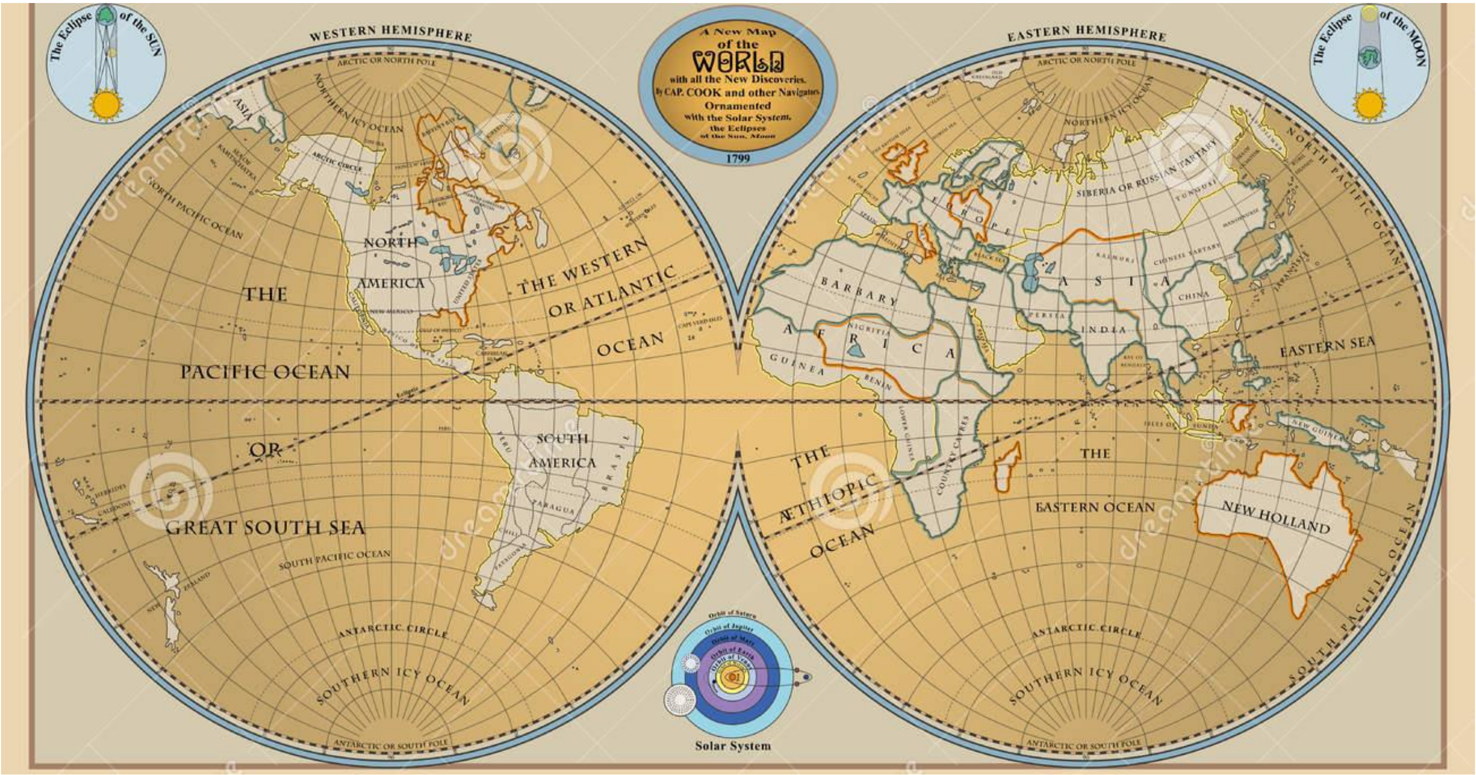
What is the aim of the Systems Biology?



The aim is to to find the cellular components and how these components are connected together.

We need genome-scale methods not small models





http://www.biokemi.org/biozoom/issues/514/articles/2285

https://www.dreamstime.com/stock-illustration-vector-old-globe-map-world-new-discoveries-image60583134

Math Background (Linear programming)

$$x + y - z = 0$$

$$x - y - z = 0,$$

x,y,z >=0

$$\xrightarrow{}$$

S v = 0

$$\begin{bmatrix} 1 & 1 & -1 \\ 1 & -1 & -1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

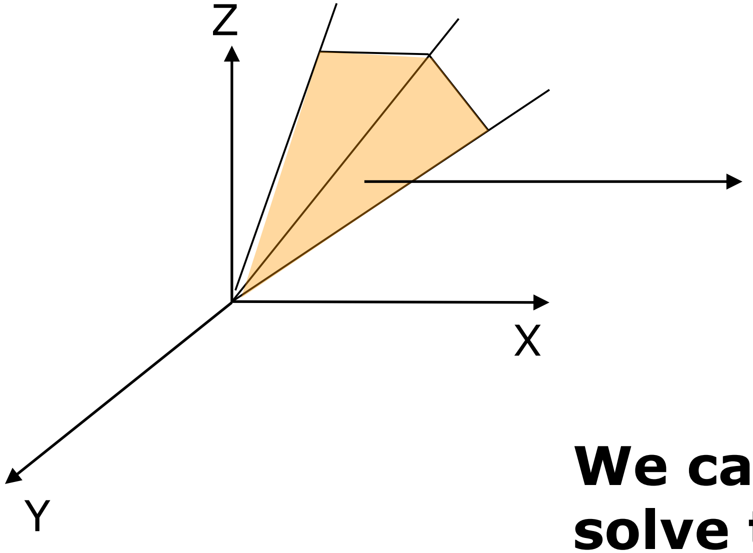
The problem
we need a point that maximizes z

$$\text{Max z}$$

$$s.t.$$

$$S\,v = 0,$$

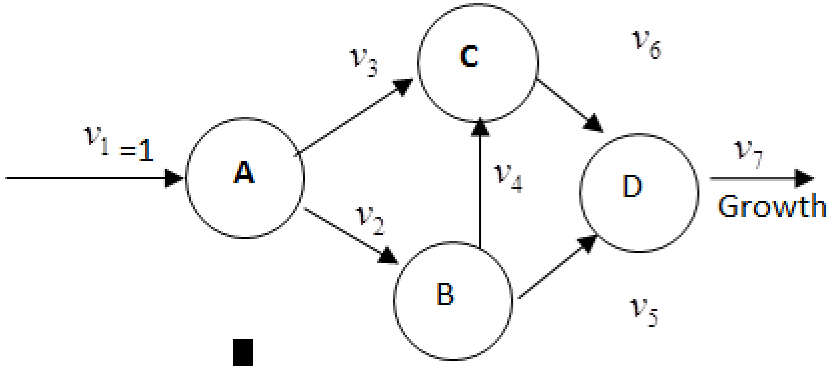
$$0 \le v \le 100$$



Solution space
each point is a solution
for this system

We can use the computer to solve this problem

Simulation: Flux balance anaylsis



$$\frac{dA}{dt} = v_1 - v_2 - v_3 \ = 0$$

$$\frac{dB}{dt} = v_2 - v_4 - v_5 \ = 0$$

$$\frac{dC}{dt} = v_3 + v_4 - v_6 \ = 0$$

$$\frac{dD}{dt} = v_3 + v_5 - v_7 \ = 0$$

$$\begin{bmatrix} \frac{dA}{dt} \\ \frac{dB}{dt} \\ \frac{dC}{dt} \\ \frac{dD}{dt} \end{bmatrix} = \underbrace{\begin{bmatrix} 1 & -1 & -1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & -1 & -1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & -1 \end{bmatrix}}_{\textbf{Stoichiometric Matrix } S} \begin{bmatrix} v_1 \\ v_2 \\ v_3 \\ v_4 \\ v_5 \\ v_6 \\ v_7 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

