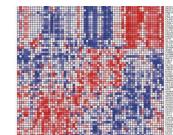
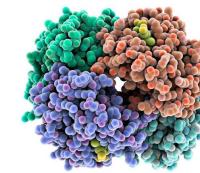
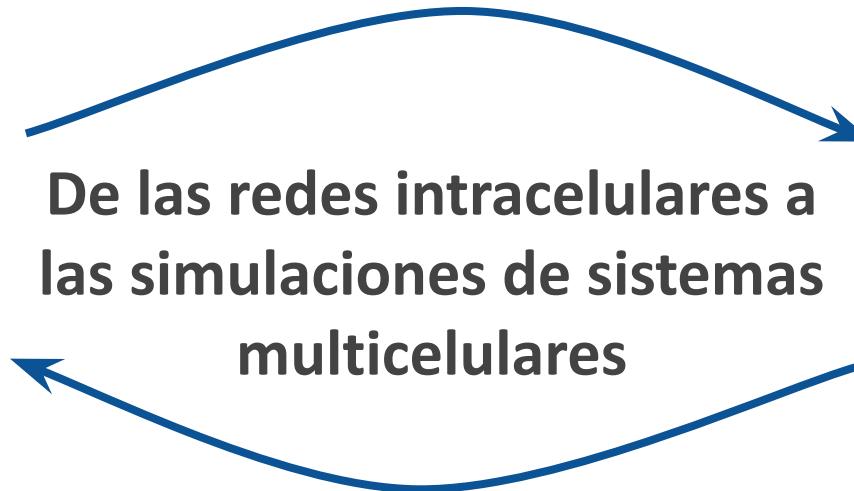
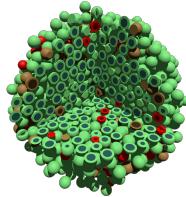
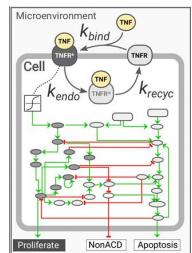
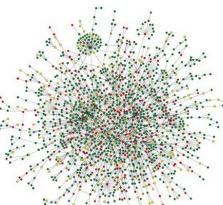




Barcelona  
Supercomputing  
Center

Centro Nacional de Supercomputación

# Curso PEDECIBA Bioinformática, Montevideo 04/2024



Profesor: Dr. Miguel Ponce de León ([miguel.ponce@bs.es](mailto:miguel.ponce@bs.es)) - BSC

Coordinador: Dr. Flavio Pazos ([flavio.pazos@gmail.com](mailto:flavio.pazos@gmail.com)) - IIBCE/IP

Apoyan:



Barcelona  
Supercomputing  
Center

Centro Nacional de Supercomputación



# Tema 5

## Modelado metabólico

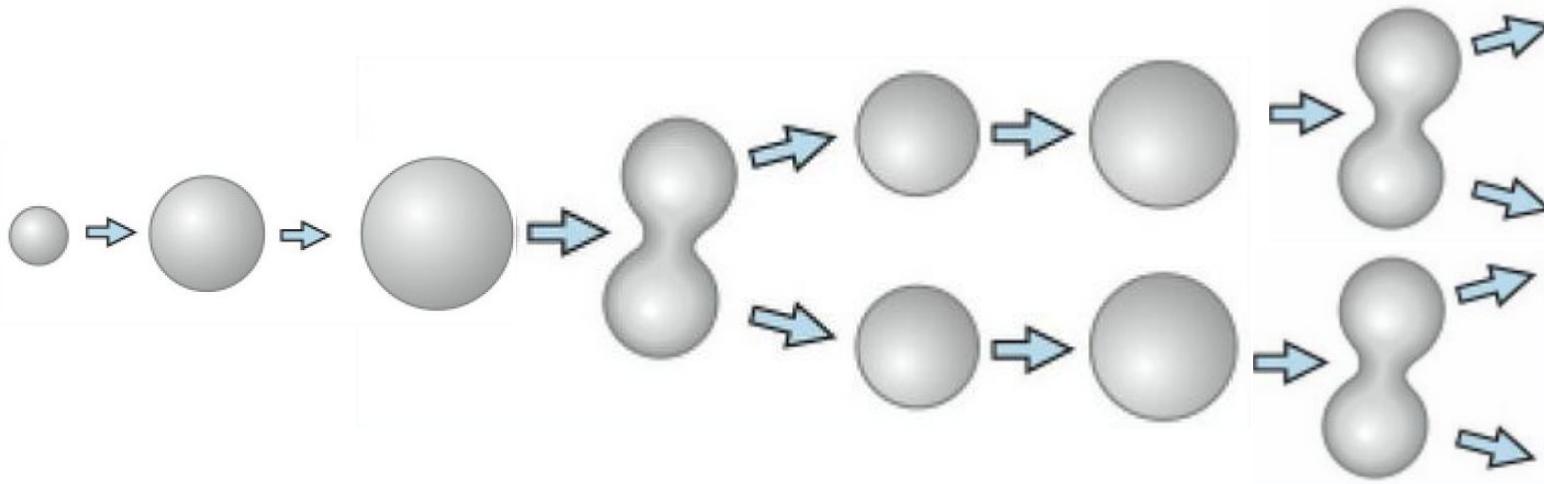
1. Redes metabólicas
2. Redes de flujos
3. Reconstrucción metabólica
4. Modelado Basado en Restricciones

# Lecture outline

- 1. Introduction**
  - a. A bit of history: from genomes to molecular catalogs of parts
  - b. Metabolic reconstruction
- 2. Constraint-based modeling**
  - a. Basic definitions
  - b. The flux space
  - c. Finding meaningful states
- 3. Tutorial Part 1**
  - a. Introduction to cobrapy
  - b. Basic Model Manipulation
- 4. Flux Balance Analysis**
  - a. What is the biomass?
  - b. Predicting growth rates
  - c. Predicting gene deletions
- 5. Tutorial Part 2**
  - a. Genome-scale modeling
  - b. Simulating metabolic models
- 6. Tutorial Part 3**
  - a. The gene protein reaction rules
  - b. Gene knockouts

# What is Life?

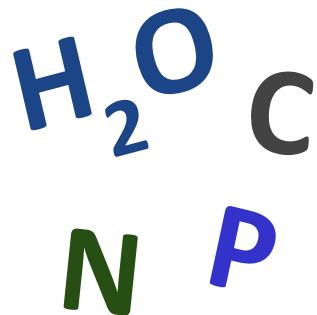
The definition of life is controversial. The current definition is that organisms are open systems that maintain homeostasis, are composed of cells, have a life cycle, undergo metabolism, can grow, adapt to their environment, respond to stimuli, reproduce and evolve.



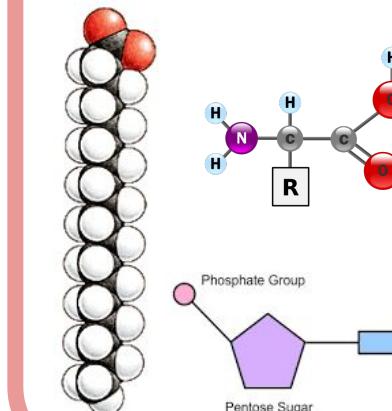
# What is a cell made of?

*Level (scale) of description*

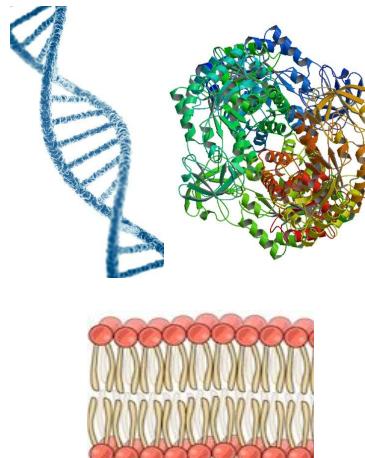
Chemistry



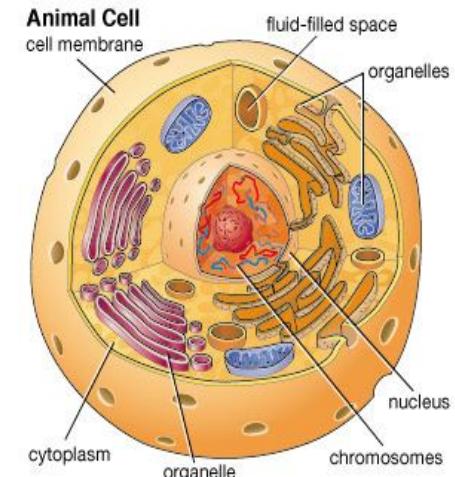
Building blocks



Macromolecules



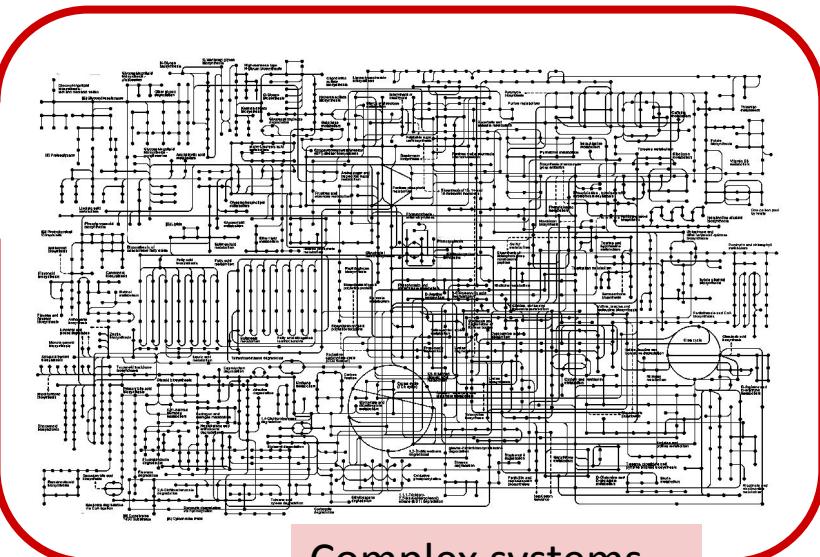
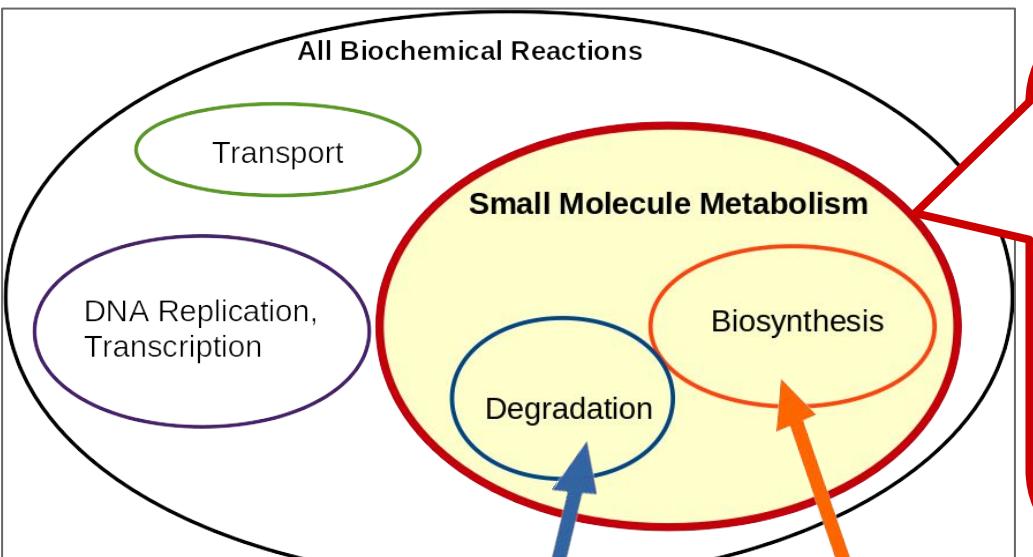
Cell



Cell's molecular factory: metabolism ← What is metabolism?

# Metabolism: the molecular factory of the cell

Is the **network** of biochemical reactions and transport processes that occur within a cell and allow **cell maintenance and growth**



- Generation of **energy (catabolism)** and **building block (anabolism)**
- Include the enzymatic reaction that act over small molecules

# Enzymes and Reactions

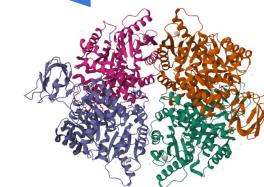
Pyruvate Kinase

EC number: 2.7.1.40

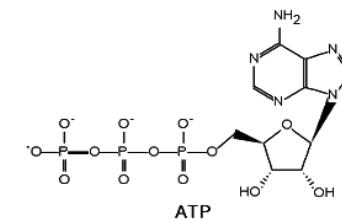
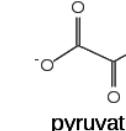
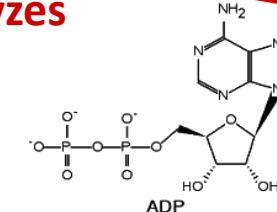
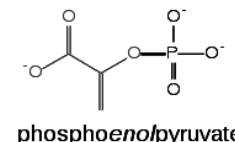
Coding genes: [PKM](#) or [PKLR](#)

Encodes

Isozymes →



Catalyzes



Reactants

Products

Most of biochemical reactions are **catalyzed** by a class of proteins called **enzymes** (and thus enzymes are encoded in one or more genes)

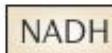
# Enzymes and Reactions

## Level 1: Metabolite specificity

Primary metabolites

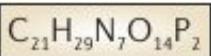
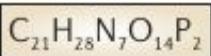
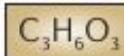


Coenzymes

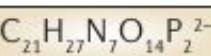
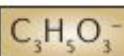


## Level 2: Metabolite formulae

Neutral formulae



Charged formulae



## Level 3: Stoichiometry



## Level 4: Thermodynamic considerations and/or directionality



## Level 5: Localization

Prokaryotes

[c]: cytoplasm

[e]: extracellular

[p]: periplasm

[n]: nucleus

[g]: golgi apparatus

[v]: vacuole

[l]: lysosome

[m]: mitochondria

[x]: peroxisome

[h]: chloroplast

[r]: endoplasmic reticulum

Eukaryotes



**1- Metabolites** (reactants and products)

**2- Stoichiometry** (molar ratio, mass conservation)

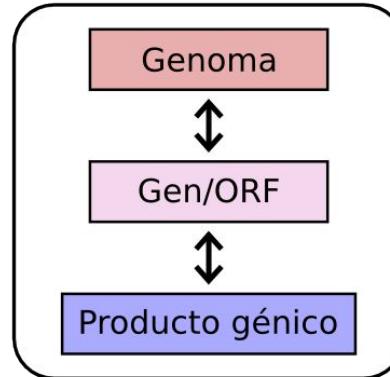
**3- Enzyme** (genes encoding the protein)

**4- Thermodynamics** (directionality)

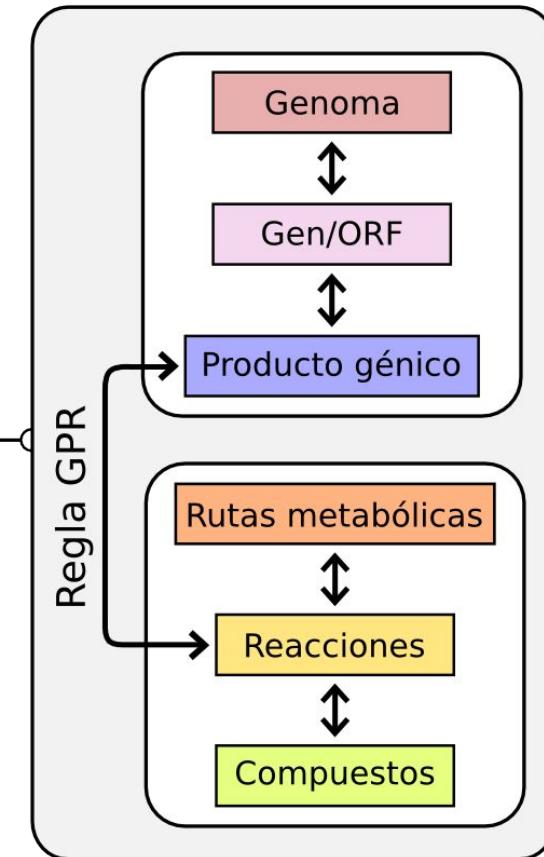
**5- Localization** (compartment)

# Pathway Genome Database

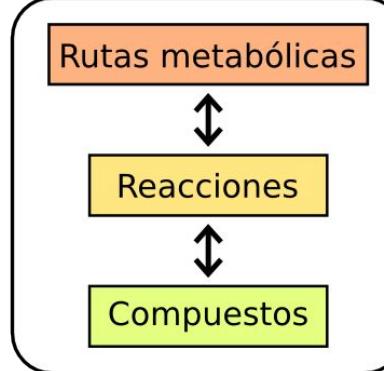
Anotación genómica



BD modelo-organismo



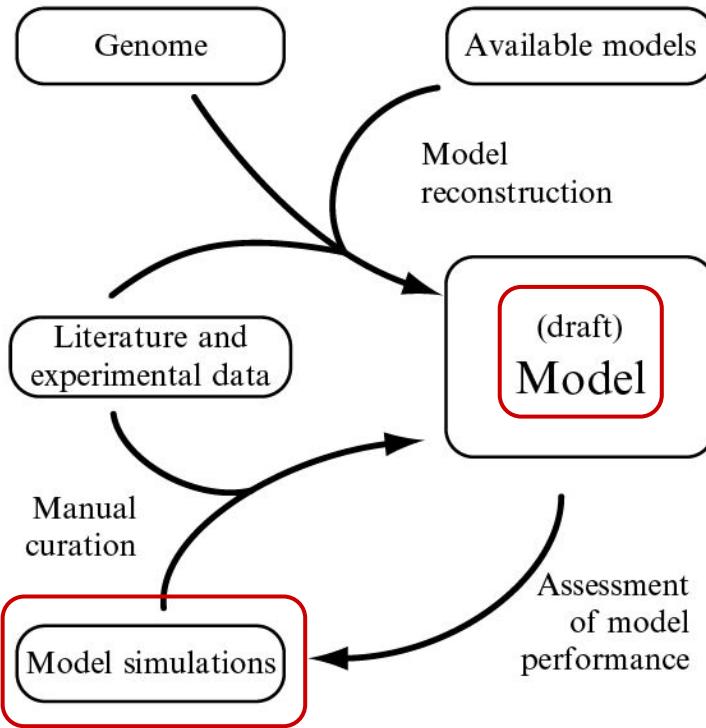
BD metabólica



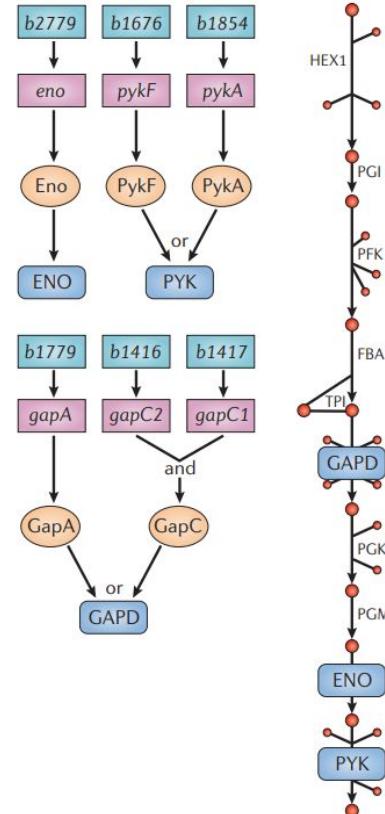
Software  
Algoritmo para  
predicción de  
rutas metabólicas