$$Max \ \ Growth = v_7$$

$$s.t.$$

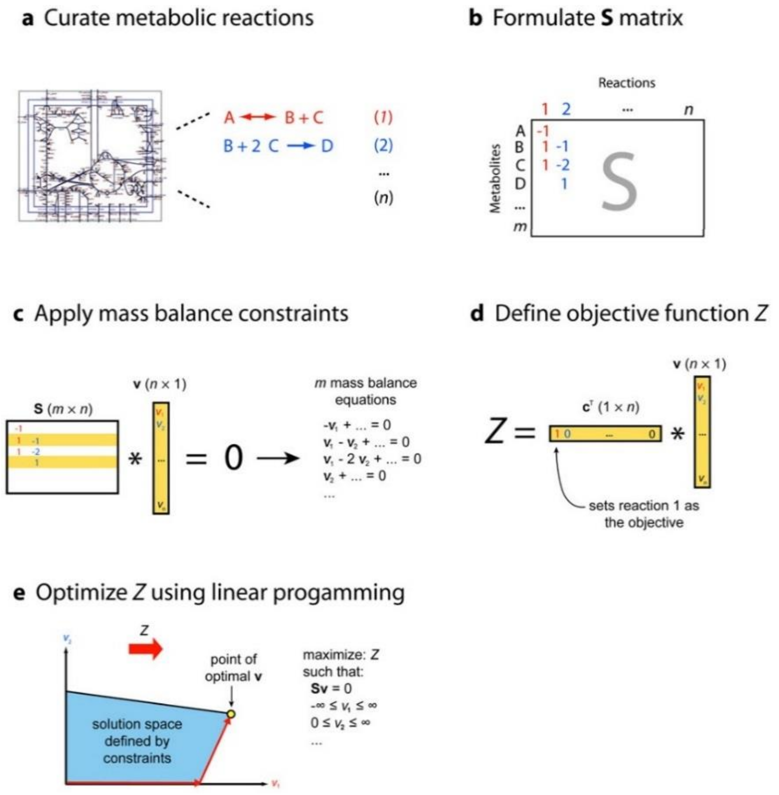
$$SV = 0$$

$$-100 \le V \le 100,$$

$$v_1 = 1$$

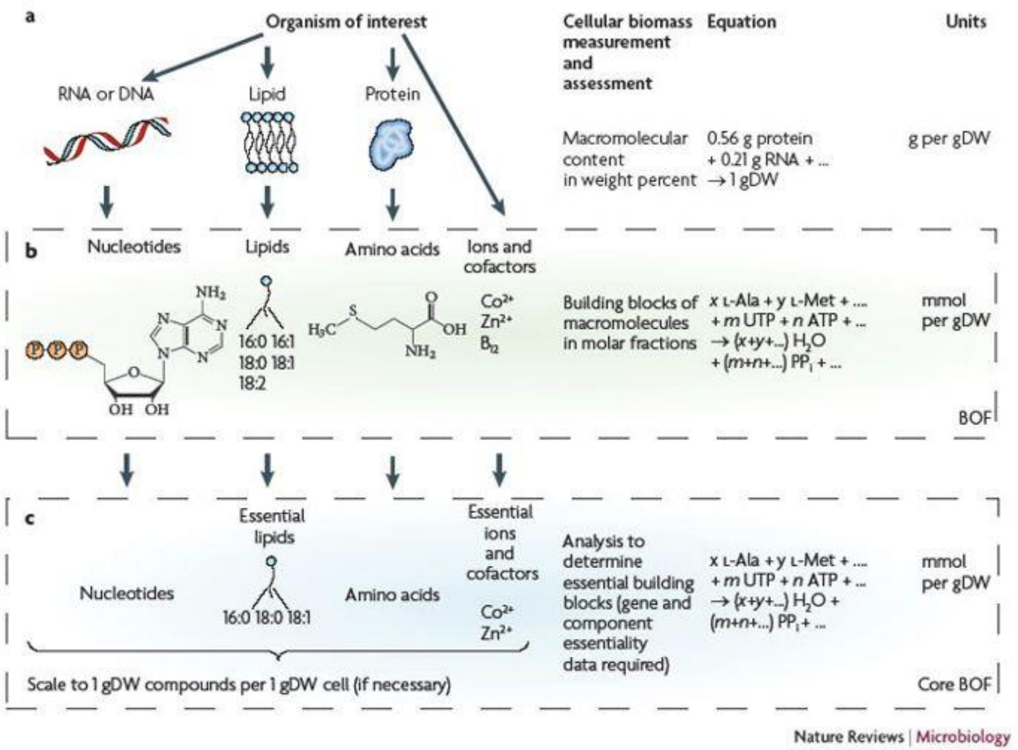
What is Flux Balance Anaylsis (FBA)

Chemical reactions



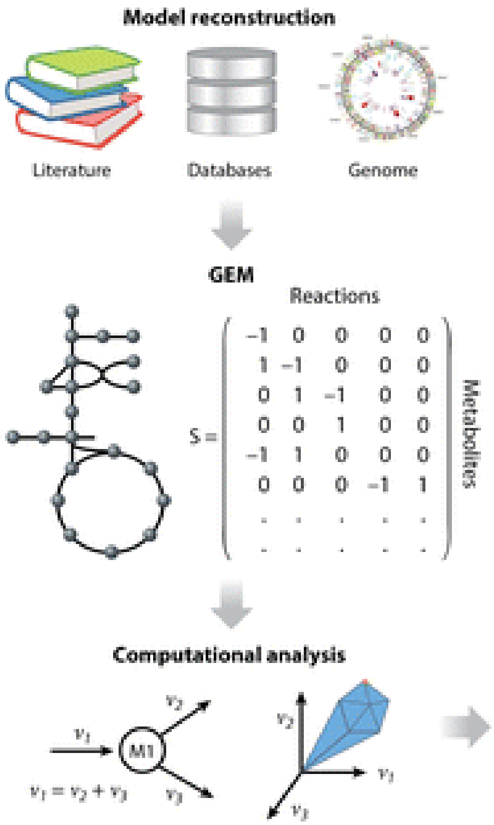
[Nat Biotechnol. 2010 Mar; 28(3): 245–248](#)

Biomass reaction



Nature Reviews Microbiology **volume7**, pages129–143 (2009)

How can you reconstruct the model



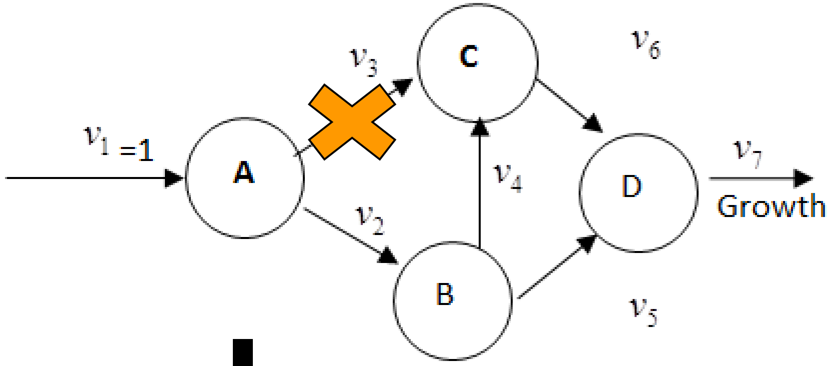
Nielsen J. 2017.
Annu. Rev. Biochem. 86:245–75

Reaction	Genes
S1 + S2 → P1 + P2	g1
S1+ S3 →P3 + P4	g1 and g2
S3 → P5	g1 or g3
…	…

Please remember this rule:

- You input the glucose uptake rate as 1 mmol/gDW/h
- FBA maximizes the growth rate
- The model predicts the growth rate