Regla GPR

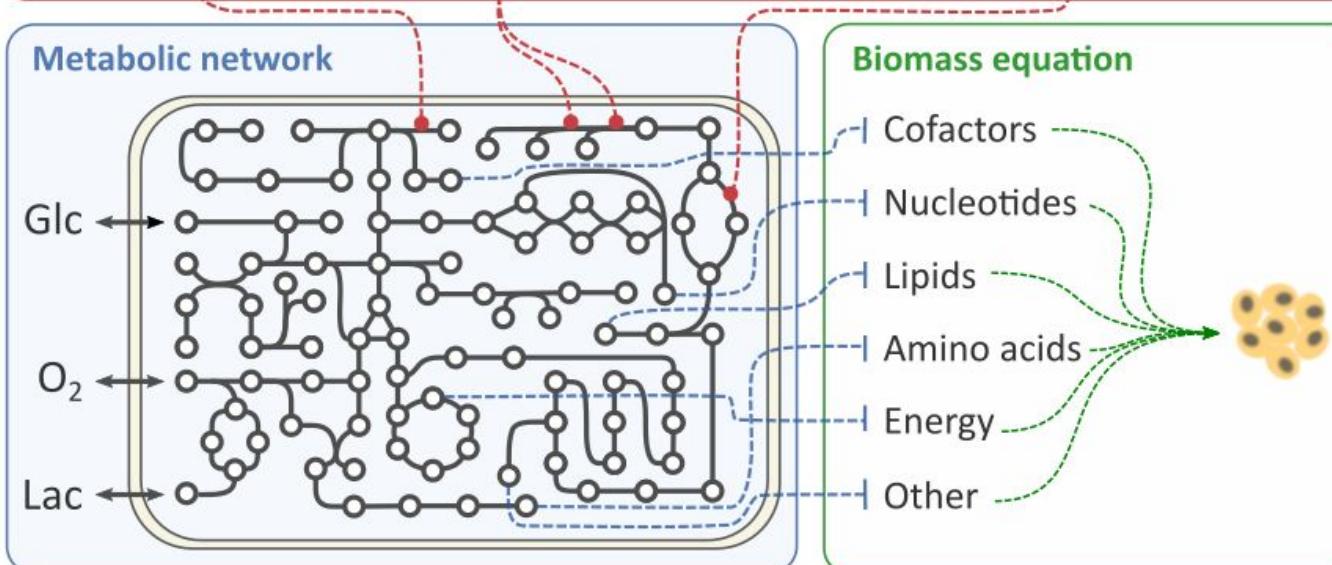
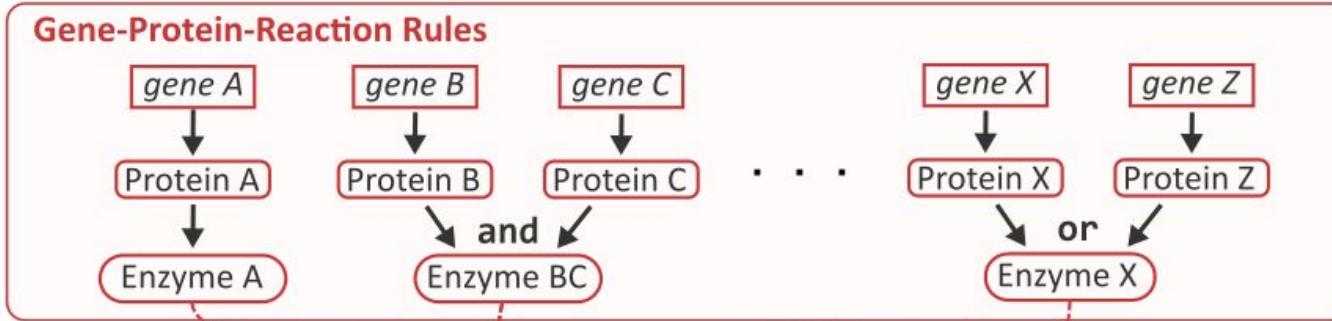
# Reconstruction of metabolic models



Abbreviation	Glycolytic reactions										Genes
HEX1	$[c]GLC + ATP \rightarrow G6P + ADP + H$										glk
PGI	$[c]G6P \leftrightarrow F6P$										pgi
PFK	$[c]ATP + F6P \rightarrow ADP + FDP + H$										pfkA, pfkB
FBA	$[c]FDP \leftrightarrow DHAP + G3P$										fbaA, fbaB
TPI	$[c]DHAP \leftrightarrow G3P$										tpiA
GAPD	$[c]G3P + NAD + PI \leftrightarrow 13DPG + H + NADH$										gapA, gapC1, gapC2
PGK	$[c]13DPG + ADP \leftrightarrow 3PG + ATP$										pgk
PGM	$[c]3PG \leftrightarrow 2PG$										gpmA, gpmB
ENO	$[c]2PG \leftrightarrow H_2O + PEP$										eno
PYK	$[c]ADP + H + PEP \rightarrow ATP + PYR$										pykA, pykF



# What is a Genome-scale metabolic model?



	Total
Genes	1675
Metabolites	5324
Reactions	7785

Biomass	Total
RNA components:	4
DNA components:	4
Protein components:	20
Carbohydrates:	1
Lipid components:	16
Cofactors and vitamins:	10

	rxn2
A	0
B	0
:	
E	-1
F	+1

# Genome-scale metabolic model in SBML Format

```
<model id="Recon2.2.1">
  <listOfUnitDefinitions>
    <unitDefinition id="mmol_per_gDW_per_hr"> ... </listOfUnitDefinitions>
  <listOfCompartments>
    <compartment id="g" name="Golgi apparatus" size="1"/>
    <compartment id="c" name="cytoplasm" size="1"/>
    ...
  </listOfCompartments>
  <listOfSpecies>
    <species id="M_10fthf5glu_c" name="10-formyltetrahydrofolate-[Glu](5)" compartment="c" charge="-6">
      <notes>FORMULA: C40H45N11O19</notes>
    </species>
    ...
  </listOfSpecies>
  <reaction id="R_ENO" name="enolase" reversible="true">
    <notes>
      <p>GENE ASSOCIATION: HGNC:3350 or HGNC:3354 or HGNC:3353</p>
      <p>CONFIDENCE LEVEL: 5</p>
      <p>SUBSYSTEM: Glycolysis/gluconeogenesis</p>
    </notes>
    <listOfReactants>
      <speciesReference species="M_2pg_c" stoichiometry="1"/>
    </listOfReactants>
    <listOfProducts>
      <speciesReference species="M_h2o_c" stoichiometry="1"/>
      <speciesReference species="M_pep_c" stoichiometry="1"/>
    </listOfProducts>
    <kineticLaw>
      <listOfParameters>
        <parameter id="UPPER_BOUND" value="1000" units="mmol_per_gDW_per_hr"/>
        <parameter id="FLUX_VALUE" value="0" units="mmol_per_gDW_per_hr"/>
        <parameter id="OBJECTIVE_COEFFICIENT" value="0" units="dimensionless"/>
        <parameter id="LOWER_BOUND" value="-1000" units="mmol_per_gDW_per_hr"/>
      </listOfParameters>
    </kineticLaw>
  </reaction>
</model>
```



<http://sbml.org/>



<https://metabolicatlas.org>



<https://www.vmh.life/>



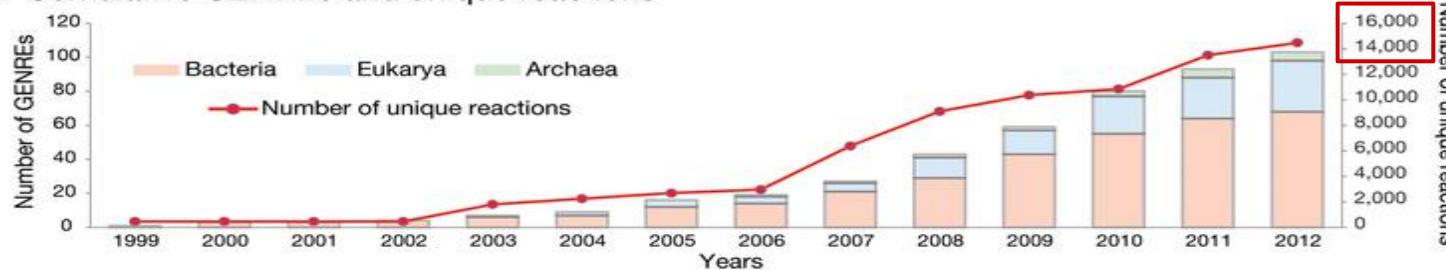
<http://bigg.ucsd.edu/>



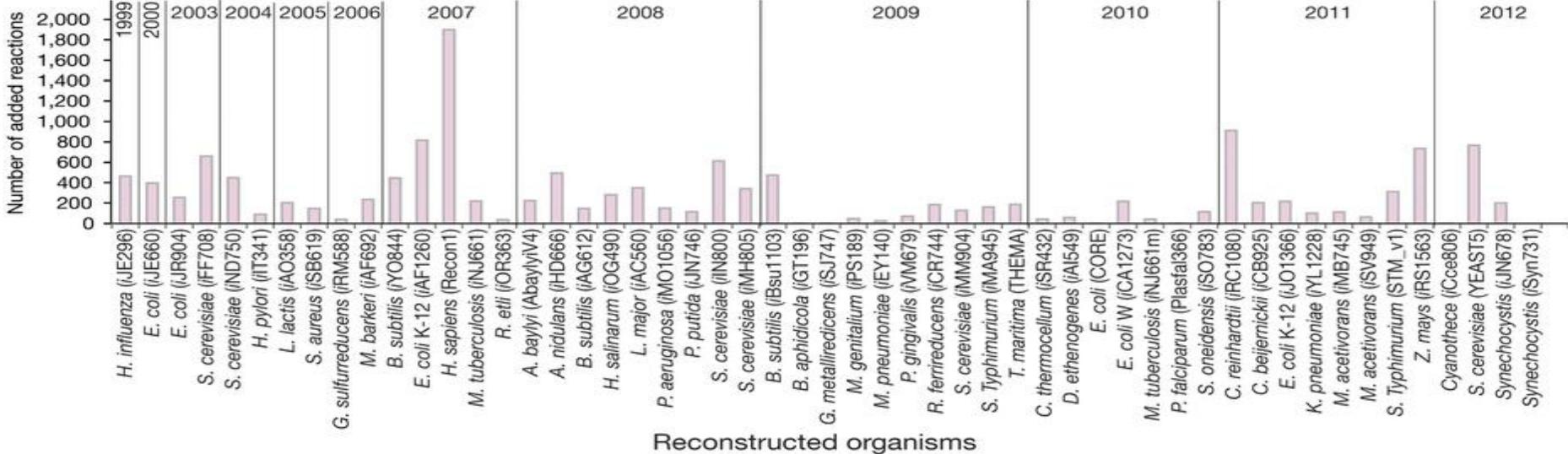
<https://www.ebi.ac.uk/biomodels-main/>

# Manually curated Genome-Scale Models

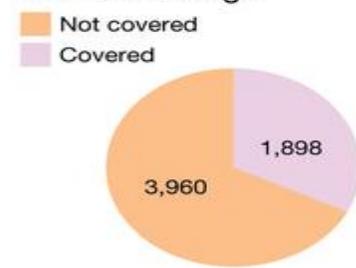
## a Cumulative GENREs and unique reactions



## C New reaction additions by model



## b EC coverage



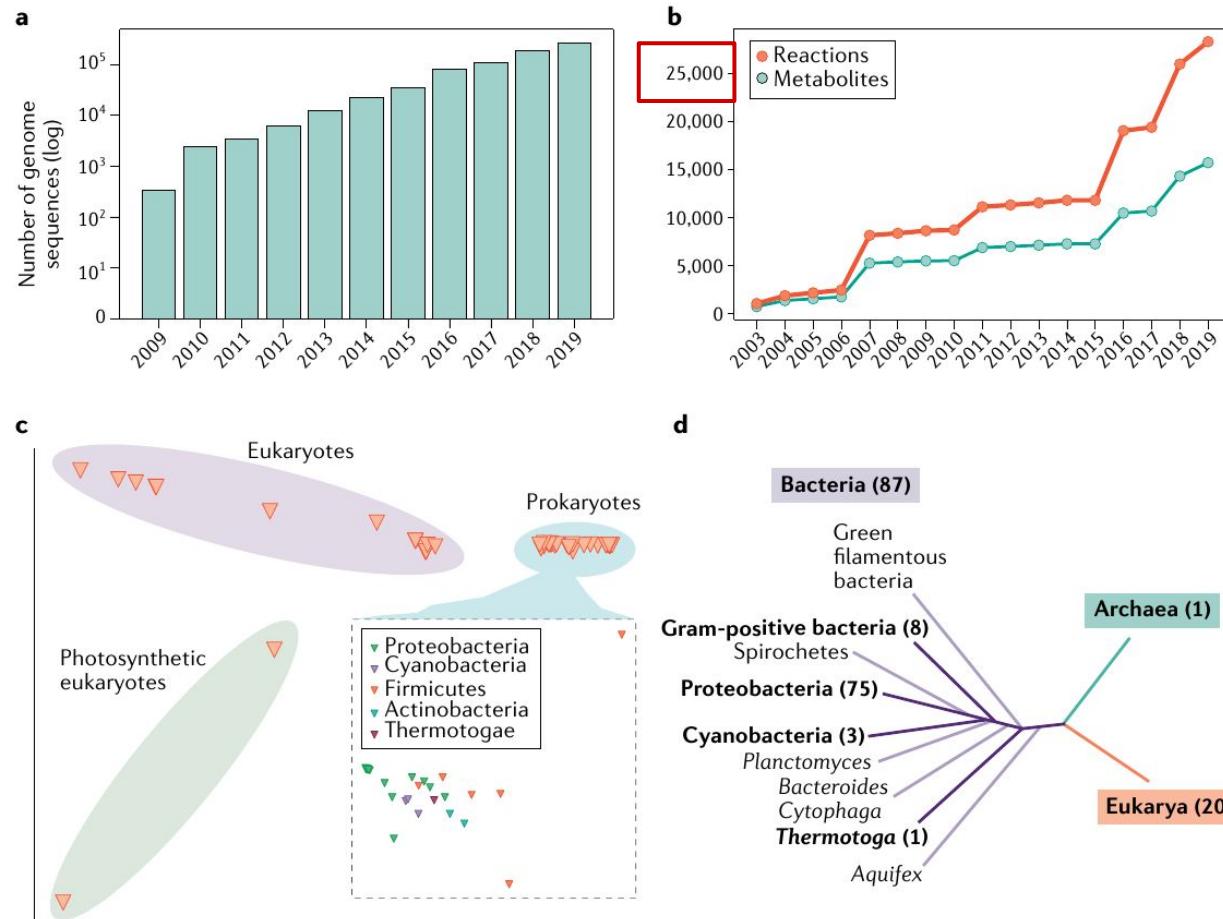
# Update no Genomes and reconstructed of genome-scale models

a) Number of public **genome sequences** in the PATRIC database.

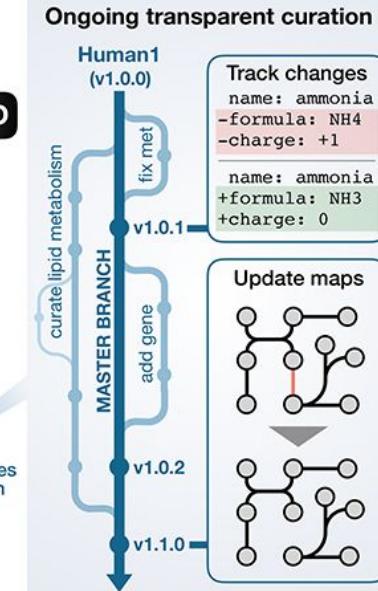
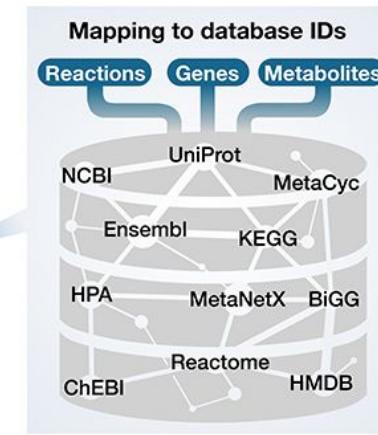
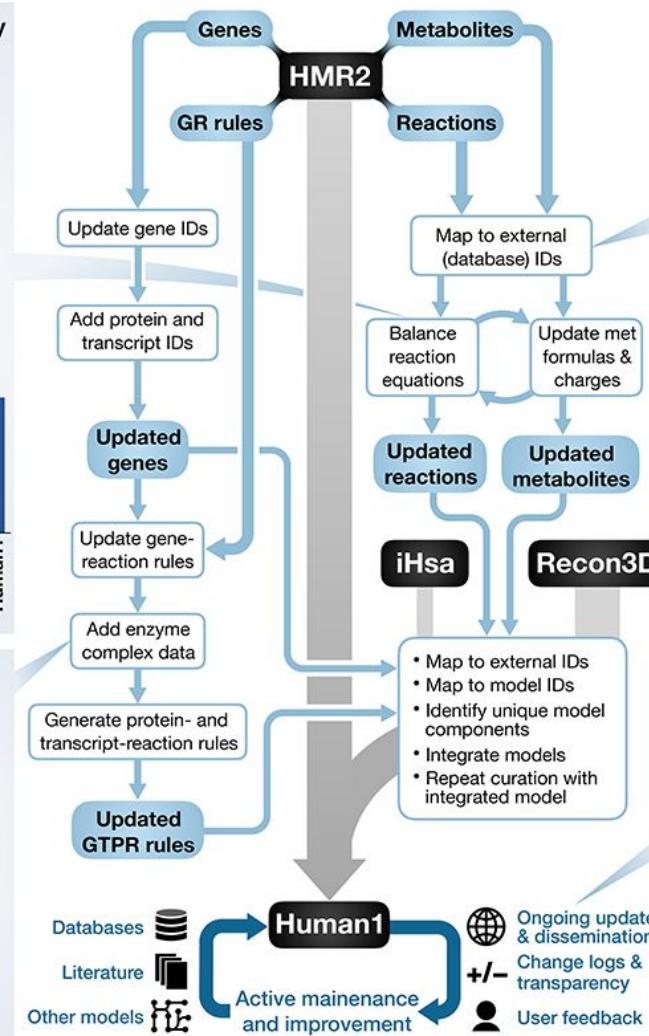
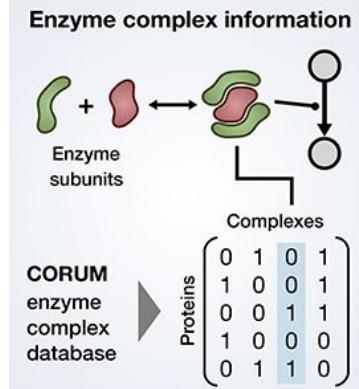
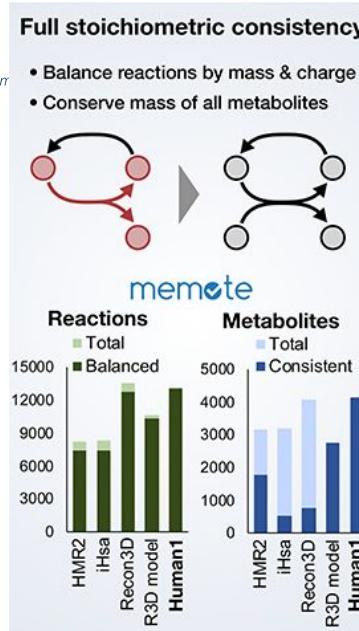
b) Number of **reactions and metabolites** represented in 108 manually curated models in the BiGG Models database.

c) Multiple correspondence analysis of the reactomes of the 108 reconstructions.