Simulation: Flux balance anaylsis



$$\begin{aligned}
&Max\ \ Growth = v_7 \\
&s.t. \\
&SV = 0 \\
&-100 \le V \le 100, \\
&v_1 = 1
\end{aligned}$$

$$\frac{dA}{dt} = v_1 - v_2 - v_3 \ = 0$$

$$\frac{dB}{dt} = v_2 - v_4 - v_5 \ = 0$$

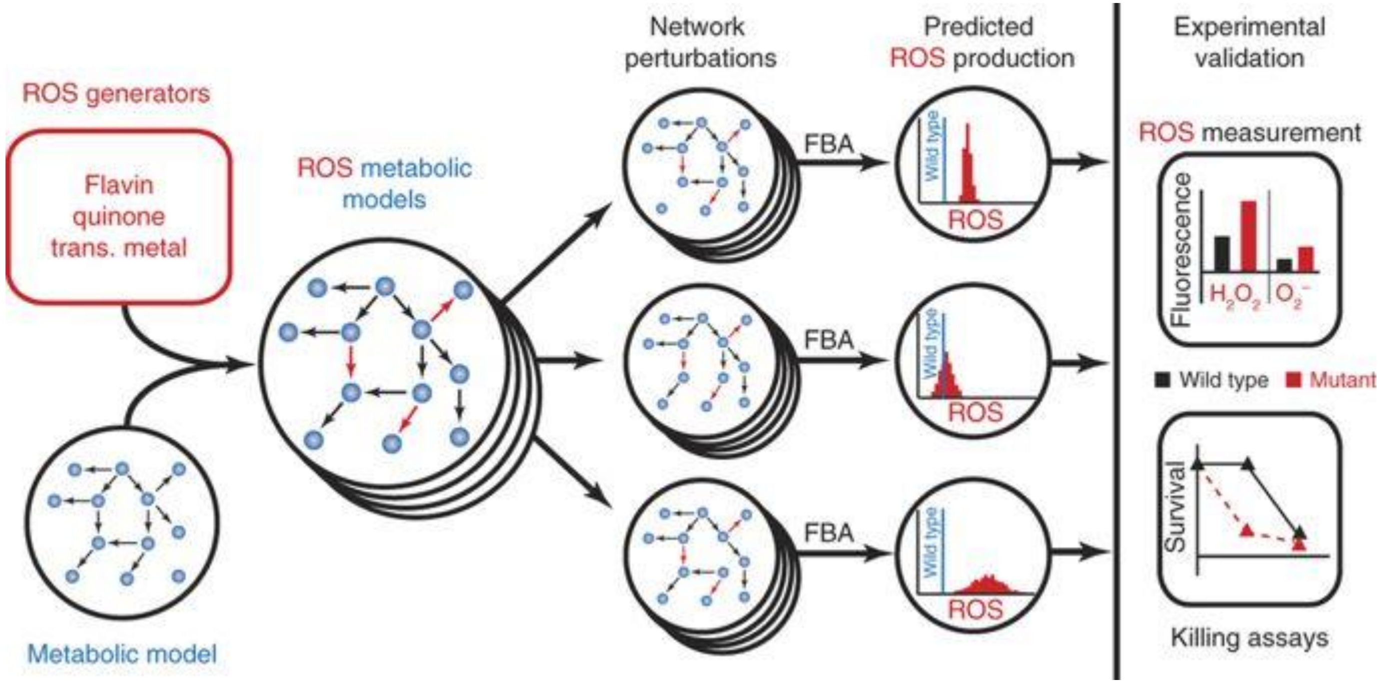
$$\frac{dC}{dt} = v_3 + v_4 - v_6 \ = 0$$

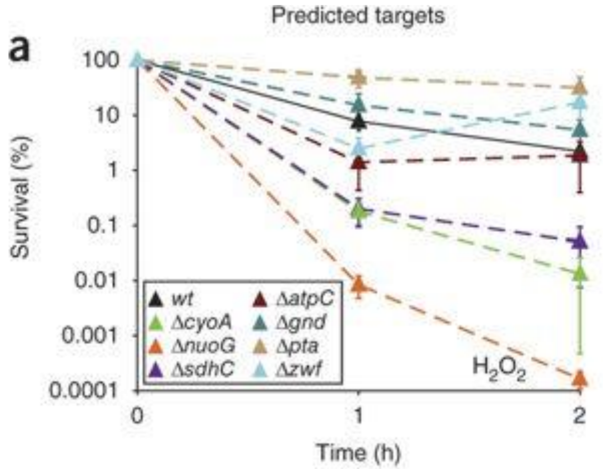
$$\frac{dD}{dt} = v_3 + v_5 - v_7 \ = 0$$

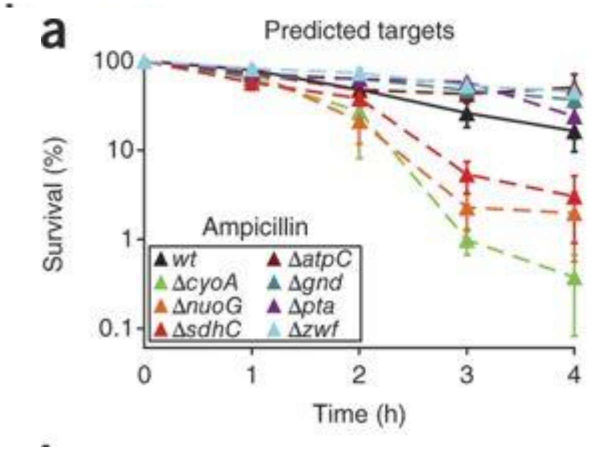
$$\begin{bmatrix} \frac{dA}{dt} \\ \frac{dB}{dt} \\ \frac{dC}{dt} \\ \frac{dD}{dt} \end{bmatrix} = \underbrace{\begin{bmatrix} 1 & -1 & -1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & -1 & -1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & -1 \end{bmatrix}}_{\textbf{Stoichiometric Matrix } S} \begin{bmatrix} v_1 \\ v_2 \\ v_3 \\ v_4 \\ v_5 \\ v_6 \\ v_7 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

Finding antibiotics (deleting one reaction or gene)

Reactive oxygen species (ROS) such as H2O2 can kill bacteria, can we attack the resisentce machanism in bacteria

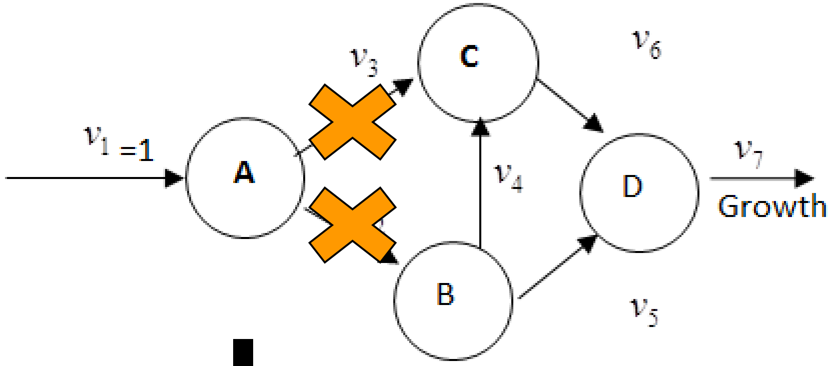






Nature Biotechnology **volume31**, pages160–165 (2013)

Simulation: Flux balance anaylsis



$$\frac{dA}{dt} = v_1 - v_2 - v_3 = 0$$

$$\frac{dB}{dt} = v_2 - v_4 - v_5 = 0$$

$$\frac{dC}{dt} = v_3 + v_4 - v_6 = 0$$

$$\frac{dD}{dt} = v_3 + v_5 - v_7 = 0$$

$$\begin{bmatrix} \frac{dA}{dt} \\ \frac{dB}{dt} \\ \frac{dC}{dt} \\ \frac{dD}{dt} \end{bmatrix} = \begin{bmatrix} 1 & -1 & -1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & -1 & -1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & -1 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ v_3 \\ v_4 \\ v_5 \\ v_6 \\ v_7 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