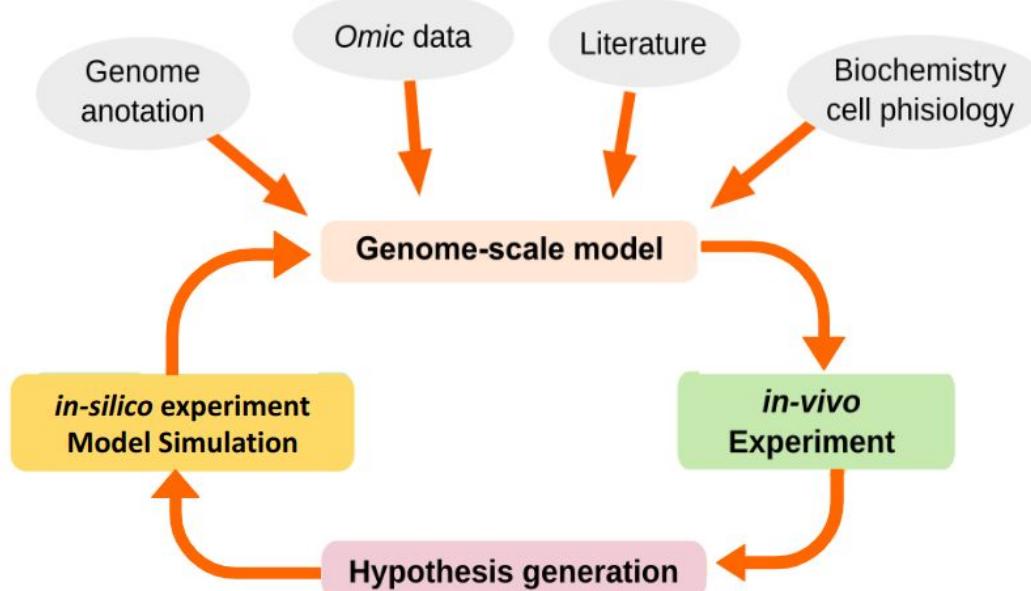
d) **Coverage** of the 108 reconstructions in the tree of life. The number in parentheses represents the number of reconstructions in each branch. Groups with at least one metabolic reconstruction are bolded



# Human Metabolism



# Reconstruction of Genome-scale Metabolic Models



## Summary

- Is the iterative process to elucidate the metabolic network of an organism.
- It can be as another dimension of genome annotation.
- Integrate information using bioinformatics to produce a PGDB.
- A PGDB are use to reconstruct genome-scale metabolic models.

## Manually curated Metabolic Models

- BiGG (Norsigian et al. 2020. doi:10.1093/nar/gkz1054)
- Browser: <http://bigg.ucsd.edu/>
- Viewer <https://escher.github.io/>
- BioModels

## Automatic Metabolic Reconstruction

- THE SEEDv2 Model (Henry et al 2023)
- CarVeMe (Machado et al. 2018)
- Pathway Tools (Paley and Karp)

## Human Metabolic Models

**Recon3D (Brunk et al. 2018. doi:10.1038/nbt.4072)**

- Browser: <https://www.vmh.life/#human/all>
- Viewer: <https://www.vmh.life/minerva/>

**Human-GEM (Robinson et al. 2021. doi:10.1126/scisignal.aaz1482 )**

- Browser: <https://metabolicatlas.org/explore/Human-GEM/gem-browser>
- Viewer: <https://metabolicatlas.org/explore/Human-GEM/map-viewer/>

# References: genome-scale metabolic reconstruction

- Stein, L. *Genome Annotation: From Sequence To Biology*. Nat. Rev. Gen. (2001).
- Ma, H. & Zeng, A. *Reconstruction of metabolic networks from genome data and analysis of their global structure for various organisms*. Bioinformatics (2003).
- Karp, P. D. *Pathway databases: a case study in computational symbolic theories*. Science (2001).
- Karp, P. D. et al. *Multidimensional annotation of the Escherichia coli K-12 genome*. Nucleic acids research (2007).
- Henry, C. S. et al. *High-throughput generation, optimization and analysis of genome-scale metabolic models*. Nat. Biotech. (2010).
- Monk, J., Nogales, J. & Palsson, B. O. *Optimizing genome-scale network reconstructions*. Nat. Biotech. (2014).
- Brunk, E. et al. *Recon3D enables a three-dimensional view of gene variation in human metabolism*. Nat. Biotech. (2018)
- Robinson, J. L. et al. *An atlas of human metabolism*. Sci. Signal. 13, (2020)

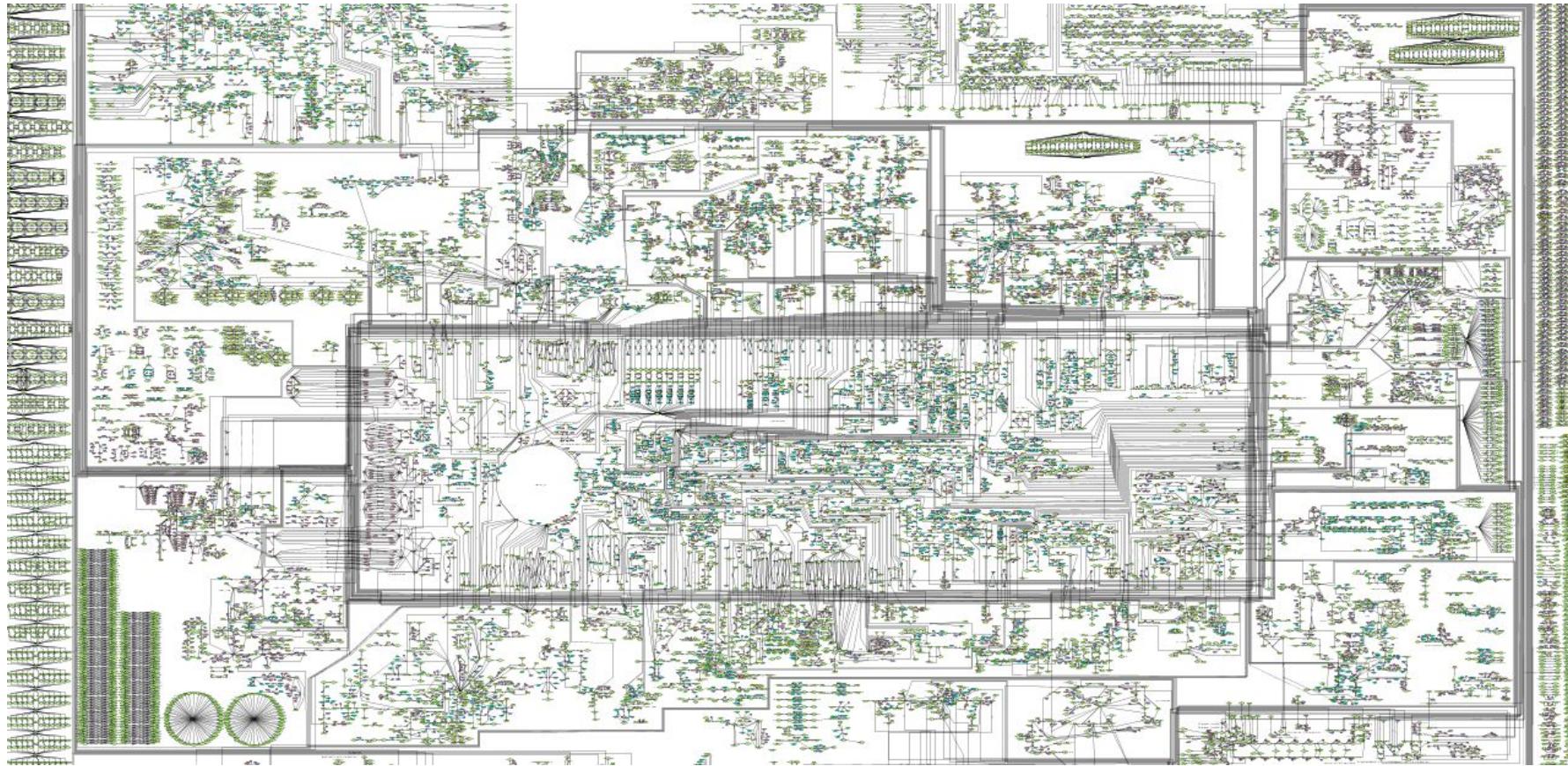


**Barcelona  
Supercomputing  
Center**

Centro Nacional de Supercomputación

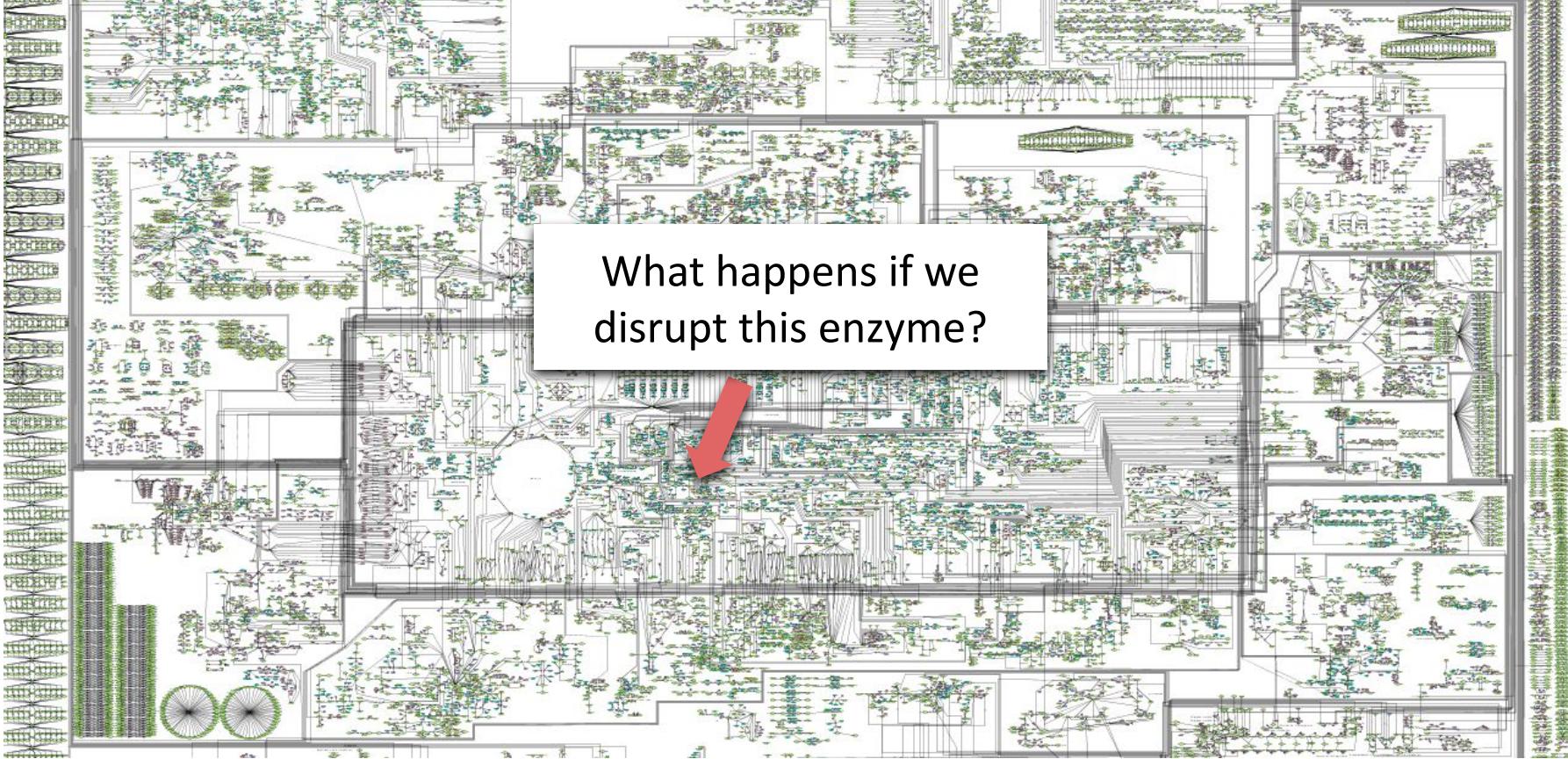
# Virtual Human Metabolic:

<https://www.vmh.life/minerva/>



# Virtual Human Metabolic:

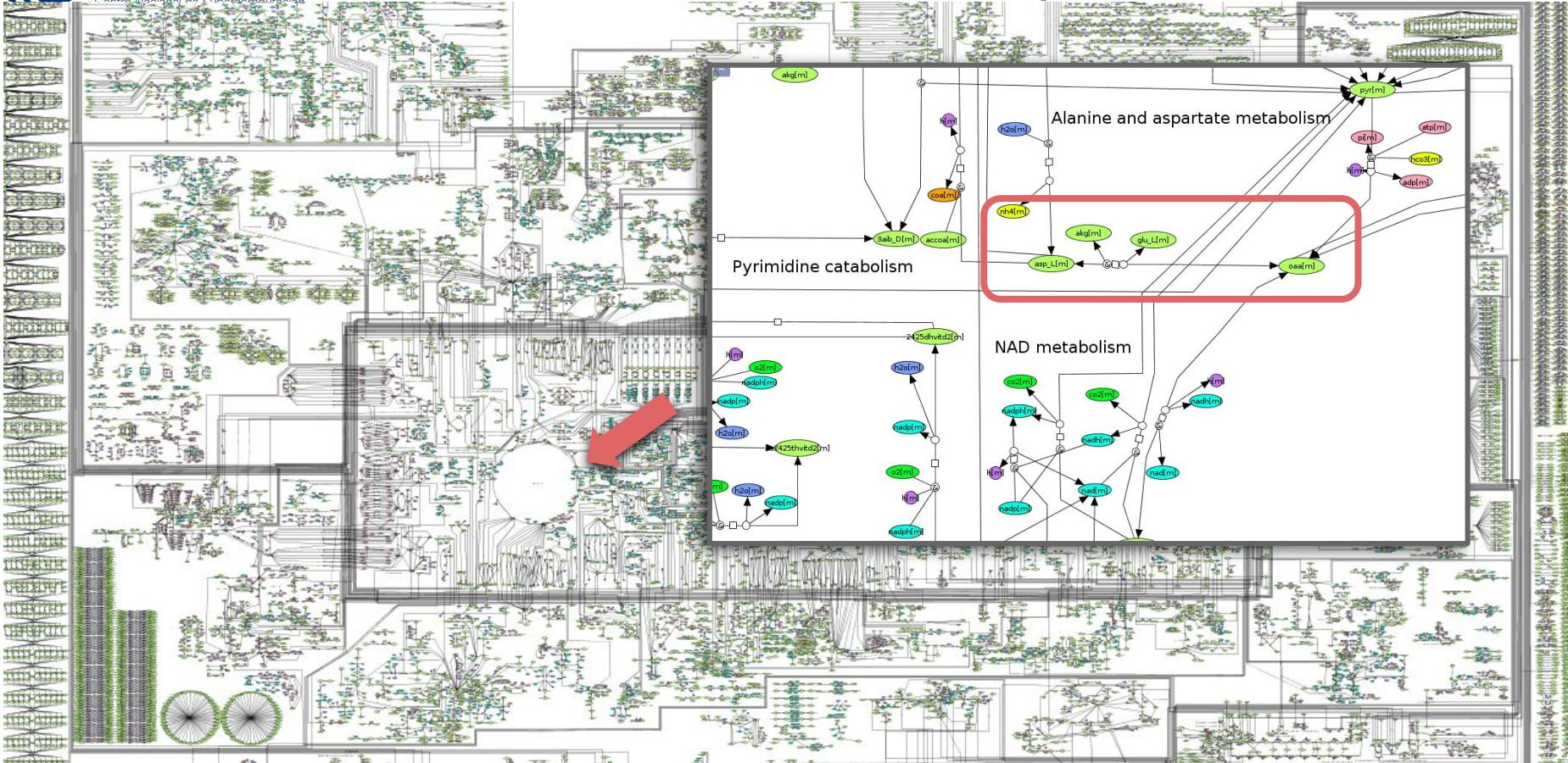
<https://www.vmh.life/minerva/>



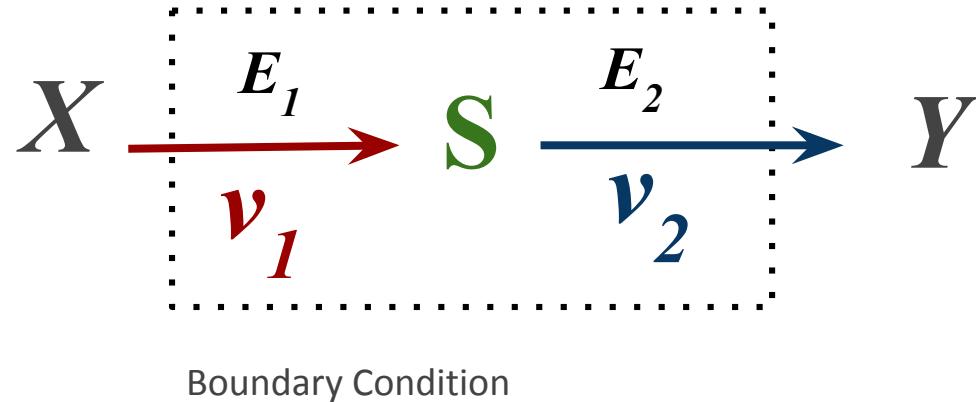
A complex metabolic pathway diagram showing numerous green nodes (enzymes) connected by a network of black lines (substrate and product interactions). A red arrow points to a specific node in the center of the diagram.

What happens if we  
disrupt this enzyme?

# Let's focus on the smallest components: reactions



# Kinetic Modeling of biochemical reactions



## Model components

Variables:  $[S]$ ,  $v_1$ ,  $v_2$

Parameters:  $X$ ,  $Y$ ,  $E_1$ ,  $E_2$

*Reaction rate equations*

## Reactions rate Equations

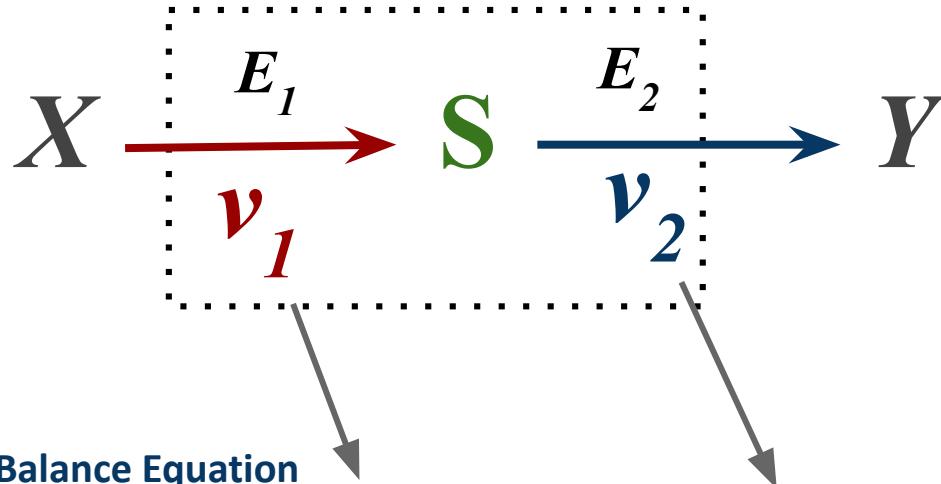
$$v_1(X, S, t, \Theta) \quad X \rightarrow S$$

$$v_2(S, Y, t, \Theta') \quad S \rightarrow Y$$

## Mass Balance Equation

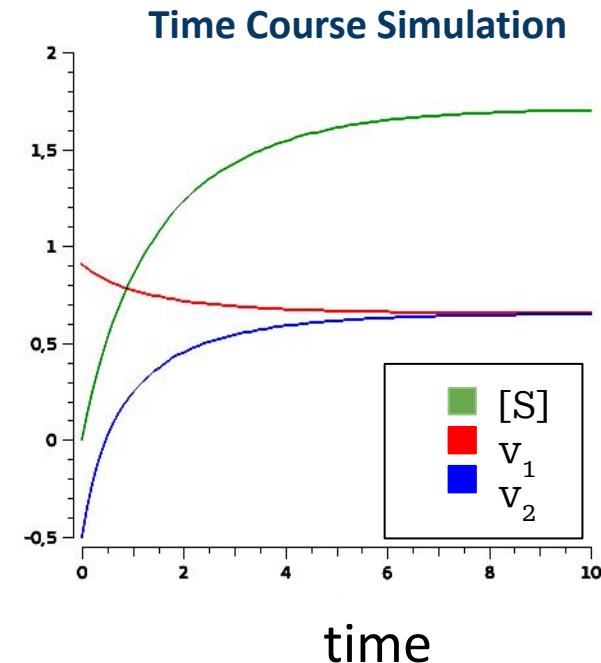
$$\frac{dS}{dt} = v_1 - v_2$$

# Kinetic Modeling of biochemical reactions

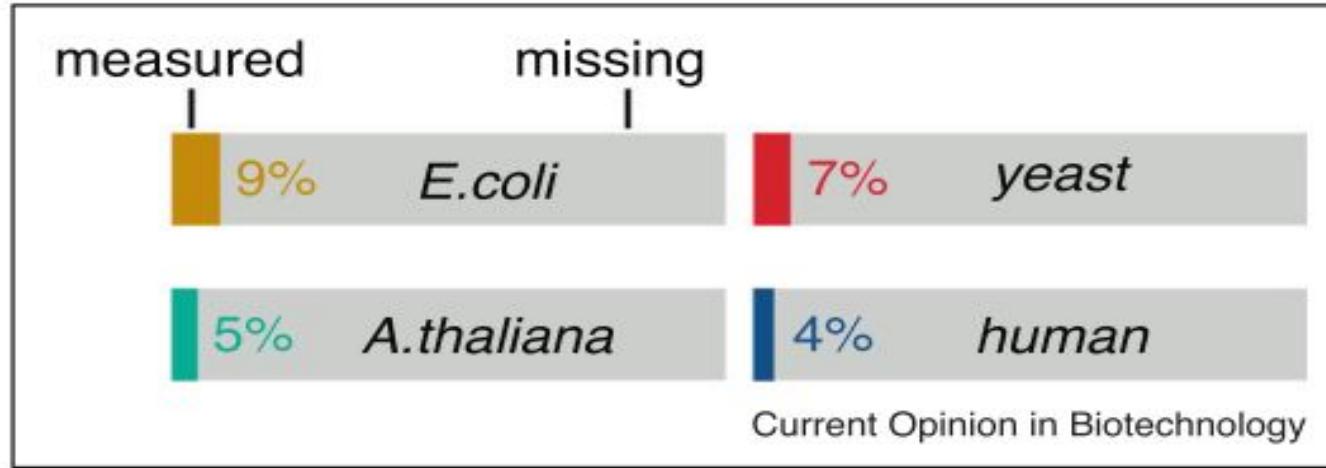


Mass Balance Equation

$$\frac{d([S])}{dt} = \left( \frac{\frac{V_f(v_1) \cdot [X]}{K_{ms}(v_1)} - \frac{V_r(v_1) \cdot [S]}{K_{mp}(v_1)}}{1 + \frac{[X]}{K_{ms}(v_1)} + \frac{[S]}{K_{mp}(v_1)}} \right) + \left( \frac{\frac{V_f(v_2) \cdot [S]}{K_{ms}(v_2)} - \frac{V_r(v_2) \cdot [Y]}{K_{mp}(v_2)}}{1 + \frac{[S]}{K_{ms}(v_2)} + \frac{[Y]}{K_{mp}(v_2)}} \right)$$



# Kinetic constants/parameters: the cursed of modeling

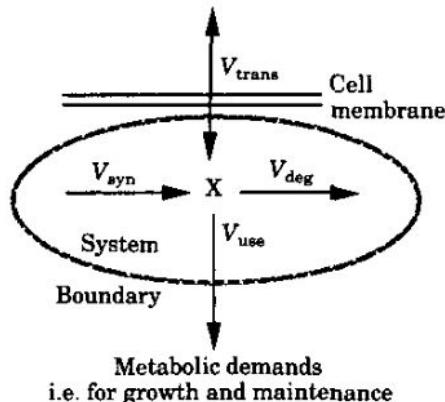
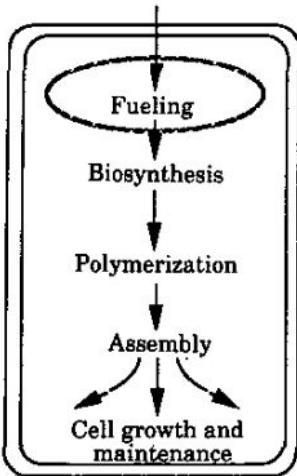


## Nº of reactions from GEMs:

- *E. coli* (*iJO1266*): 2,251
- Budding yeast (*iND750*): 1,149
- Arabidopsis (–): 1,363
- Human (Human1): 13,085

# The dawn of the constraint-based modeling

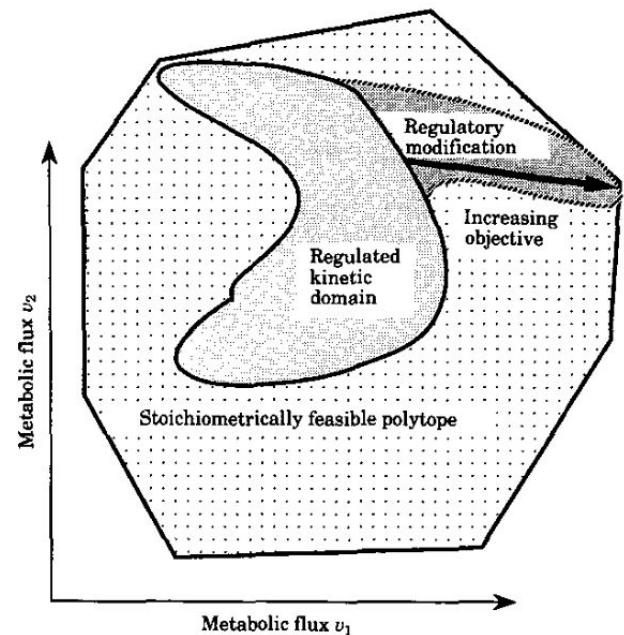
Carbon source and minerals



Living systems operate under a wide range of **constraints**

- Physical constraints (thermodynamics)
- Environmental constraints (availability of nutrients)
- Kinetic and Regulatory constraints

Constraints can be used to exclude impossible cellular states!

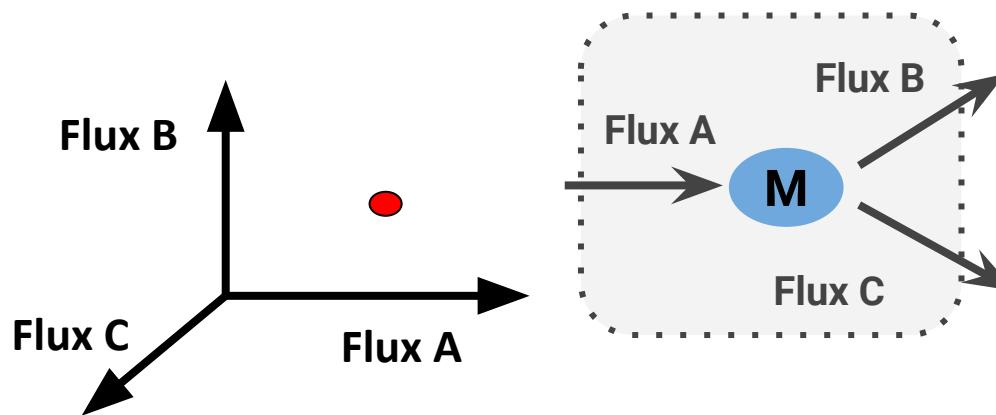


“... When you have eliminated all which is impossible then whatever remains, however improbable, must be the truth...”. Sherlock Holmes

# Two approaches for simulating metabolic systems

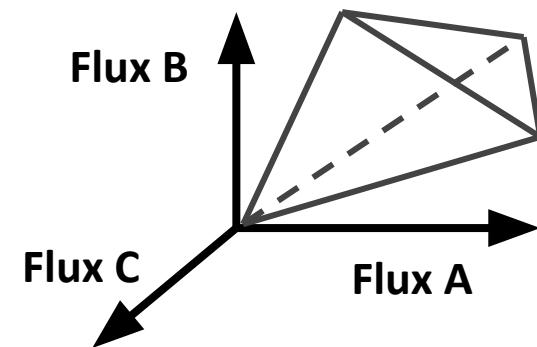
## Detailed Kinetic Model

- Information on enzyme kinetics
- Detailed knowledge about reaction mechanisms
- Solution is a single dot in the phase space  
(steady state), or a trajectory  
(oscillations, complex attractors)



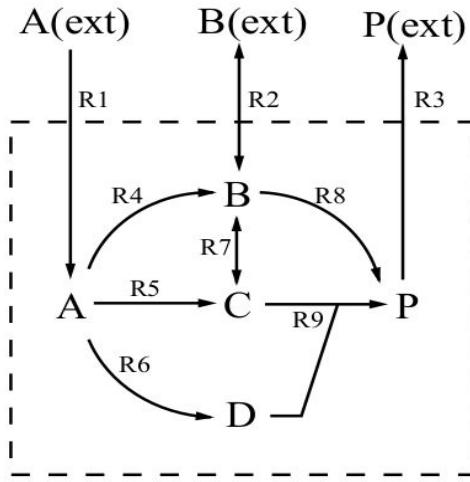
## Constraint-Based Model

- Stoichiometric Network
- Constraints → Flux space
- Explore the flux space
  - Enumeration
  - Optimization
  - Sampling



# Basic definitions: the stoichiometric matrix

The biochemical system



Mass balance equations  
(matrix notation)

$$\begin{aligned} d[\mathbf{A}] / dt \\ d[\mathbf{B}] / dt \\ d[\mathbf{C}] / dt \\ d[\mathbf{D}] / dt \\ d[\mathbf{P}] / dt \\ d[\mathbf{A}_{\text{ext}}] / dt \\ d[\mathbf{B}_{\text{ext}}] / dt \\ d[\mathbf{P}_{\text{ext}}] / dt \end{aligned}$$

=

$$\begin{array}{c|ccccccccc} & \mathbf{R1} & \mathbf{R2} & \mathbf{R3} & \mathbf{R4} & \mathbf{R5} & \mathbf{R6} & \mathbf{R7} & \mathbf{R8} & \mathbf{R9} \\ \mathbf{A} & 1 & 0 & 0 & -1 & -1 & -1 & 0 & 0 & 0 \\ \mathbf{B} & 0 & 1 & 0 & 1 & 0 & 0 & -1 & -1 & 0 \\ \mathbf{C} & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & -1 \\ \mathbf{D} & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & -1 \\ \mathbf{P} & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 1 & 1 \\ \mathbf{A}_{\text{ext}} & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \mathbf{B}_{\text{ext}} & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \mathbf{P}_{\text{ext}} & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \end{array} \cdot$$

Flux vector  $\mathbf{v}$

Mass balance equations

$$d[\mathbf{A}] / dt = \mathbf{R1} - \mathbf{R4} - \mathbf{R5} - \mathbf{R6}$$

...

$$d[\mathbf{D}] / dt = \mathbf{R6} - \mathbf{R9}$$

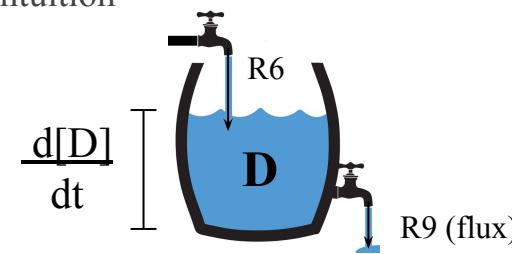
...

$\mathbf{N}$ : *m*x*n*

$\mathbf{m}$ : metabolites

$\mathbf{n}$ : reactions

Intuition



# Basic definitions: mass balance constraints

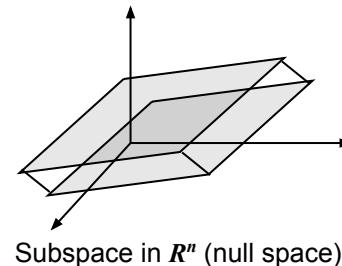
## Mass Balance

- First Law of Thermodynamics
- Conservation of mass
- Quasi steady-state condition
- Hypothesis: time scale of metabolism is faster than regulation

### Constraint equations

$$N \cdot v = 0$$

Linear system of equations



d[A] / dt	Stoichiometric Matrix									Flux vector $v$
	R1	R2	R3	R4	R5	R6	R7	R8	R9	
d[B] / dt	1	0	0	-1	-1	-1	0	0	0	2
d[C] / dt	0	1	0	1	0	0	-1	-1	0	0
d[D] / dt	0	0	0	0	1	0	1	0	-1	1
d[P] / dt	0	0	-1	0	0	0	0	1	1	0
d[A <sub>ext</sub> ] / dt	-1	0	0	0	0	0	0	0	0	1
d[B <sub>ext</sub> ] / dt	0	-1	0	0	0	0	0	0	0	1
d[P <sub>ext</sub> ] / dt	0	0	1	0	0	0	0	0	0	0

=  $\cdot = 0$

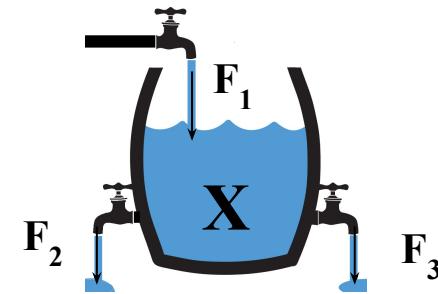
### Network structure

$N: mxn$

$m: \text{metabolites (rows)}$

$n: \text{reactions (columns)}$

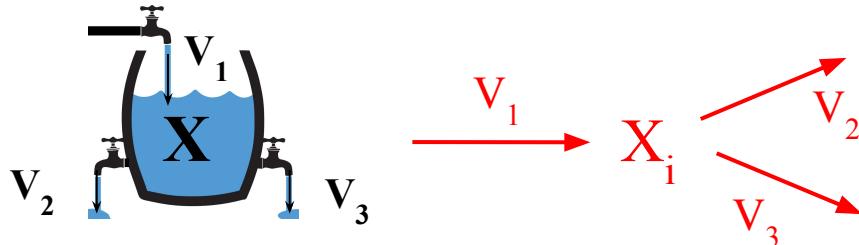
$n > m$



$$\frac{dX}{dt} = F_1 - F_2 - F_3 = 0$$

# Hipótesis del estado estacionario

Para cada metabolito de la red se escribe la ecuación de balance:



El balance de flujos en el componente  $X_i$ :

$$V_1 = V_2 + V_3 \quad \boxed{V_1 - V_2 - V_3 = 0}$$

- El resultado es un sistema con **m** ecuaciones (metabolitos) and **n** incognitas (flujos)

Matrix notation: **N . v = 0**

**N** = Matriz estequiométrica (**m** x **n**)

**v** = Vector de flujos metabólicos (**n**)

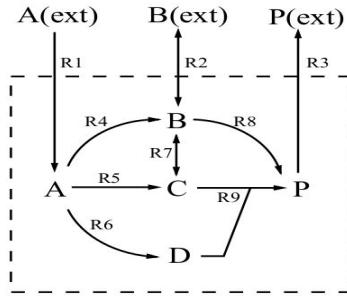
En general las redes metabólicas tienen más reacciones que metabolitos:

$$n > m$$

de manera que el sistema es indeterminado

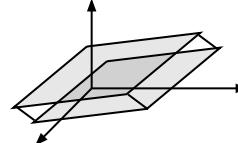
No hay una única solución!!!