Stoichiometric Matrix S

$$Max\ \ Growth = v_7$$

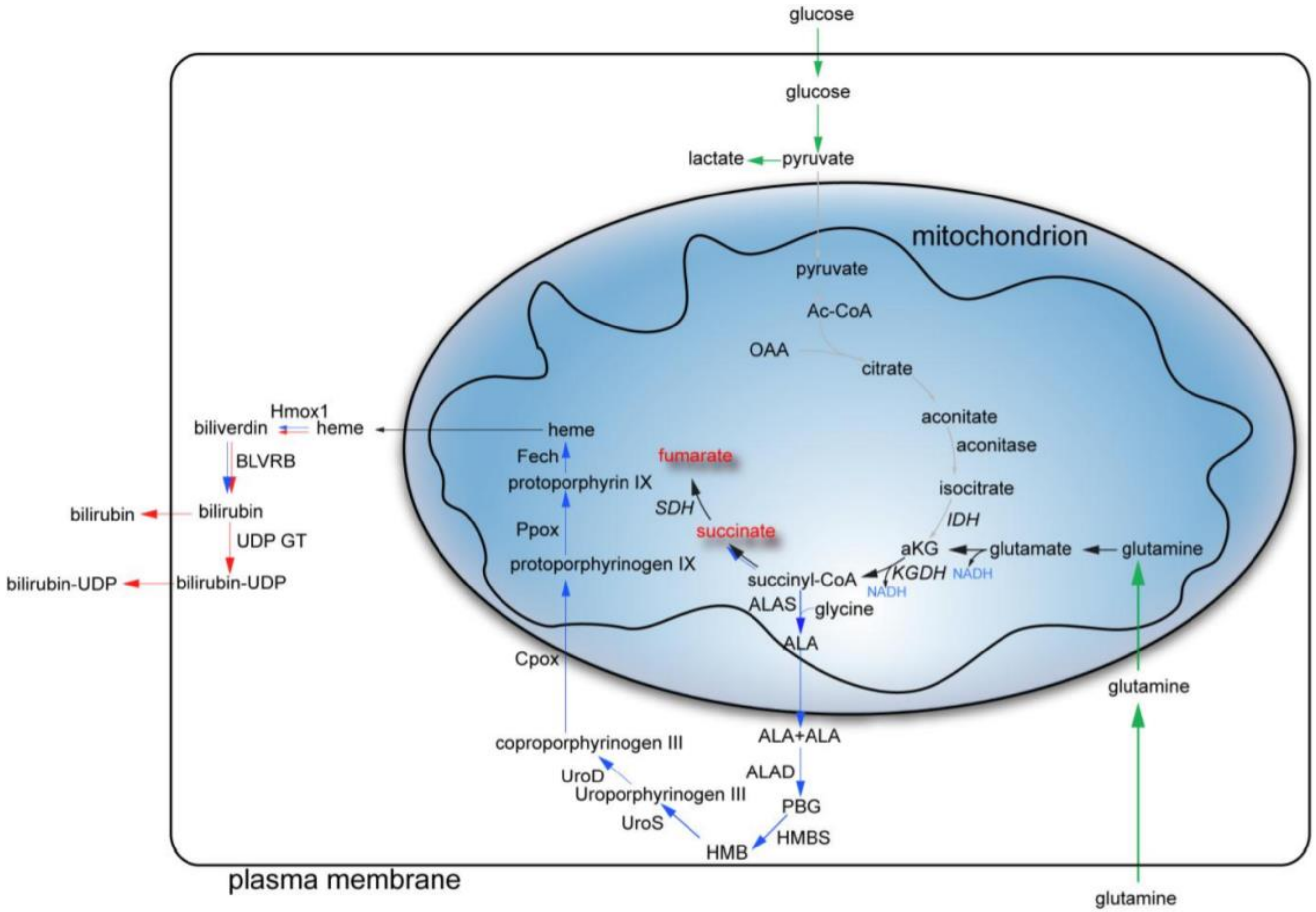
$s.t.$

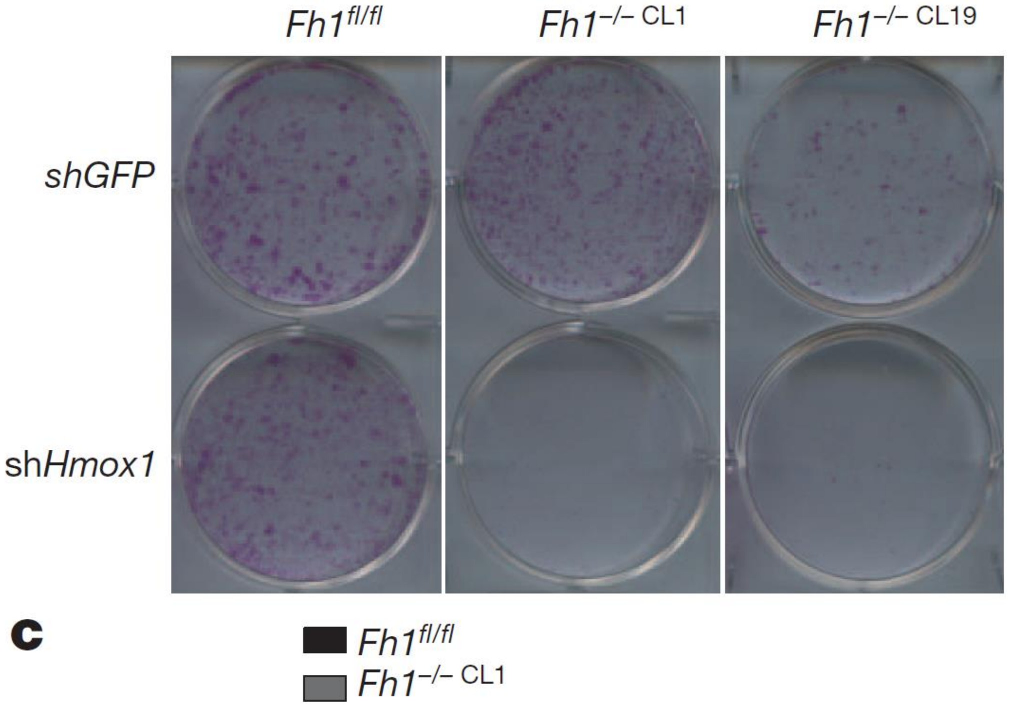
$$SV = 0$$

$$-100 \le V \le 100,$$

$$v_1 = 1$$

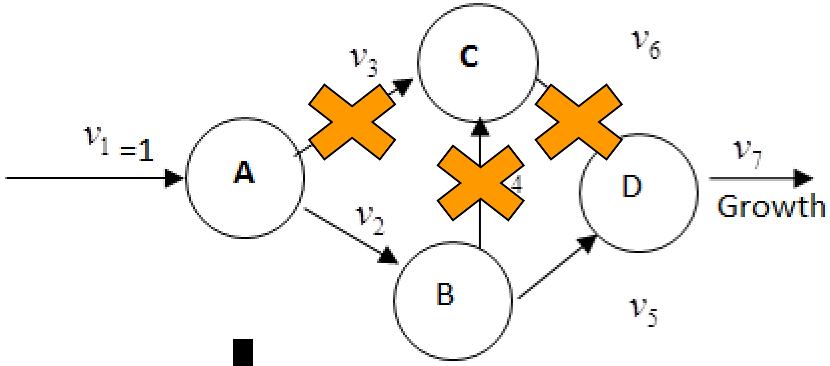
GEM Applications: Deleting one reaction (gene)





Nature **volume477**, pages225–228 (08 September 2011)