# Thermodynamics constraints → the flux cone



## Mass Balance

$$N \cdot v = 0 \quad \rightarrow$$



Subspace in  $R^n$

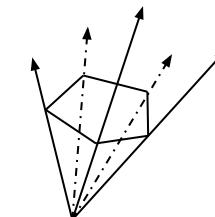
## Thermodynamics

Irreversible reactions must have non-negative flux

$$v_i \geq 0 \quad i \text{ Irrev}$$

## Irreversible reactions

$\{R1, R3, R4, R5, R6, R8, R9\}$



Convex cone

## Mass Balance

$$\begin{array}{|c|c|c|c|c|c|c|c|c|} \hline & 1 & 0 & 0 & -1 & -1 & -1 & 0 & 0 & 0 \\ \hline & 0 & 1 & 0 & 1 & 0 & 0 & -1 & -1 & 0 \\ \hline & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & -1 \\ \hline & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & -1 \\ \hline & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 1 & 1 \\ \hline & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \hline & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \hline & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ \hline \end{array} \cdot \begin{array}{|c|c|c|c|c|c|c|c|c|} \hline & 2 & 0 & 1 & 0 & 1 & 1 & 0 & 1 \\ \hline & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \hline & 1 & 0 & 1 & 0 & 1 & 1 & 0 & 0 \\ \hline & 0 & 0 & 1 & 0 & 1 & 1 & 0 & 0 \\ \hline & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ \hline & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ \hline & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \hline & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ \hline \end{array} = 0$$

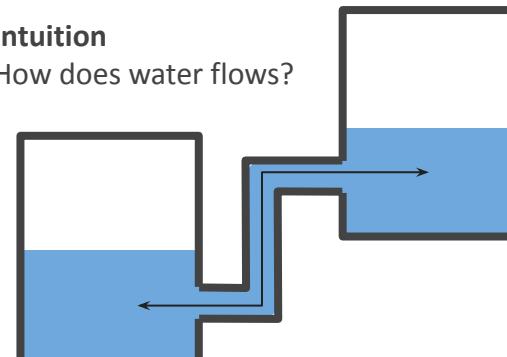
## Thermodynamics

R1	2
R3	1
R4	0
R5	1
R6	1
R8	0
R9	1

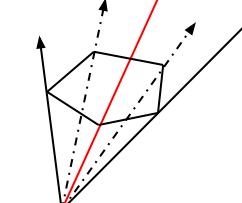
$$\geq 0$$

## Intuition

How does water flows?

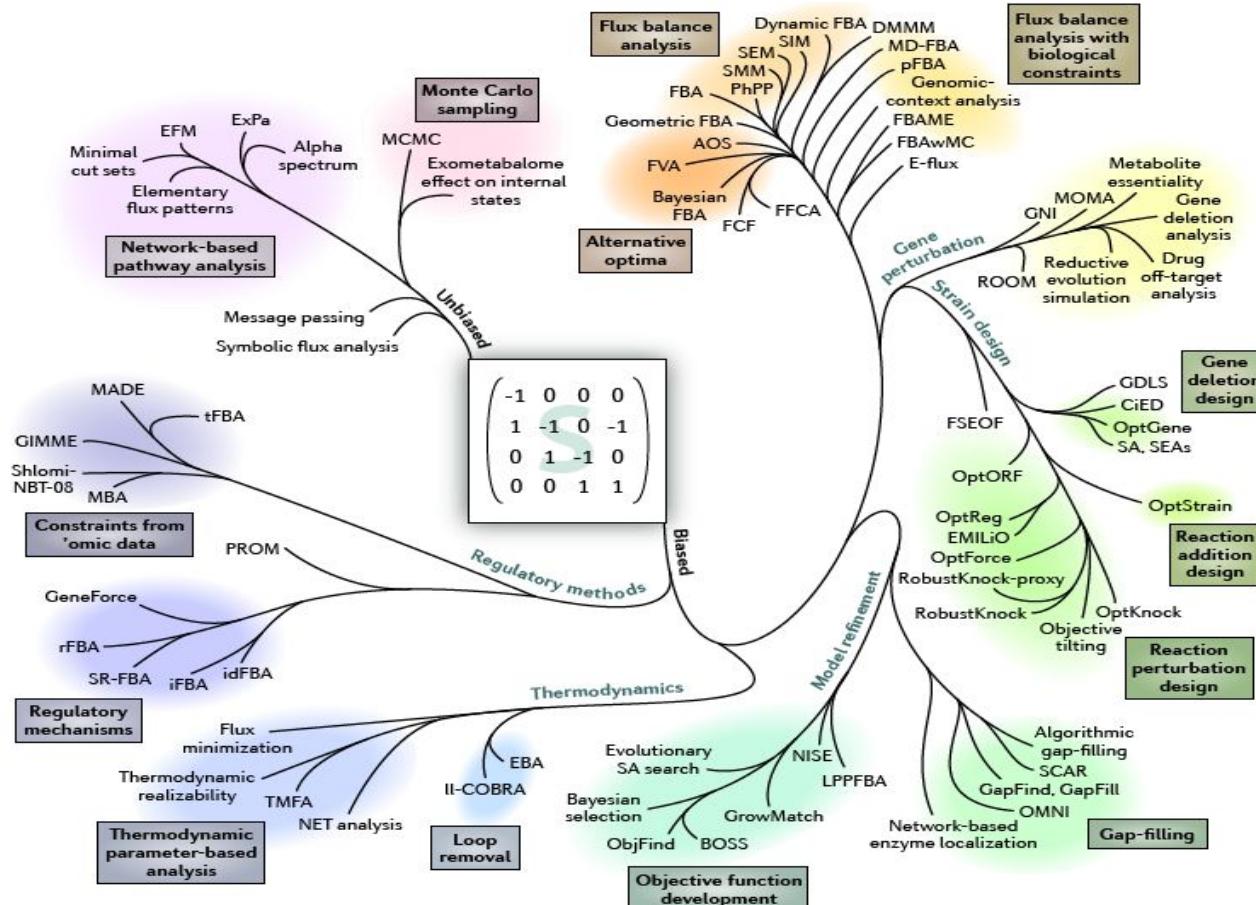


## Flux Cone



Flux distribution  $v$

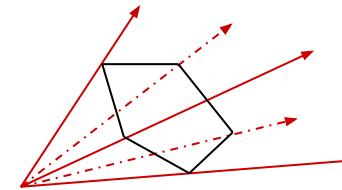
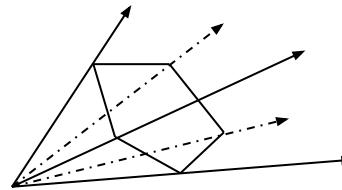
# The family of Constraint-Based Methods



# Constraint-based modelling

## Pathways Analysis (PA):

Caracterizar el espacio de flujos determinando todos los modos de funcionamiento mínimos (o modos elementales) → Vías metabólicas

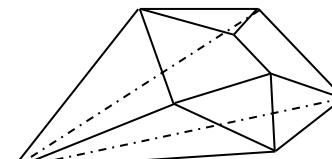


## Flux Balance Analysis (FBA)

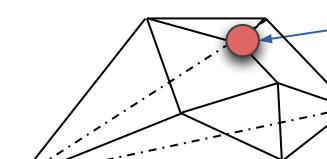
Identificar una única solución (vector de flujos) que optimice una función objetivo (por lo general la tasa de crecimiento).

## Metabolic Flux Analysis (MFA)

Emplea medidas de flujos experimentales para intentar “determinar” el valor de las velocidades no conocidas.



Espacio factible



Espacio factible

Solución particular

# Modos de Flujo Elemental

## ¿Por qué emplearlos?

1. Descomposición de la red en las unidades funcionales o módulos (vías metabólica)
2. No requiere conocimientos de parámetros cinética !!
3. Utiliza coeficientes estequiométricos y reversibilidad/irreversibilidad de las reacciones

Concepto relacionado: vía Extrema (Extreme Pathway)

(C.H. Schilling, D. Letscher y B.O. Palsson, *J. theor. Biol.* 203 (2000) 229)

- La diferencia está en que se distinguen dos tipos de reacciones: reacciones internas y de intercambio, y que todas las reacciones reversibles internas se dividen en un paso directo y otro reverso (irreversibles).

# Modos de Flujo Elemental

## Definición:

Conjunto **no-descomponible** (minimal) de reacciones capaz de operar en estado estacionario, obedeciendo restricciones de irreversibilidad

## Propiedades:

1. Conjunto **único**
2. El conjunto EFM<sub>s</sub> de una subred es un subconjunto de EFM<sub>s</sub> de la red original.
3. Cualquier modo  $v$  puede descomponerse en:

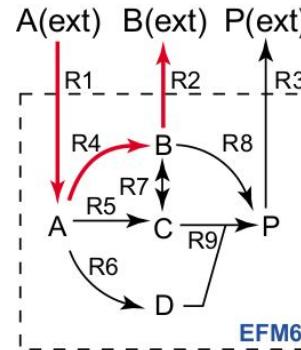
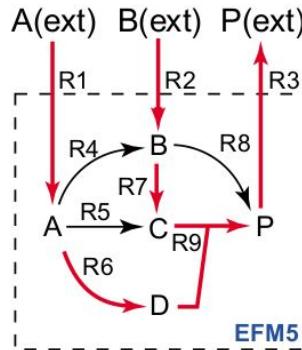
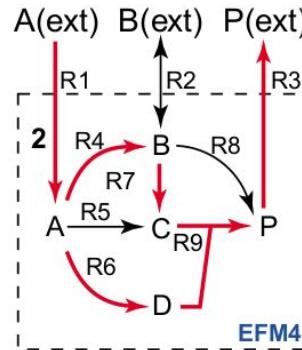
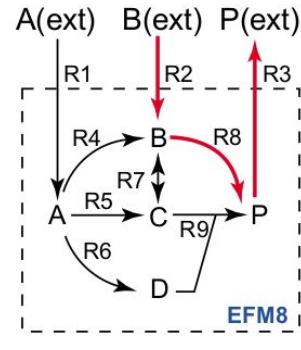
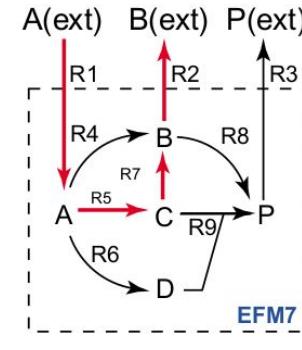
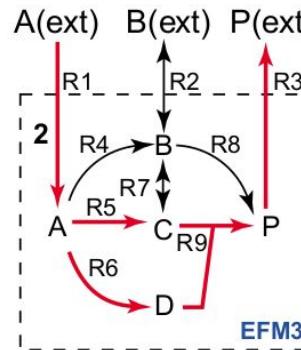
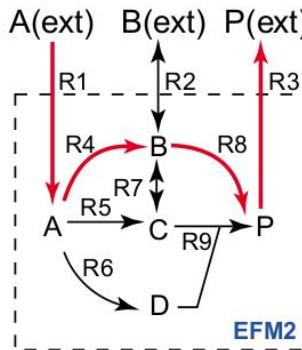
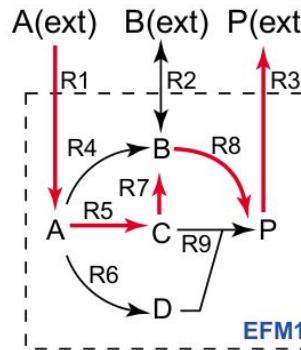
$$\vec{v} = \sum_i \lambda_i \vec{e}_i \quad \lambda_i \geq 0$$

El conjunto EFM<sub>s</sub> nos permite describir el espacio de flujos:

$$F = \left\{ v \mid N \cdot v = 0, \quad v_j \geq 0 \quad \forall j \in J_{Irrev} \right\}$$

en términos unidades funcionales más simples!!

# Elementary Flux Modes



## Reactions

R1	R2	R3	R4	R5	R6	R7	R8	R9
1	0	1	0	1	0	-1	1	0
1	0	1	1	0	0	0	1	0
2	0	1	0	1	1	0	0	1
2	0	1	1	0	1	1	0	1
1	1	1	0	0	1	1	0	1
1	-1	0	1	0	0	0	0	0
1	-1	0	0	1	0	-1	0	0
0	1	1	0	0	0	0	1	0



**Barcelona  
Supercomputing  
Center**

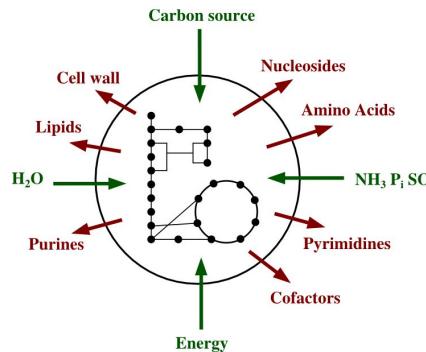
Centro Nacional de Supercomputación

# COPASI Hands-on!

# Environmental constraints → growth conditions



- RPMI
- Leibovitz-LT15

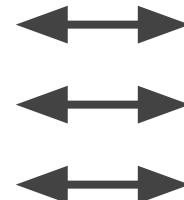
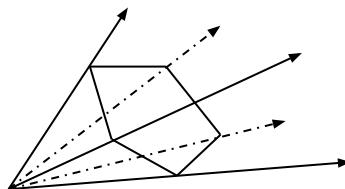


**Exchange fluxes:** artificial variables that represent the organism's exchanges with the environment:

- Nutrient incorporation (net consumption)
- Product secretion (net production)

	$\text{GLC}_{\text{Ex}}$	$\text{O}_{\text{Ex}}$	$\text{CO2}_{\text{Ex}}$
Glc	-1		
O <sub>2</sub>		-1	
CO <sub>2</sub>			-1

**Flux Cone**



Glucose  
Oxygen  
 $\text{CO}_2$

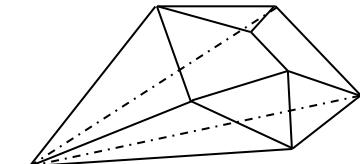
**Min**

- 6  
- 20  
0

**Max**

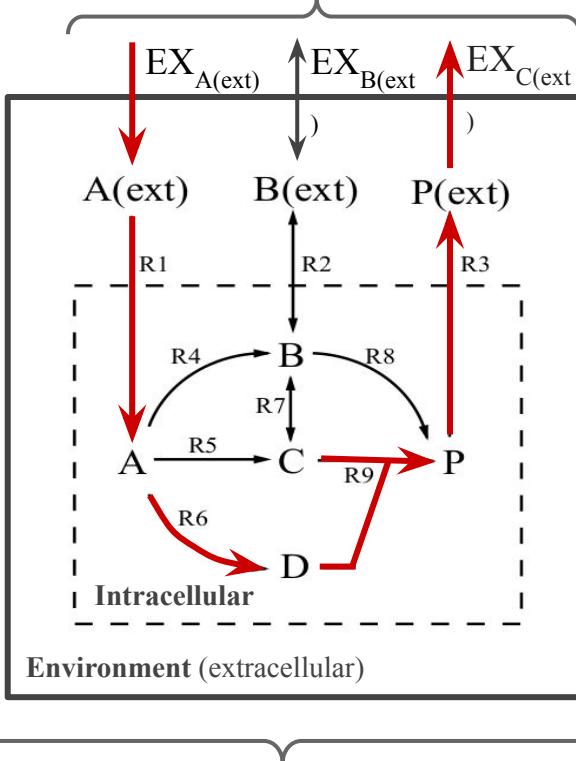
0  
0  
Inf

**Flux Space**



# Altogether → the Flux Space

Artificial variables (exchanges fluxes)



The model is expanded to consider environment

	R1	R2	R3	R4	R5	R6	R7	R8	R9	EX <sub>A</sub>	EX <sub>B</sub>	EX <sub>C</sub>
A	1	0	0	-1	-1	-1	0	0	0	0	0	0
B	0	1	0	1	0	0	-1	-1	0	0	0	0
C	0	0	0	0	1	0	1	0	-1	0	0	0
D	0	0	0	0	0	1	0	0	-1	0	0	0
P	0	0	-1	0	0	0	0	1	1	0	0	0
A <sub>ex</sub>	-1	0	0	0	0	0	0	0	0	-1	0	0
B <sub>ex</sub>	0	-1	0	0	0	0	0	0	0	0	-1	0
P <sub>ex</sub>	0	0	1	0	0	0	0	0	0	0	0	-1

Internal fluxes                                  Exchanges fluxes

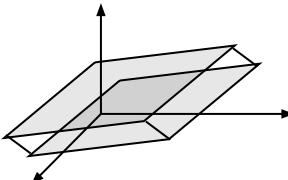
$$\begin{array}{c}
 \left[ \begin{array}{c|cc|cc}
 0 & 2 & & & 10 \\
 -10 & 0 & 10 & & 10 \\
 0 & 0 & 0 & 10 & 10 \\
 0 & 1 & 0 & 10 & 10 \\
 0 & 1 & 1 & 0 & 10 \\
 0 & 0 & 0 & 10 & 10 \\
 -10 & 0 & 0 & 10 & 10 \\
 0 & 0 & 0 & 10 & 10 \\
 0 & 0 & 10 & 0 & 10 \\
 0 & -10 & 0 & 0 & 10 \\
 0 & 0 & 0 & 10 & 10 \\
 0 & 0 & 0 & 10 & 10 \\
 \end{array} \right] \leq \\
 \left[ \begin{array}{c|cc|cc}
 2 & & & & 10 \\
 0 & 0 & & & 10 \\
 1 & & 10 & & 10 \\
 0 & 0 & 0 & 10 & 10 \\
 1 & 1 & 0 & 10 & 10 \\
 1 & 1 & 10 & 0 & 10 \\
 0 & 0 & 0 & 10 & 10 \\
 0 & 0 & 0 & 10 & 10 \\
 1 & 10 & 0 & 10 & 10 \\
 0 & 0 & 0 & 10 & 10 \\
 0 & 0 & 0 & 10 & 10 \\
 0 & 0 & 0 & 10 & 10 \\
 \end{array} \right]
 \end{array}$$

Exchanges fluxes' Lb and Ub define the growth media condition

# Altogether → the Flux Space

## C1. Mass Balance

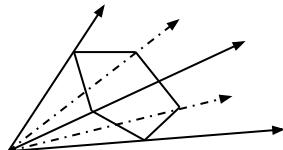
$$N \cdot v = 0$$



$C1 \leftarrow$  Subspace in  $R^n$  (null space)

## C2. Thermodynamics

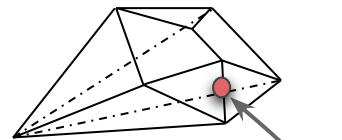
$$v_i \geq 0 \quad i \text{ Irrev}$$



$C1 \& C2 \leftarrow$  Flux cone (convex cone)

## C3. Capacity Constraints

$$Lb_i \leq v_i \leq Ub_i$$



$C1 \& C2 \& C3 \leftarrow$  Flux space  
(convex polyhedra)

Flux vector

$v_i$ : flux through reaction  $i$

$Lb_i$ : flux lower bound

$Ub_i$ : flux upper bound

## Mass Balance Constraint

	R1	R2	R3	R4	R5	R6	R7	R8	R9	R1	2
A	1	0	0	-1	-1	-1	0	0	0	R2	0
B	0	1	0	1	0	0	-1	-1	0	R3	1
C	0	0	0	0	1	0	1	0	-1	R4	0
D	0	0	0	0	0	1	0	0	-1	R5	1
P	0	0	-1	0	0	0	0	1	1	R6	1
$A_{\text{ex}}$	-1	0	0	0	0	0	0	0	0	R7	0
$B_{\text{ex}}$	0	-1	0	0	0	0	0	0	0	R8	0
$P_{\text{ex}}$	0	0	1	0	0	0	0	0	0	R9	1

$$= 0$$

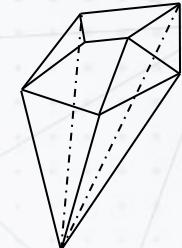
## Thermodyn.

## Capacity constraints

R1	2	0	10
R3	1	0	10
R4	0	0	10
R5	1	0	10
R6	1	1	10
R8	0	0	10
R9	1	0	10

$\geq 0$        $\leq$        $\leq$

## Flux Space



# Constraint-based modeling

C1. Mass Balance

$$N \cdot v = 0$$

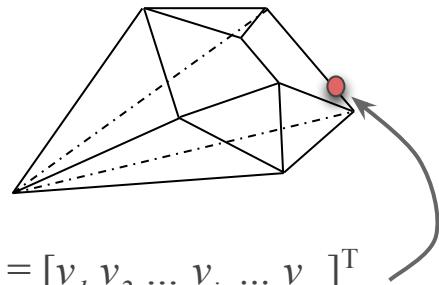
C2. Thermodynamics

$$v_i \geq 0 \quad i \text{ Irrev}$$

C3. Flux bounds

$$Lb_i \leq v_i \leq Ub_i$$

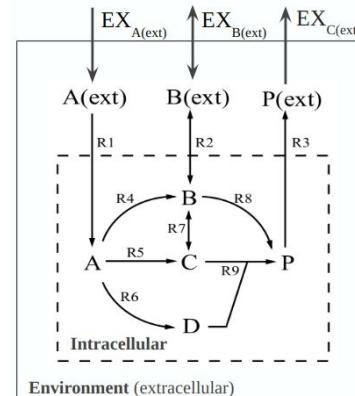
C1 & C2 & C3 ← Flux space



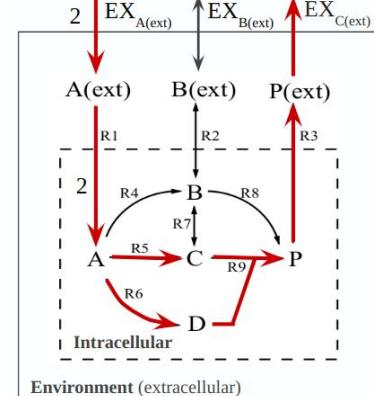
Predicting meaningful metabolic states

- Define biological goals ← optimization
- Apply further constraints (omics, regulation)
- Sampling methods
- Other approaches

“Genotype”



“Phenotype”



	R1	R2	R3	R4	R5	R6	R7	R8	R9	EX <sub>A</sub>	EX <sub>B</sub>	EX <sub>C</sub>	
A	1	0	0	-1	-1	-1	0	0	0	0	0	0	1
B	0	1	0	1	0	0	-1	-1	0	0	0	0	0
C	0	0	0	0	1	0	1	0	-1	0	0	0	1
D	0	0	0	0	0	1	0	0	-1	0	0	0	0
P	0	0	-1	0	0	0	0	1	1	0	0	0	0
A <sub>ex</sub>	-1	0	0	0	0	0	0	0	0	-1	0	0	1
B <sub>ex</sub>	0	-1	0	0	0	0	0	0	0	0	-1	0	-2
P <sub>ex</sub>	0	0	1	0	0	0	0	0	0	0	0	-1	1

$$= 0$$

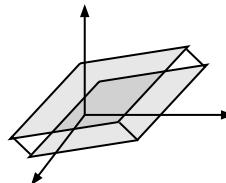
Stoichiometric Matrix  $N$

Flux vector  $v$

# Constraint-based modeling summary

## Mass Balance

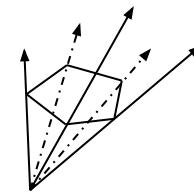
$$N \cdot v = 0$$



Subspace in  $R^n$  (null space)

## Thermodynamics

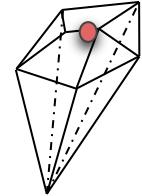
$$v_i \geq 0 \quad i \text{ Irrev}$$



Polyhedral cone (Flux cone)

## Flux bounds

$$Lb_i \leq v_i \leq Ub_i$$



Polytope (Flux space)

## Genome-Scale Metabolic Modeling

- **Kinetic modeling** is currently not possible at genome-scale because unknown kinetic parameter and reactions mechanisms.
- **Constraint-Based Modeling (CBM)** really uses the use of (linear) constraints to define a feasible space (flux space) of solutions (phenotype) → Family o Methods
- **CBM** only requires the stoichiometric matrix (reconstructed from a **PGDB**) and the set of irreversible reactions and thus it can be applied at genome-scale.
- **Linear Programming** can be used to explore the flux space (**Flux Balance Analysis**)

# Using optimization criteria to find metabolic states

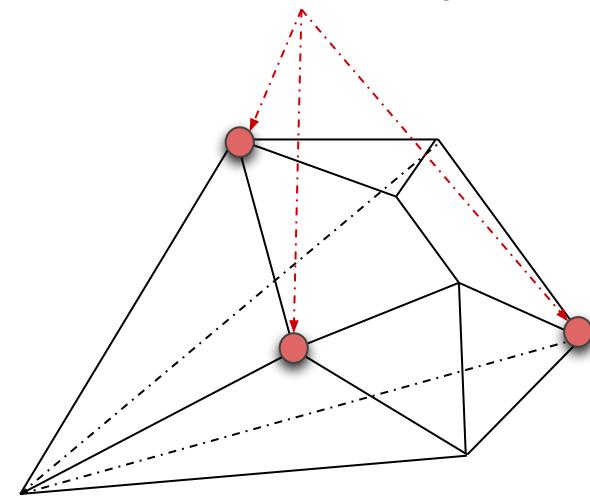
Candidate phenotypes

## Biological meaningful targets

- . Maximize biomass production (growth rate)
- . Maximize ATP production (energy)
- . Minimize total flux (resources efficiency)
- . Multiple criteria (Pareto efficiency)

## Fitting with experimental data

- . Metabolic Flux Balancing (measure fluxes)
- . Exo-metabolomic (constrain exchanges)
- . Gene expression data (context-specific)



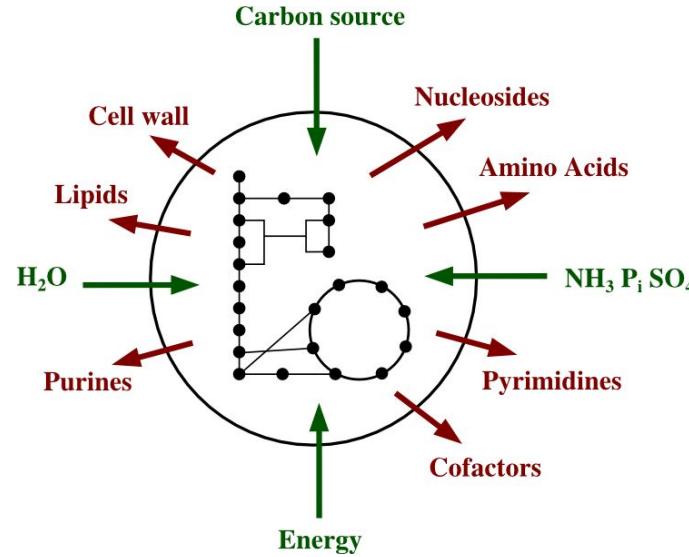
$N + \text{Constraints} \rightarrow \text{Flux Space}$

$$N \cdot v = 0$$

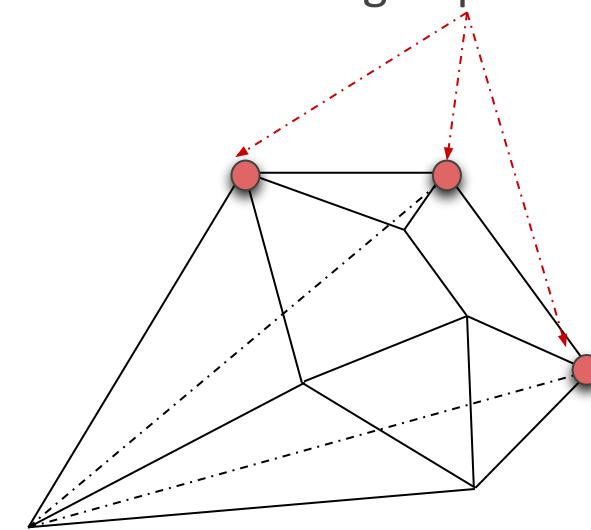
$$v_i \geq 0 \quad i \text{ Irrev}$$

$$Lb_i \leq v_i \leq Ub_i$$

# Flux Balance Analysis



Meaningful phenotypes?



## Hypothesis (idealization)

A growing cell will use its metabolism to maximize its growth rate

Maximize growth rate → maximize biomass production

What is the biomass?

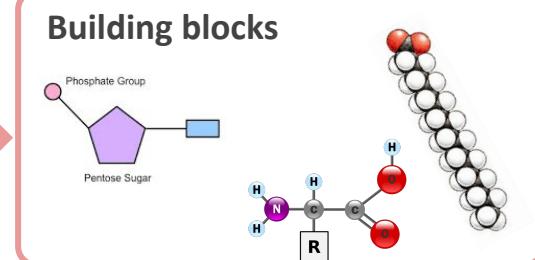
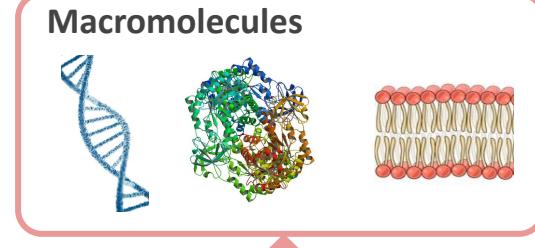
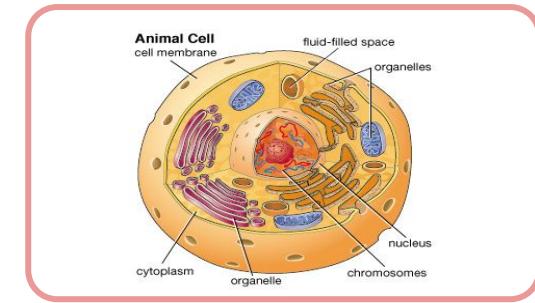
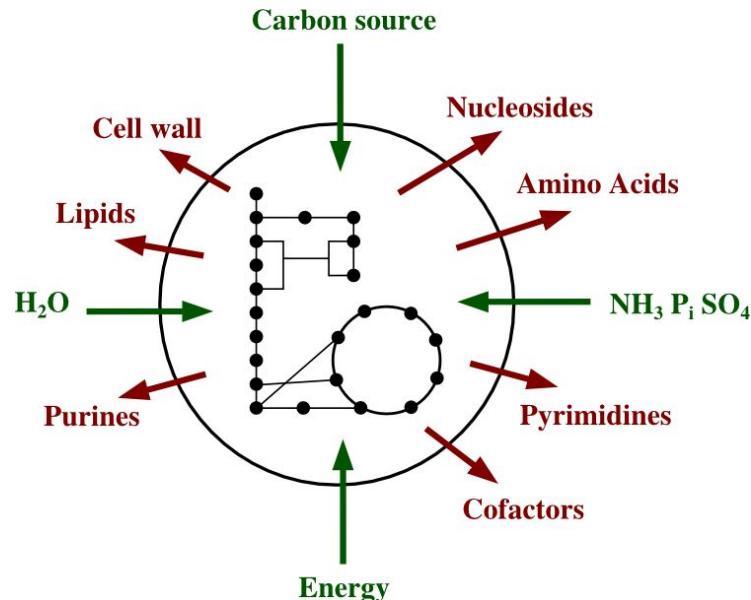
# The Biomass Equation: a molecular recipe of a cell

Stoichiometric demands of each building blocks (+energy) required to produce 1 gram of dry weight (gDW) cell

Metabolite	Demand (mmol)
ATP	41.2570
NADH	-3.5470
NADPH	18.2250
G6P	0.2050
F6P	0.0709
R5P	0.8977
E4P	0.3610
T3P	0.1290
3PG	1.4960
PEP	0.5191
PYR	2.8328
AcCoA	3.7478
OAA	1.7867
AKG	1.0789

Neidhardt, et al. *Physiology of the Bacterial Cell* (1990)

## Metabolism the cell's molecular



# Flux Balance Analysis

CB approach that finds a steady state flux distribution that maximizes biomass production (growth rate)

$$\nu_{biomass} : \sum \% \text{ precursor} \rightarrow 1\text{g DW cell}$$

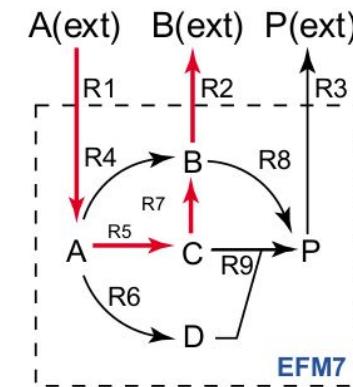
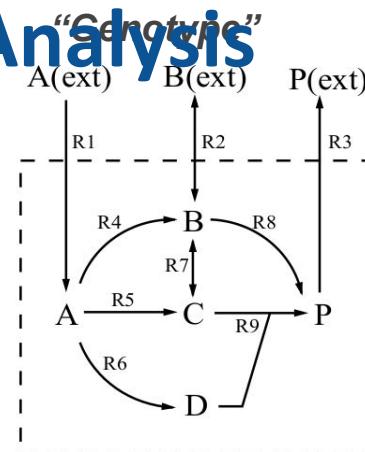
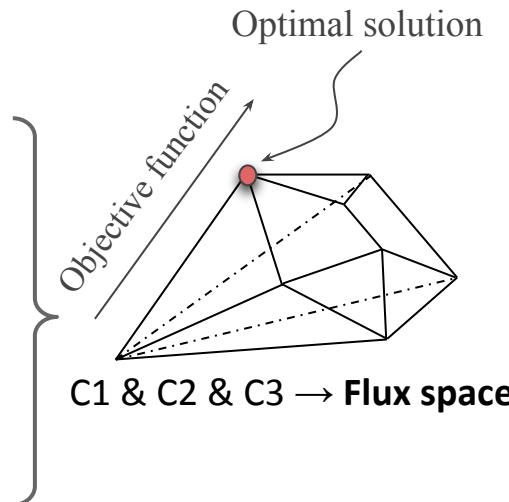
$$\text{Max: } \nu_{biomass}$$

s.t.

$$N \cdot \nu = 0$$

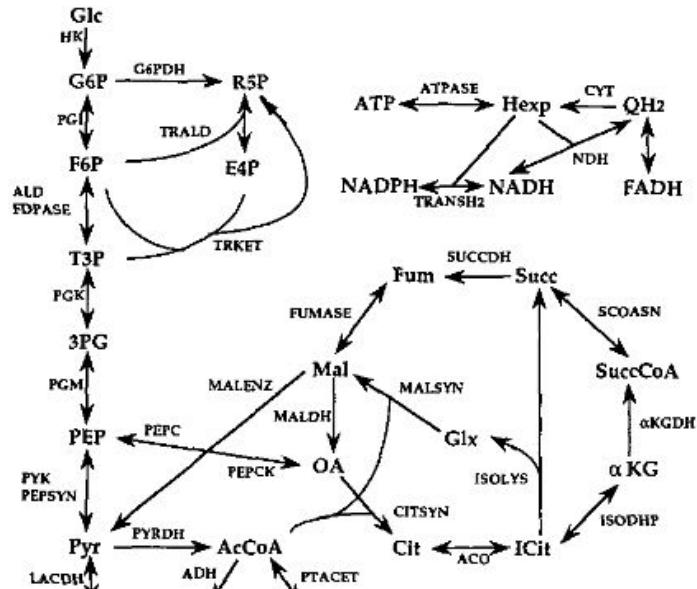
$$\nu_i \geq 0 \quad i \text{ Irrev}$$

$$Lb_i \leq \nu_i \leq Ub_i$$



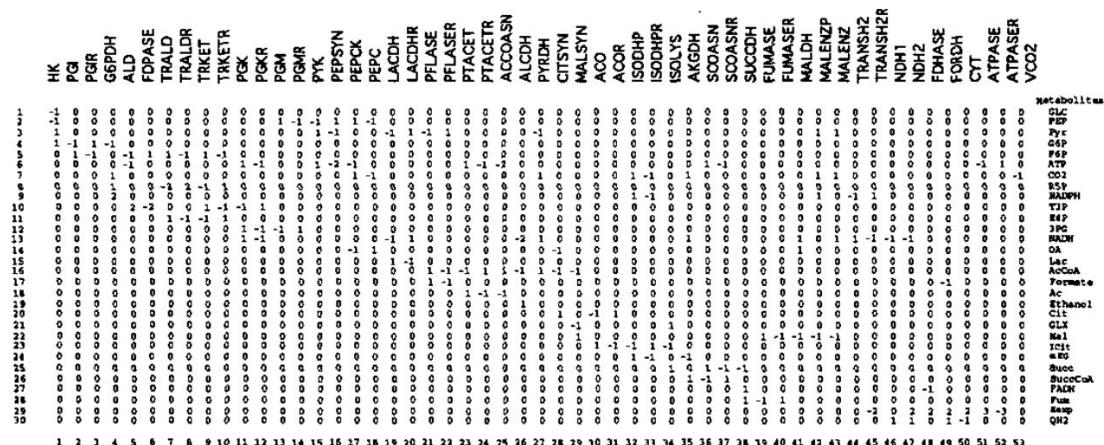
Optimal flux distribution “phenotype”

# Flux Balance Analysis

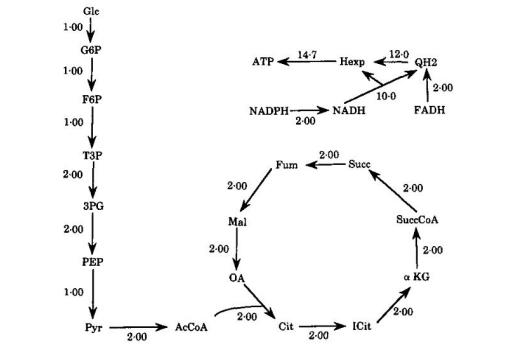
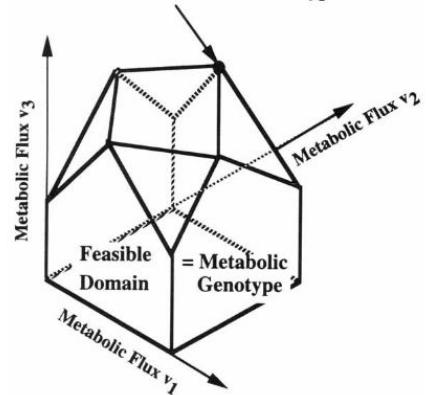


$$\sum_{\text{all } M} d_M \cdot M^{\mathbf{V}_{\text{gro}}} \rightarrow \text{biomass}$$

$$\mathbf{S} \cdot \mathbf{v} = \mathbf{b},$$



Optimum = Metabolic Phenotype



# Linear Programming Basics

Consider a systems that has to metabolites **A** and **B**.