Anti-metabolites: Deleting all reactions around one metabolite



$$\frac{dA}{dt} = v_1 - v_2 - v_3 = 0$$

$$\frac{dB}{dt} = v_2 - v_4 - v_5 = 0$$

$$\frac{dC}{dt} = v_3 + v_4 - v_6 = 0$$

$$\frac{dD}{dt} = v_3 + v_5 - v_7 = 0$$

$$\begin{bmatrix} \frac{dA}{dt} \\ \frac{dB}{dt} \\ \frac{dC}{dt} \\ \frac{dD}{dt} \end{bmatrix} = \underbrace{\begin{bmatrix} 1 & -1 & -1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & -1 & -1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & -1 \end{bmatrix}}_{\text{Stoichiometric Matrix } S} \begin{bmatrix} v_1 \\ v_2 \\ v_3 \\ v_4 \\ v_5 \\ v_6 \\ v_7 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

$$Max\ \ Growth = v_7$$

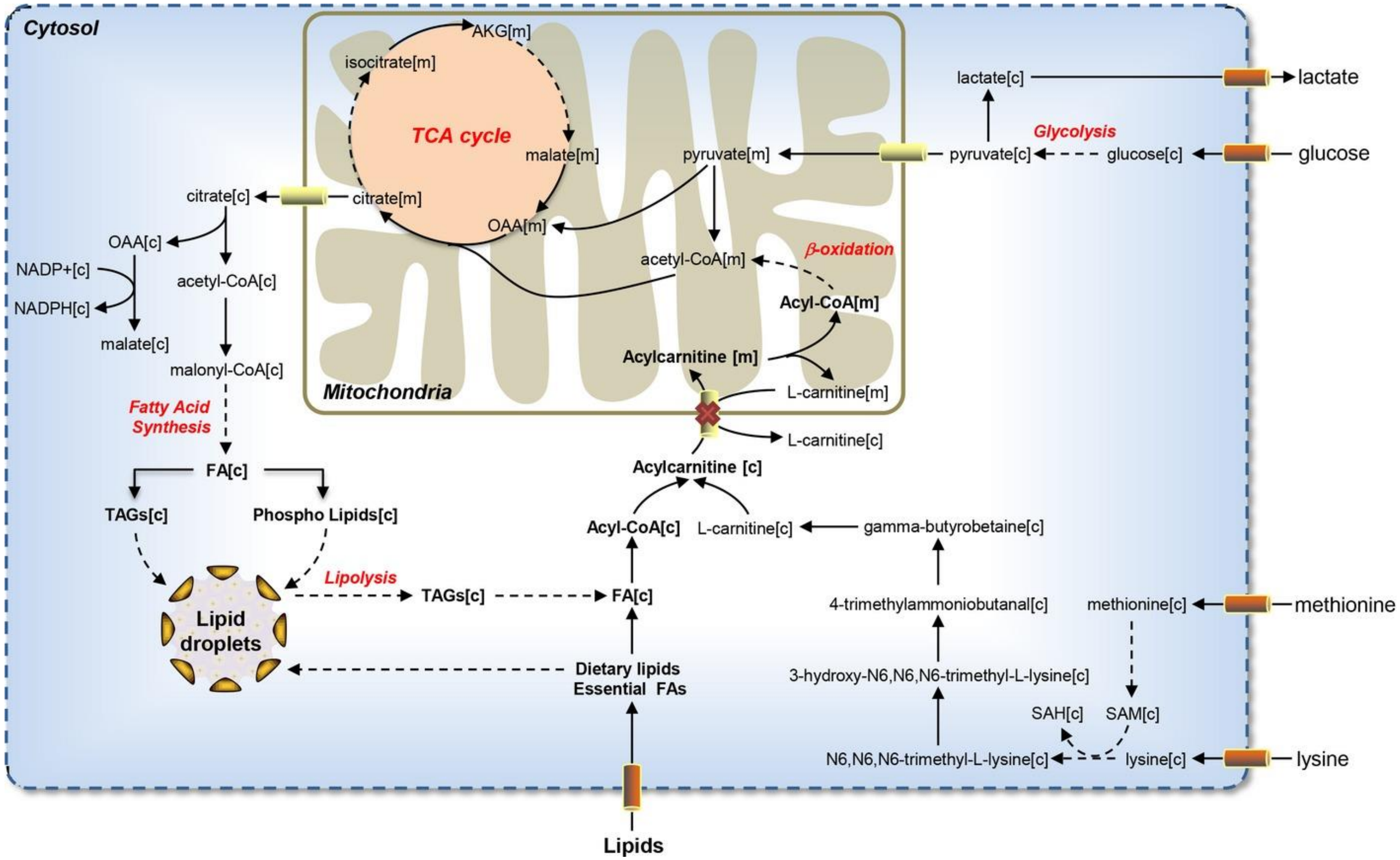
$$s.t.$$

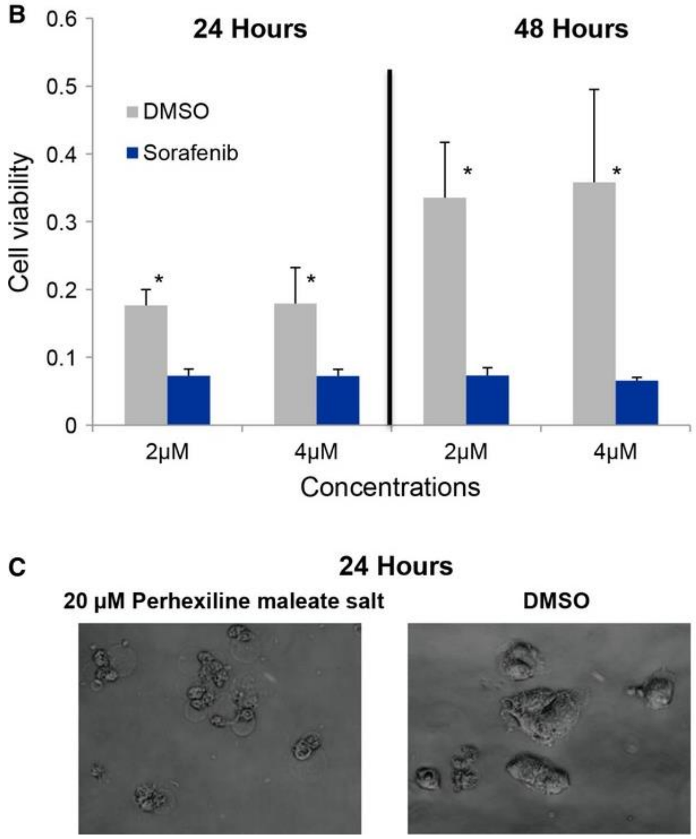
$$SV = 0$$

$$-100 \le V \le 100,$$

$$v_1 = 1$$

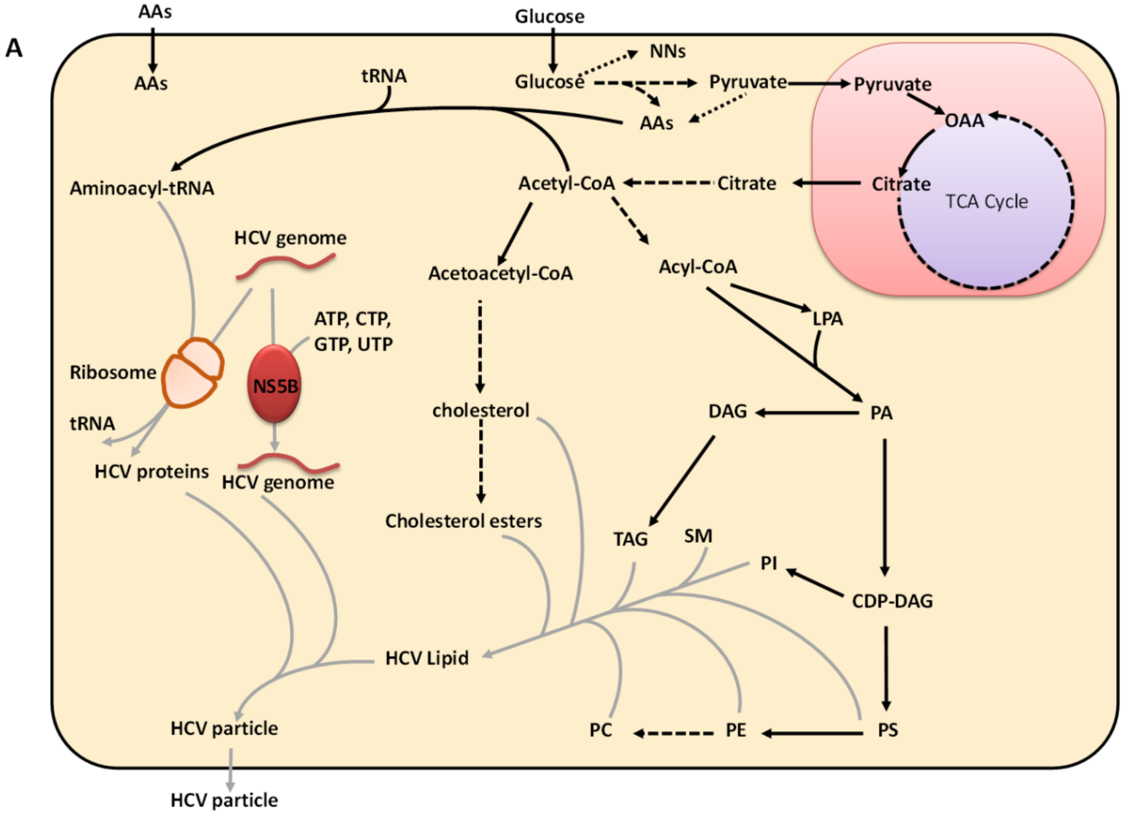
Anti-metabolite: Can the cell grow when all reaction around one metabolite are deleted