The **production constraints** on them are:

- $0 \leq A \leq 60$ , and  $0 \leq B \leq 50$

Additionally the **capacity** for producing them simultaneously is limited by:

- $A + 2B \leq 120$

The **objective function** is:

- $Z = 20A + 30B$

**Linear Program**

$$\text{Max } Z = 20A + 30B$$

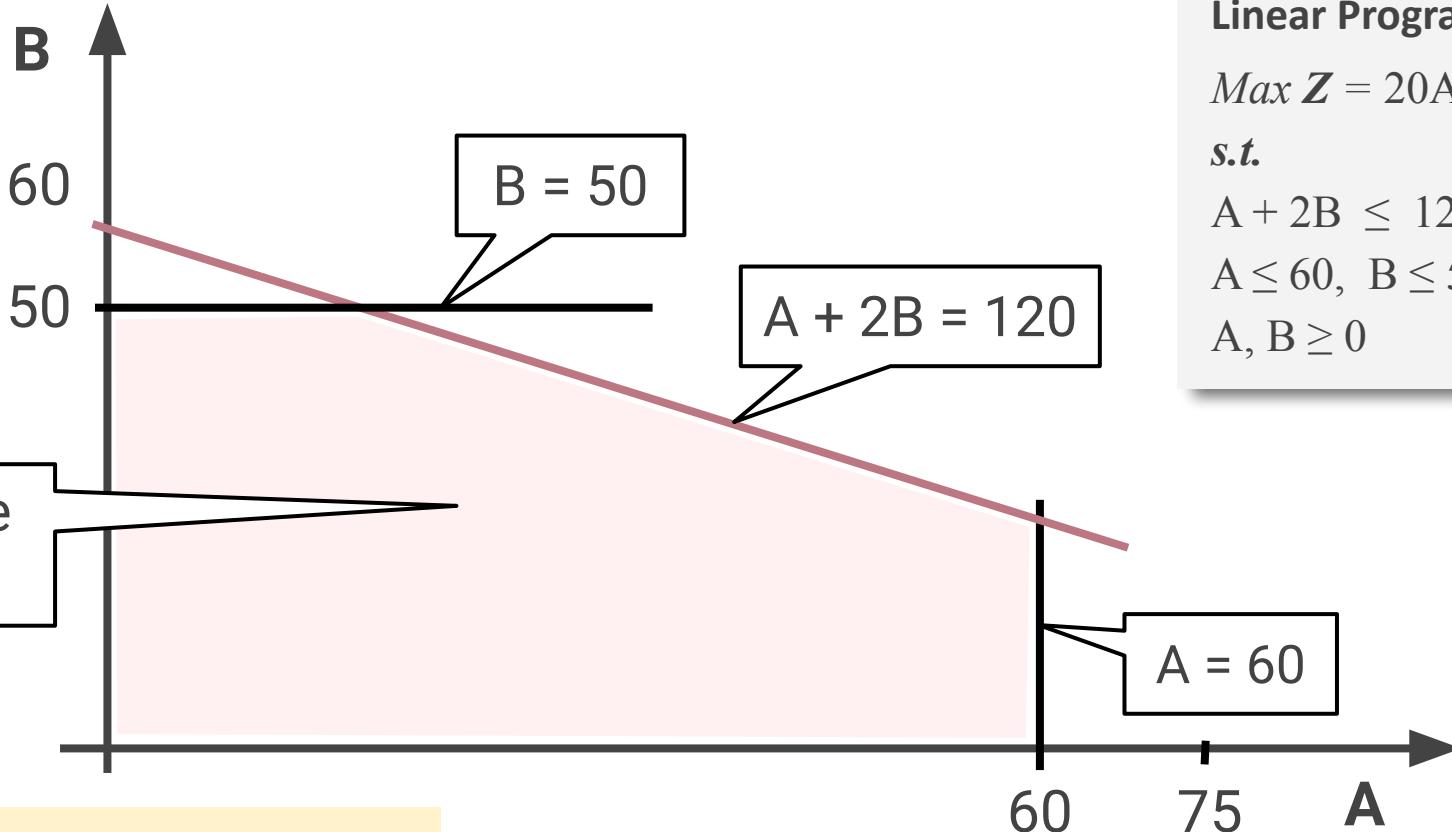
*s.t.*

$$A + 2B \leq 120$$

$$A \leq 60, B \leq 50$$

$$A, B \geq 0$$

# Linear Programming Basics



Exercise: solve the problem

**Linear Program**

$$\text{Max } Z = 20A + 30B$$

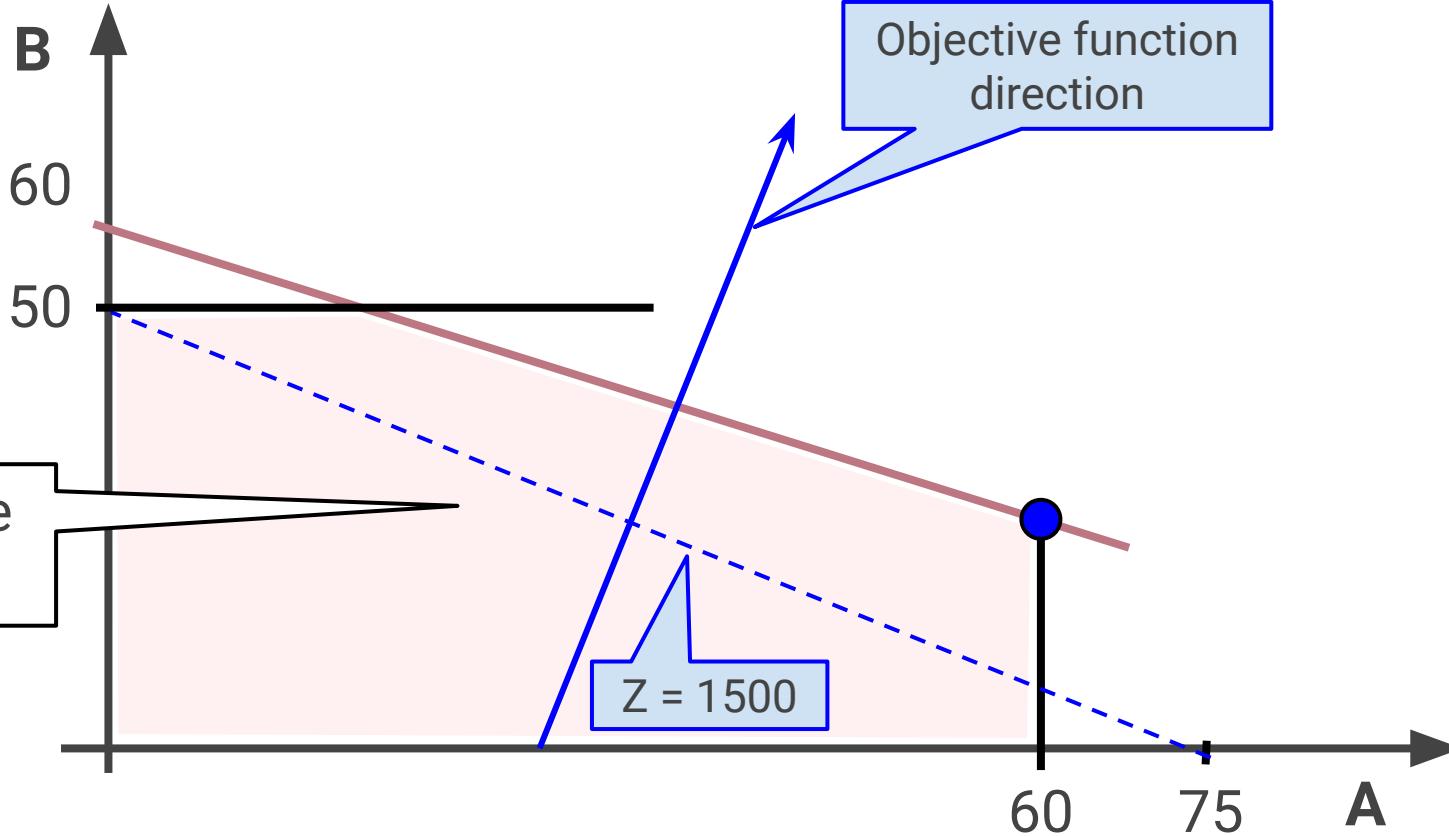
s.t.

$$A + 2B \leq 120$$

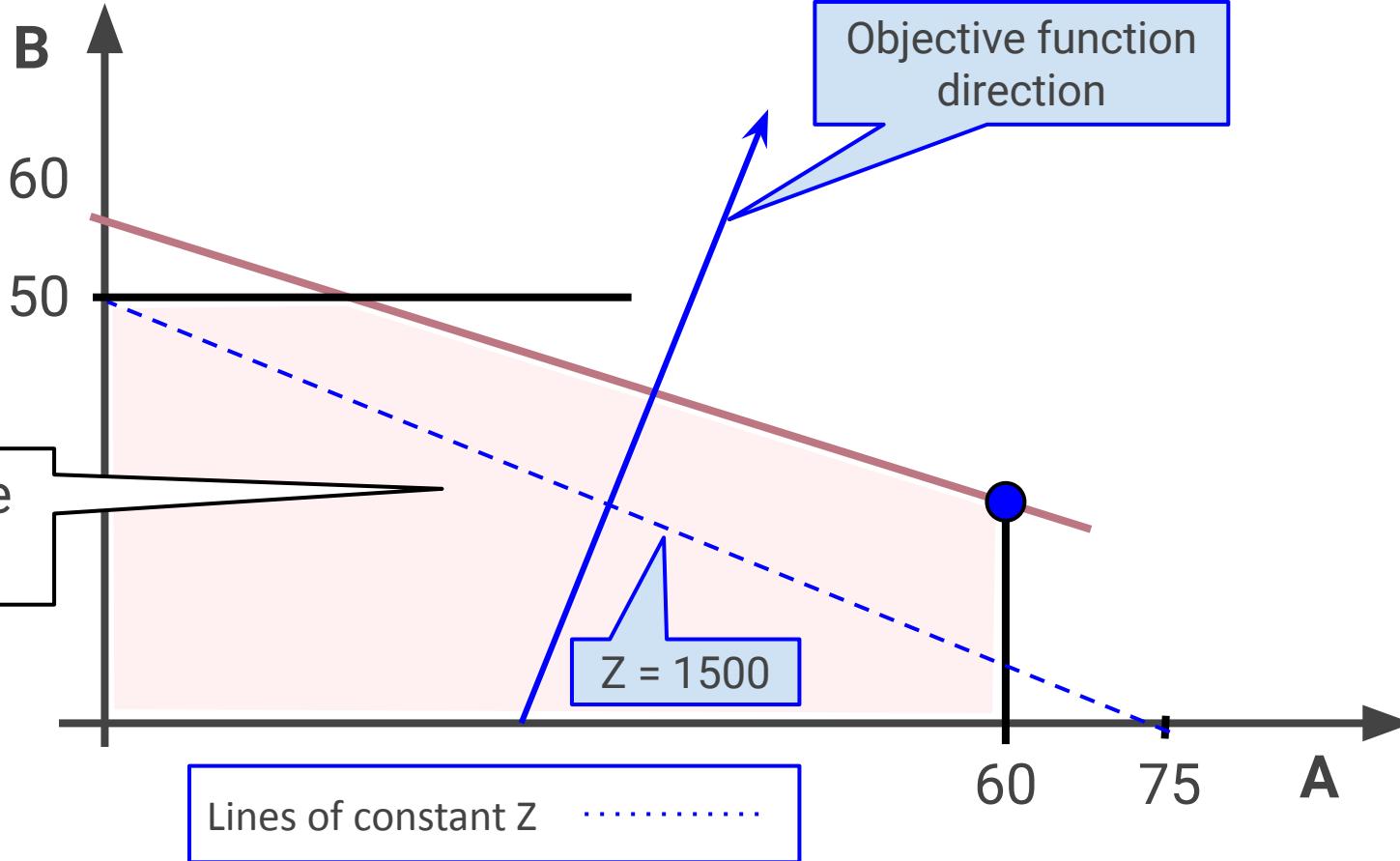
$$A \leq 60, B \leq 50$$

$$A, B \geq 0$$

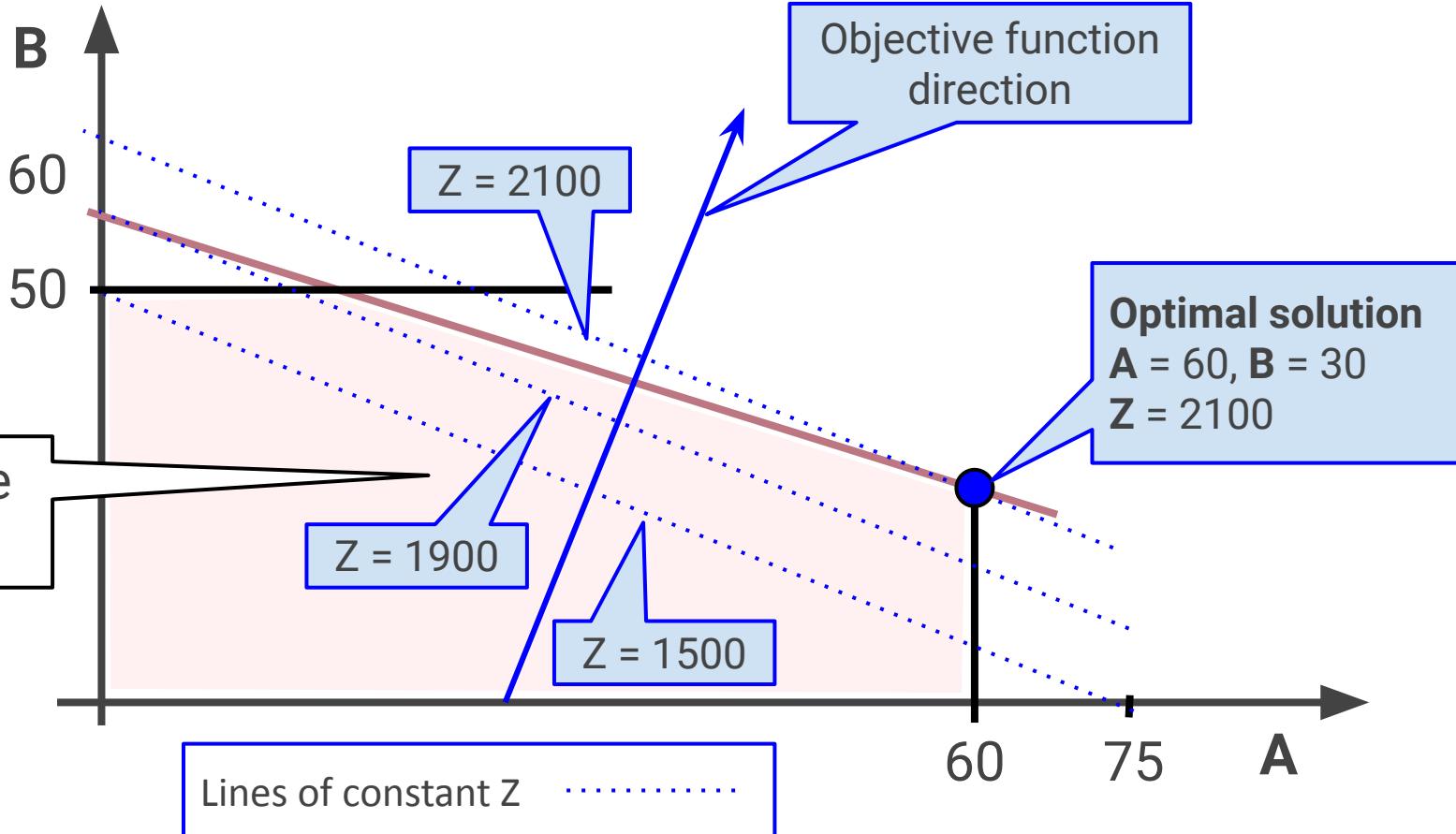
# Linear Programming Basics



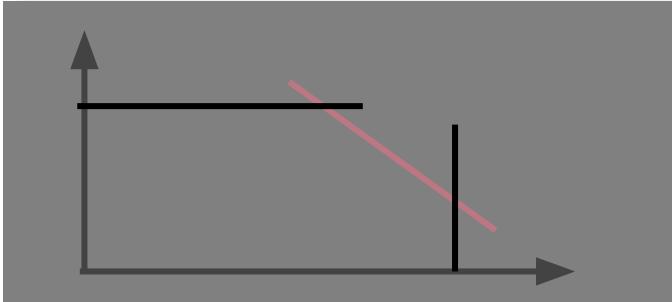
# Linear Programming Basics



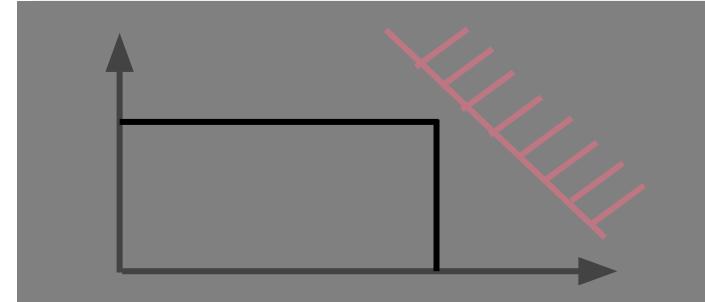
# Linear Programming Basics



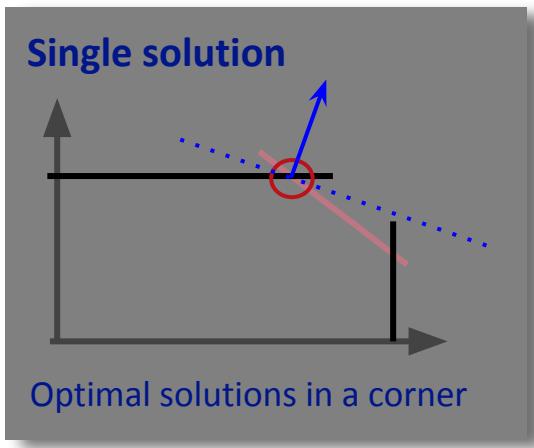
# Linear Programming Basics



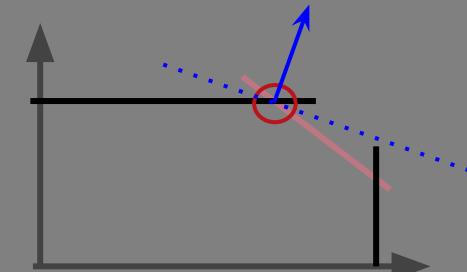
**Feasible Problem:** solutions within all the imposed constraints