Rasmus Agren et al MSB 2014

HCV



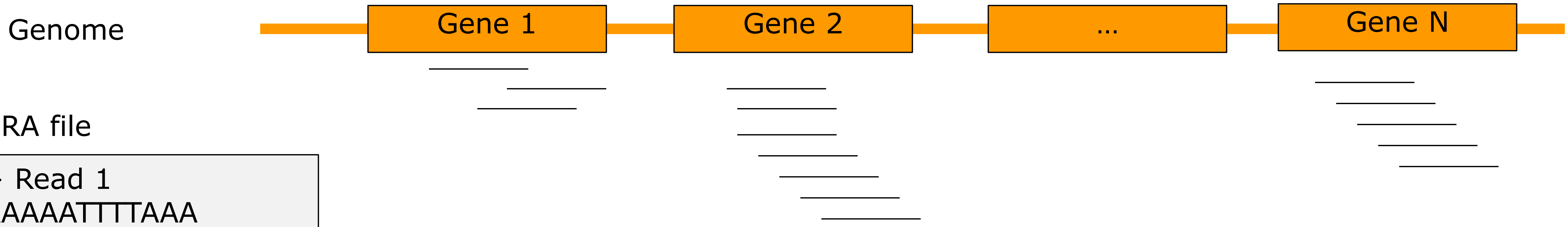
Host-targets	Summary	Ref
Acyl-CoA	Liefhebber et al downregulated the acyl-CoA synthetase long-chain family member 1 gene to reduce HCV assembly.	(3)
Acyl-glycerol		
cholesterol esters		
(R)-5-diphosphomevalonate	Jin Ye reviewed the role of mevalonate pathway in HCV assembly.	(4)
(R)-5-phosphomevalonate		
(R)-mevalonate		
Presqualene-PP	Saito et al downregulated the squalene synthase gene to reduce HCV assembly.	(5)
Squalene		
Squalene 2,3-oxide	Owens et al suggested that both of oxidosqualene cyclase and lanosterol demethylase can be used as antiviral targets for HCV assembly.	(6)
lanosterol		
dimethylallyl-PP	Ye et al and Kapadia and Chisari reported the role of geranylgeranylation in HCV replication.	(7, 8)
farnesyl-PP		
geranyl-PP		
isopentenyl-pPP		
cholesterol	Takano et al and Rodgers et al downregulated the DHCR24 and DHCR7 genes to reduce HCV assembly.	(9, 10)
inositol	Bishé et al reviewed the role of phosphoinositides in the HCV assembly.	(11)
1D-myo-inositol-3-phosphate		
3-dehydrosphinganine	Sakamoto et al downregulated the serine palmitoyltransferase gene to reduce HCV assembly.	(12)
sphinganine		
sn-glycerol-3-phosphate	Sn-glycerol-3-phosphate may prevent lipid biosynthesis through Kennedy pathway.	
1,2-diacylglycerol-LD-TAG pool (DAG)	Herker et al downregulated the DGAT1 gene to reduce HCV assembly.	(13)

Omics Analysis

Condition 1 Condition 2

	P1	P2	…	Pm	N1	N2	N3	…	Nz	Fold Change	pvalue	Adjust pvalue
Gene1	1	1	1	1	10	10	10	10	10	=log2 average P / average N	t-test (P,N)	0.004
Gene 2	30	30	30	30	1	1	1	1	1			
Gene 3	20	20	20	20	19	19	19	19	19			
…												
Gene N												

What is the biological meaning of these numbers



SRA file

> Read 1
AAAAATTTTAAA

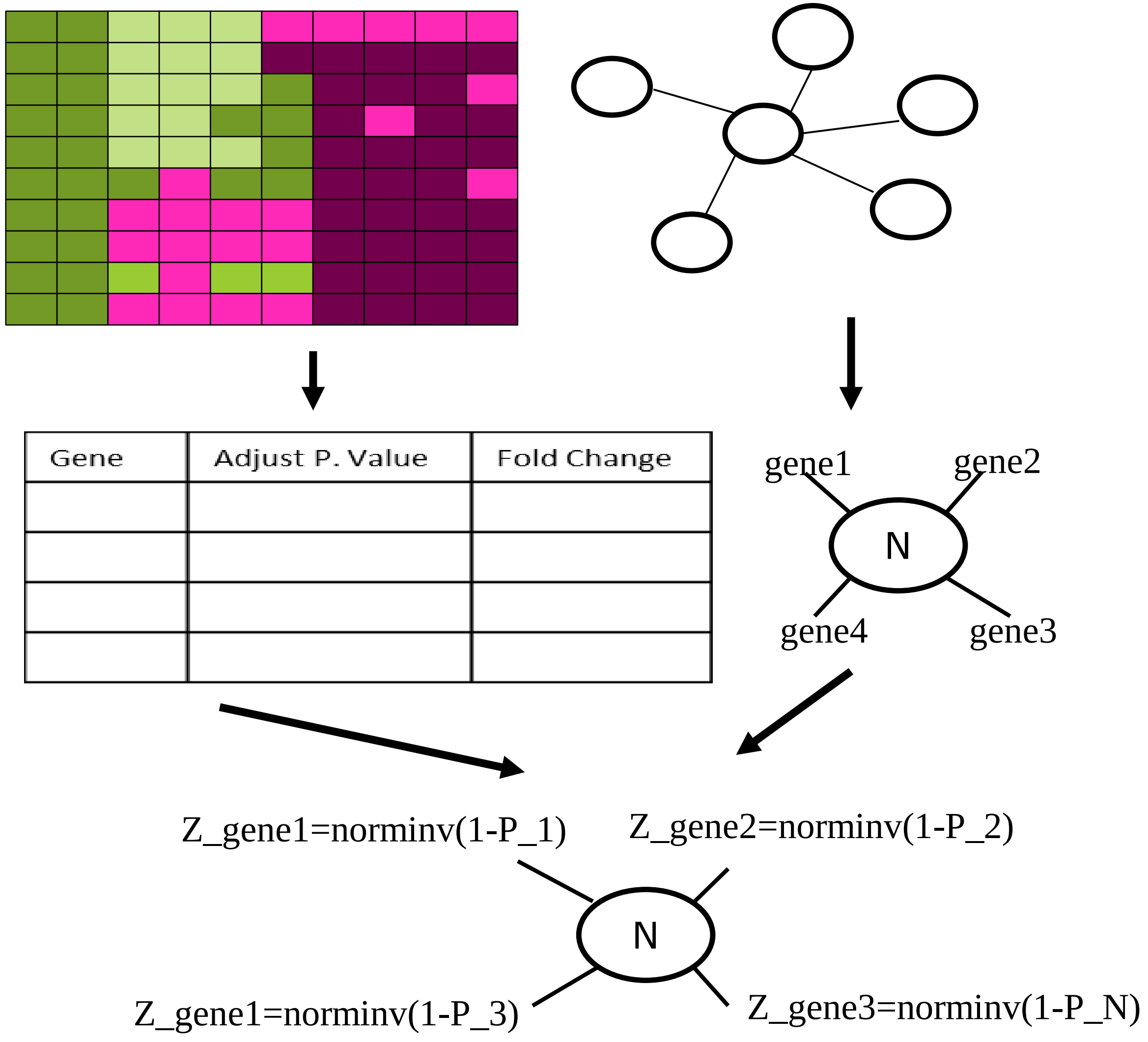
>Read 2
AAAACCCCCCC

>Read 3
AAAACCCCCCC

>Read 4
AAAACCCCCCC

RNA-seq
these numbers are number of reads

Reporter Metabolites and Subnetworks



Gene	Adjust P. Value	Fold Change

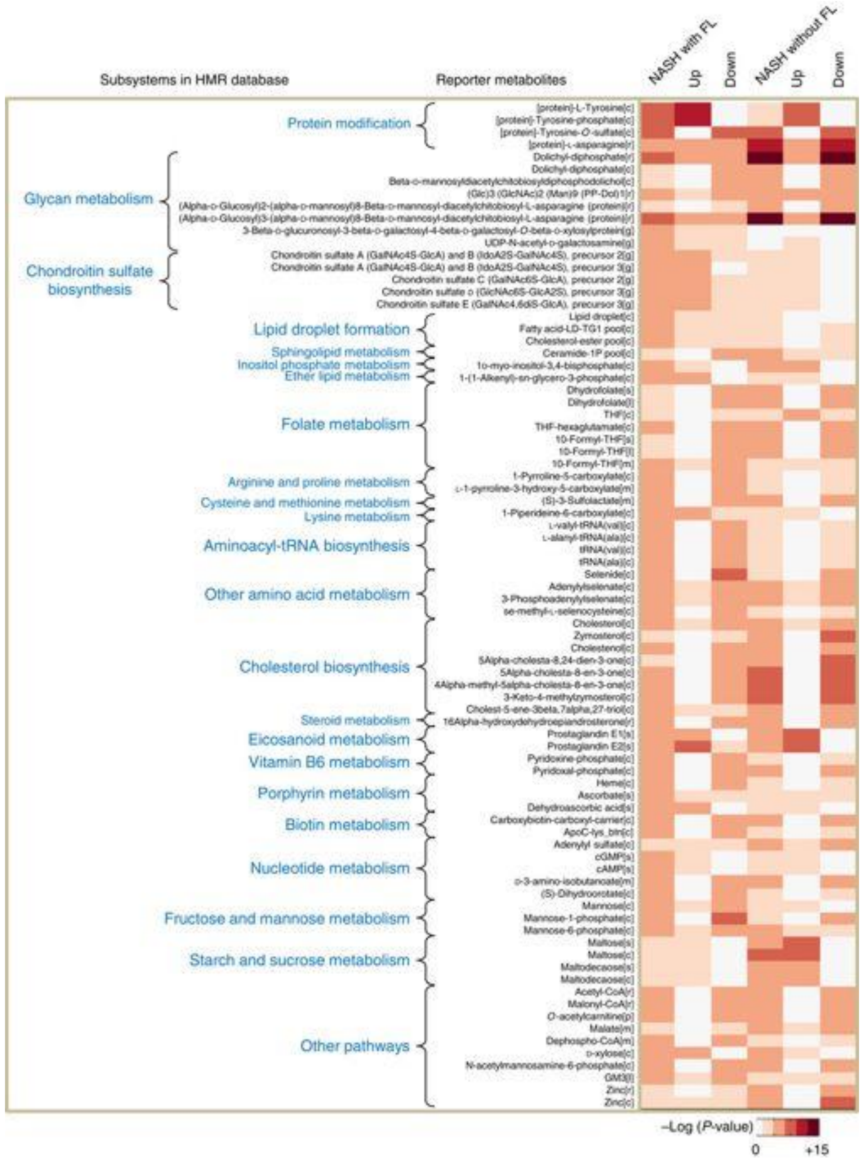
Z_gene1=norminv(1-P_1)

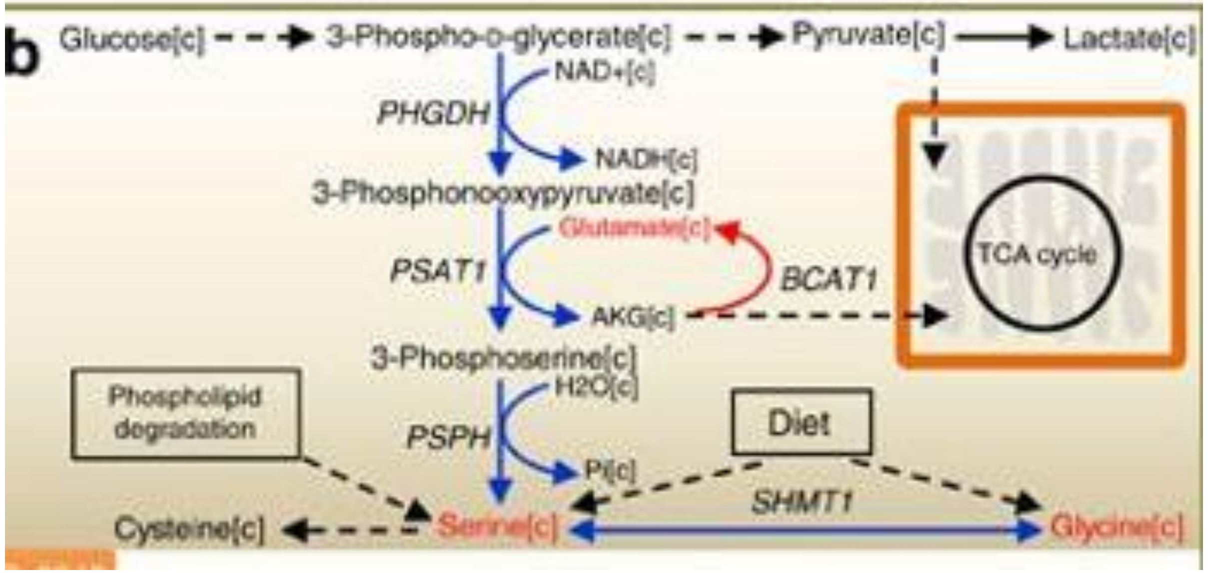
Z_gene2=norminv(1-P_2)

Z_gene1=norminv(1-P_3)

Z_gene3=norminv(1-P_N)

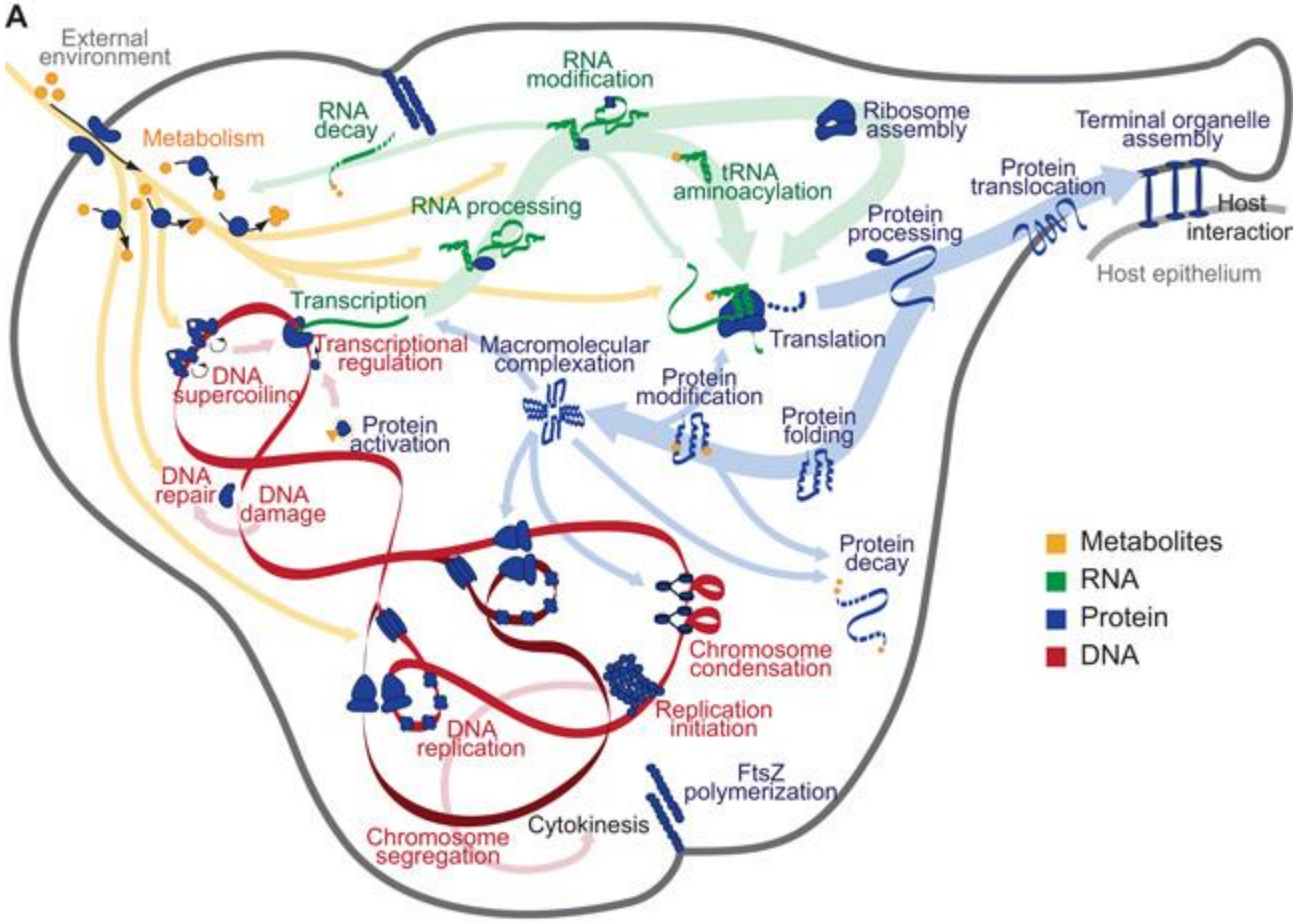
Non-alcoholic fatty liver disease





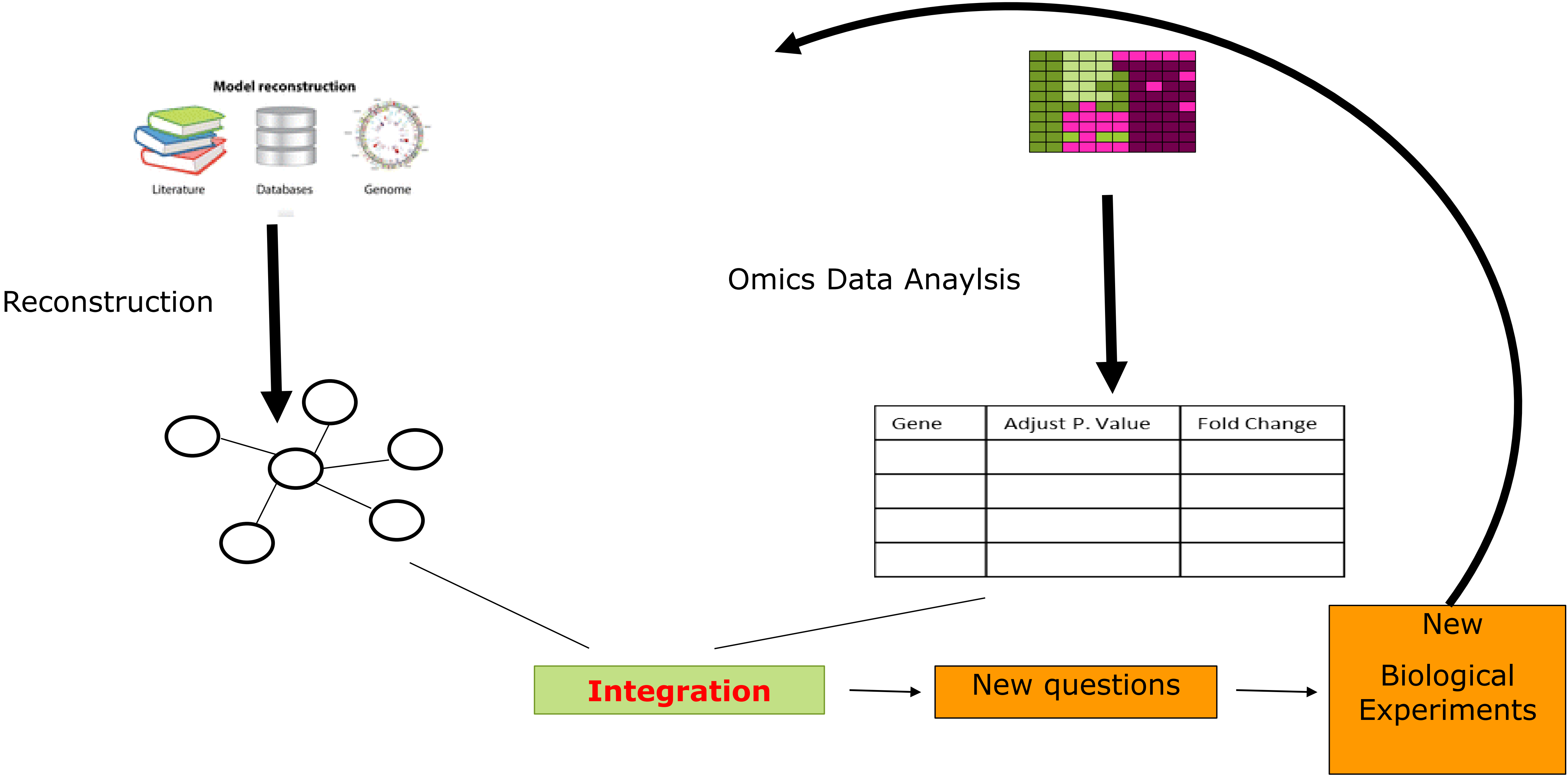
Nature Communications **volume5**, Article number: 3083 (2014)

Whole Cell Modeling



Cell. 2012 Jul 20; 150(2): 389–401.

Please Remember this Systems Biology Cycle



Reconstruction

Omics Data Anaylsis

Gene	Adjust P. Value	Fold Change

Integration

New questions

New

Biological Experiments

Omics Data and tools

https://cancergenome.nih.gov/

https://www.ebi.ac.uk/ena

https://galaxyproject.github.io/training-material/

http://kbase.us/

Genome-Scale Metabolic Models

http://biomet-toolbox.chalmers.se/

http://www.metabolicatlas.com/

http://sbrg.ucsd.edu/optimizing-genres/

https://vmh.uni.lu/

More Free Courses

https://www.youtube.com/watch?v=nHkYlEWNxeM

https://www.edx.org/course/case-studies-functional-genomics-harvardx-ph525-7x-0

https://www.video.ethz.ch/lectures/d-math/2014/spring/402-0112-14L/015a0d57-9372-4635-868c-342c0a1e61d6.html?autoplay=true

Acknowledgements

- Prof. Jens Nielsen
- Egypt Scholars

  Zeinab Hefny