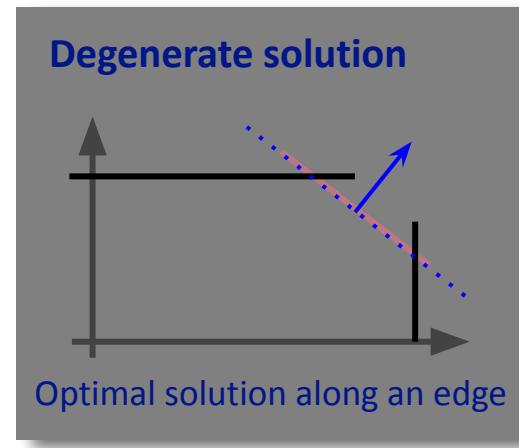
**Infeasible Problem:** solutions cannot fulfill all the imposed constraints



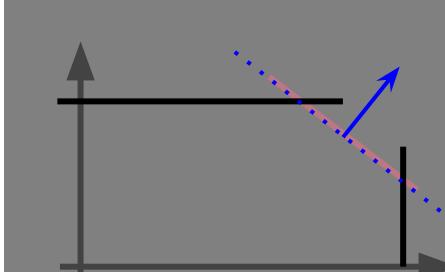
**Single solution**



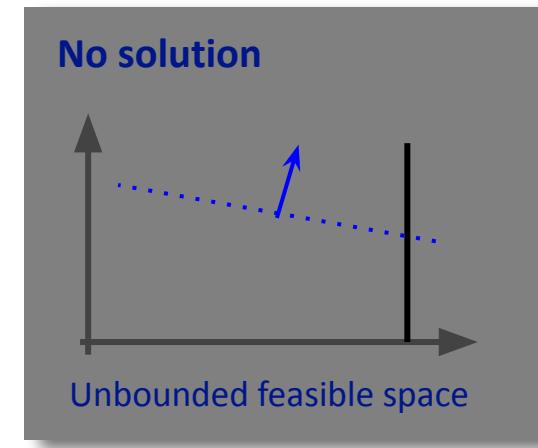
Optimal solutions in a corner



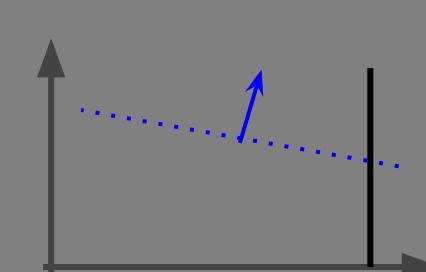
**Degenerate solution**



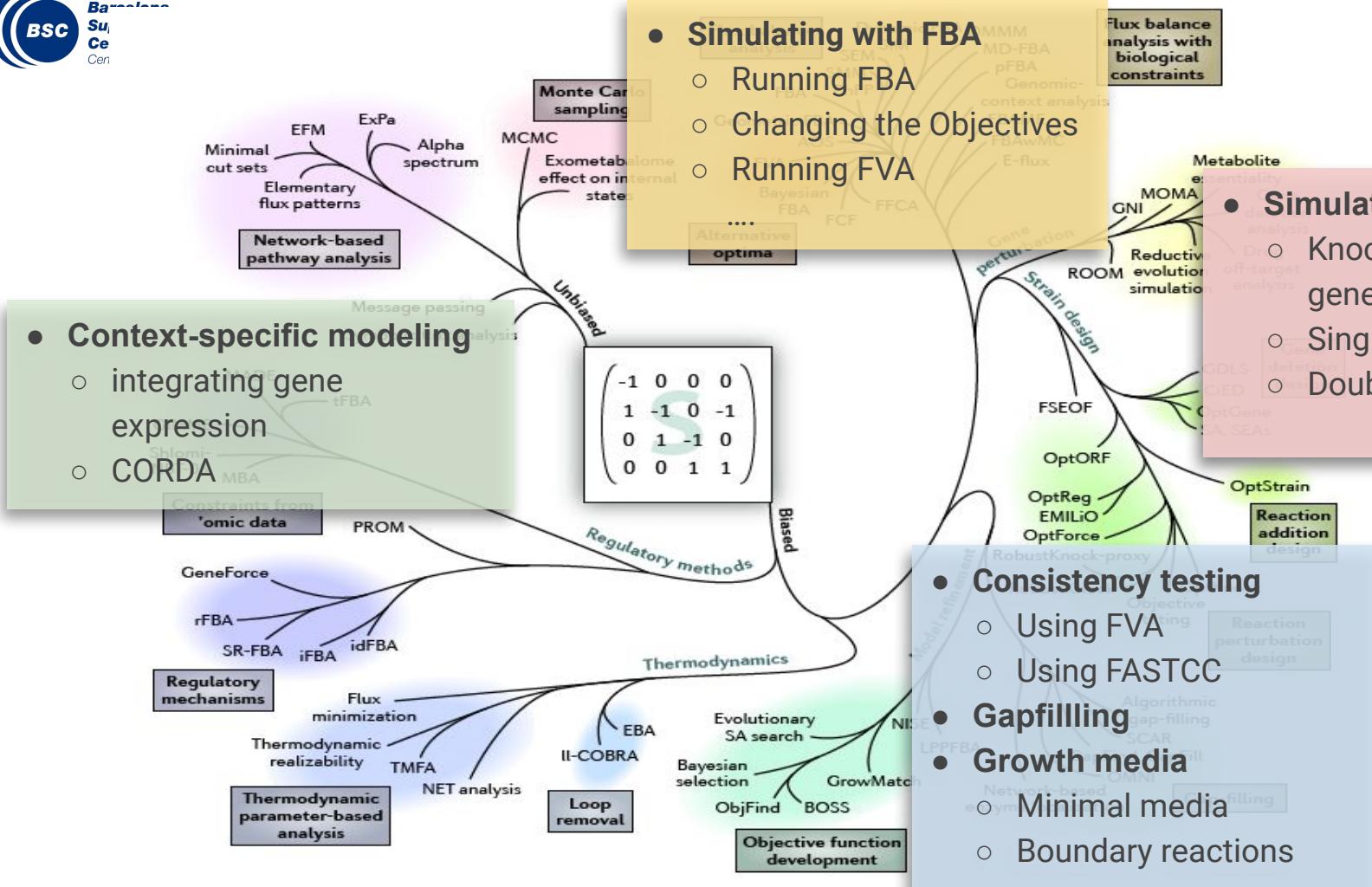
Optimal solution along an edge



**No solution**



Unbounded feasible space



# Constraint-Based Modeling: computational tools



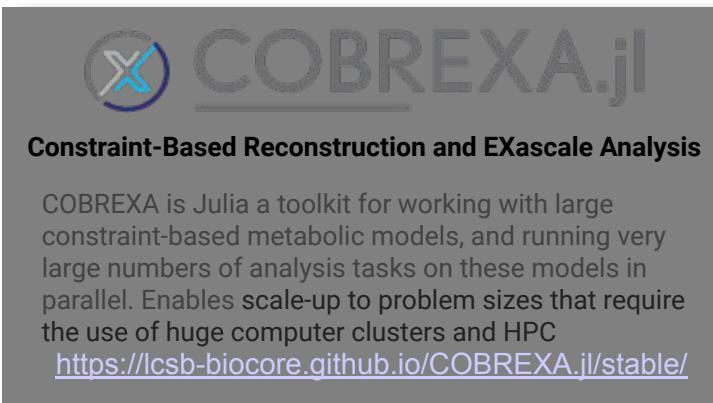
The COBRA Toolbox

The COnstraint-Based Reconstruction and Analysis (COBRA) Toolbox written in MATLAB.



COBRAPy

COBRAPy is a package for constraint-based modeling of biological networks written in Python.



**COBREXA.jl**

**Constraint-Based Reconstruction and EXascale Analysis**

COBREXA is Julia a toolkit for working with large constraint-based metabolic models, and running very large numbers of analysis tasks on these models in parallel. Enables scale-up to problem sizes that require the use of huge computer clusters and HPC

<https://lcsb-biocore.github.io/COBREXA.jl/stable/>



THE RAVEN  
TOOLBOX<sup>2</sup>

Reconstruction, Analysis and Visualization of Metabolic Networks Toolbox is a software suite for Matlab that allows for semi-automated reconstruction of genome-scale models

# Flux Balance Analysis I

*Predicting growth rates, uptakes and secretions*

## Tutorial Part 2

1.1 - Exploring models

1.2 - Predicting phenotypes

# Flux Balance Analysis

## Predicting Growth rate

Steps:

1. Set exchange fluxes bounds to define the media ( $Lb_i, Ub_i$ )
2. Solve the FBA linear problem:

$$\text{Max: } v_{\text{biomass}}$$

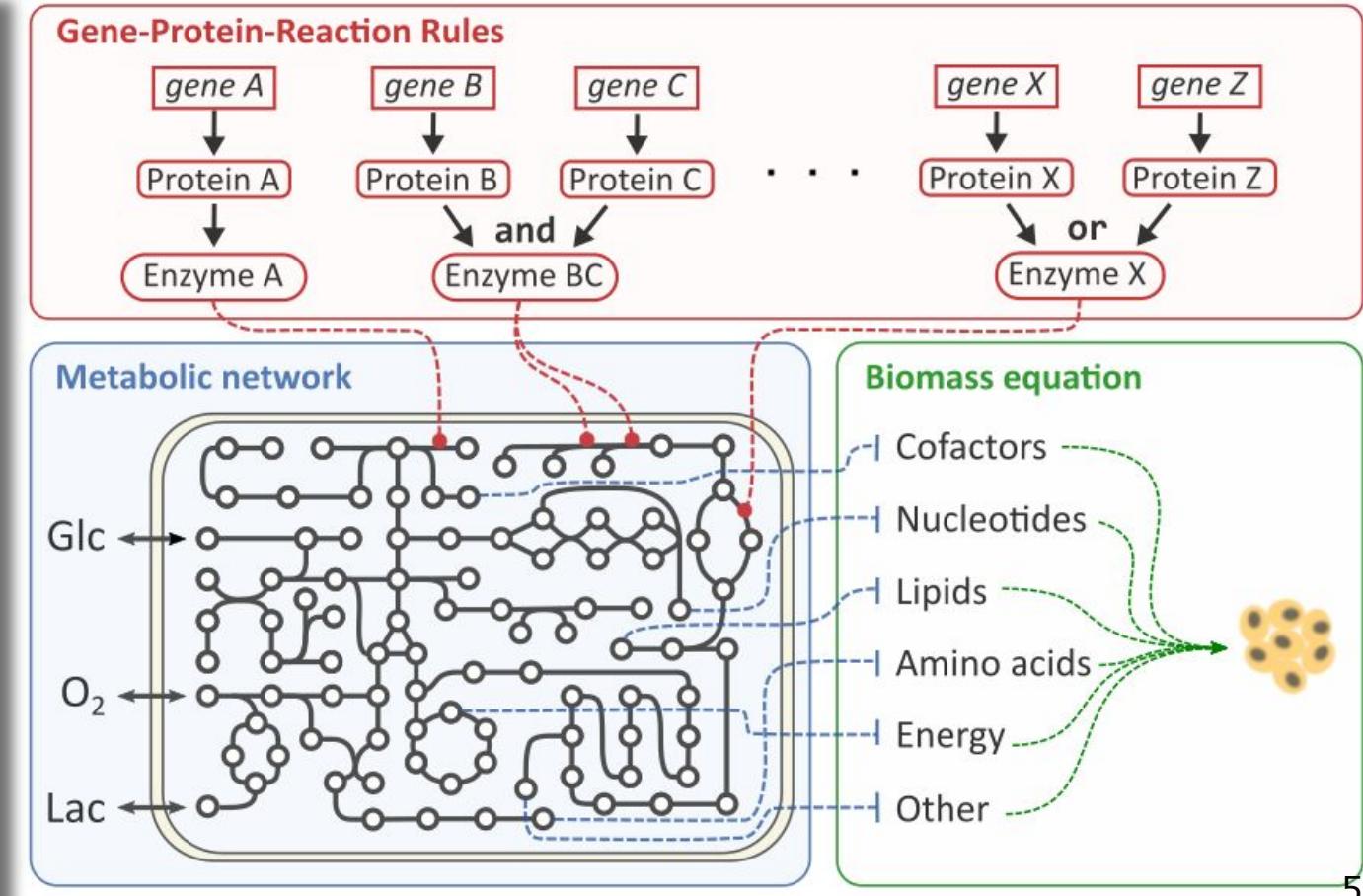
s.t.

$$N \cdot v = 0$$

$$v_i \geq 0 \quad i \text{ Irrev}$$

$$Lb_i \leq v_i \leq Ub_i$$

3. Check the value of the  $v_{\text{biomass}}$  and exchanges fluxes.



# Flux Balance Analysis

## Predicting Growth rate

Steps:

1. Set exchange fluxes bounds to define the media ( $Lb_i, Ub_i$ )
2. Solve the FBA linear problem:

$$\text{Max: } v_{\text{biomass}}$$

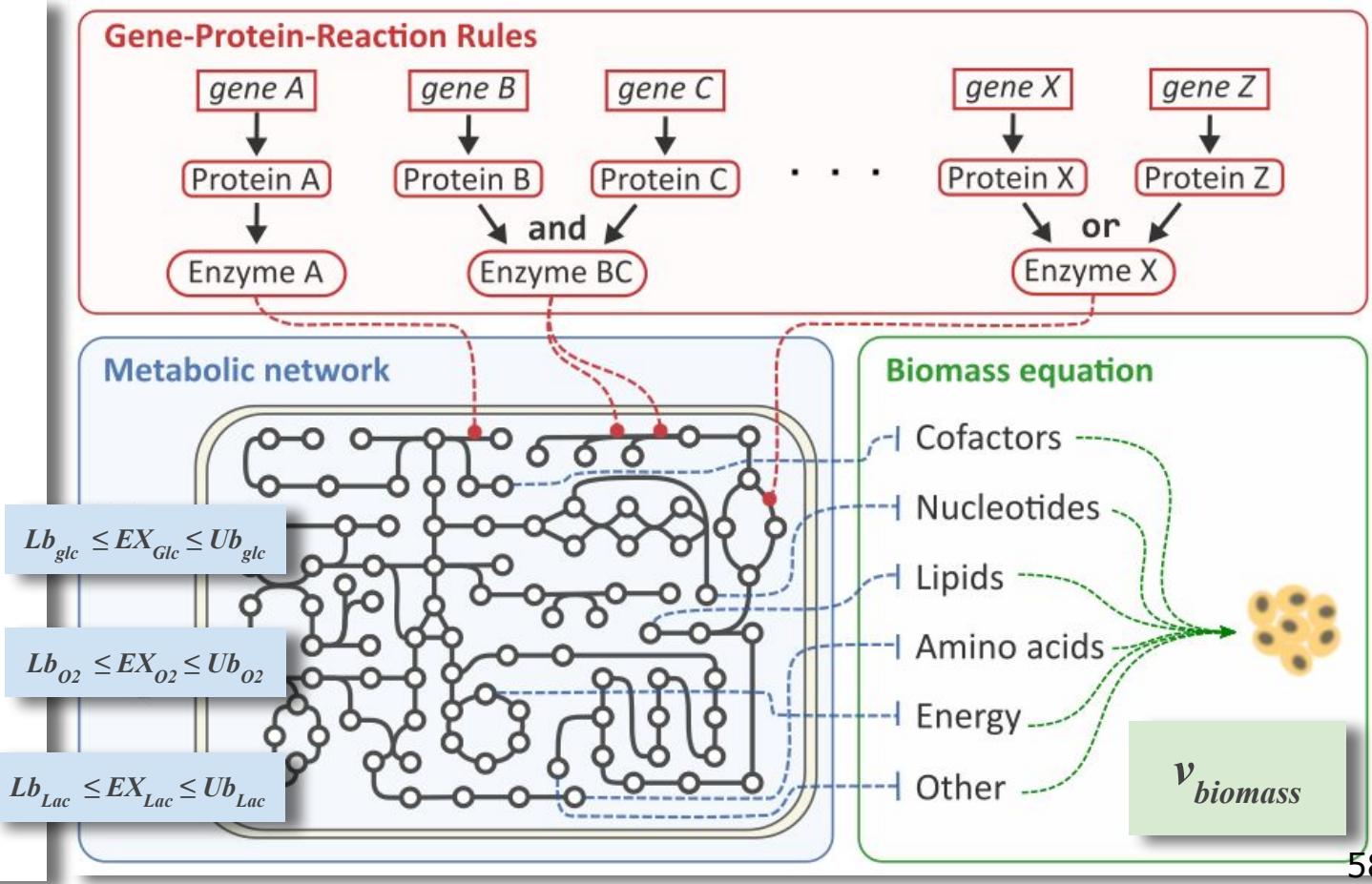
s.t.

$$N \cdot v = 0$$

$$v_i \geq 0 \quad i \text{ Irrev}$$

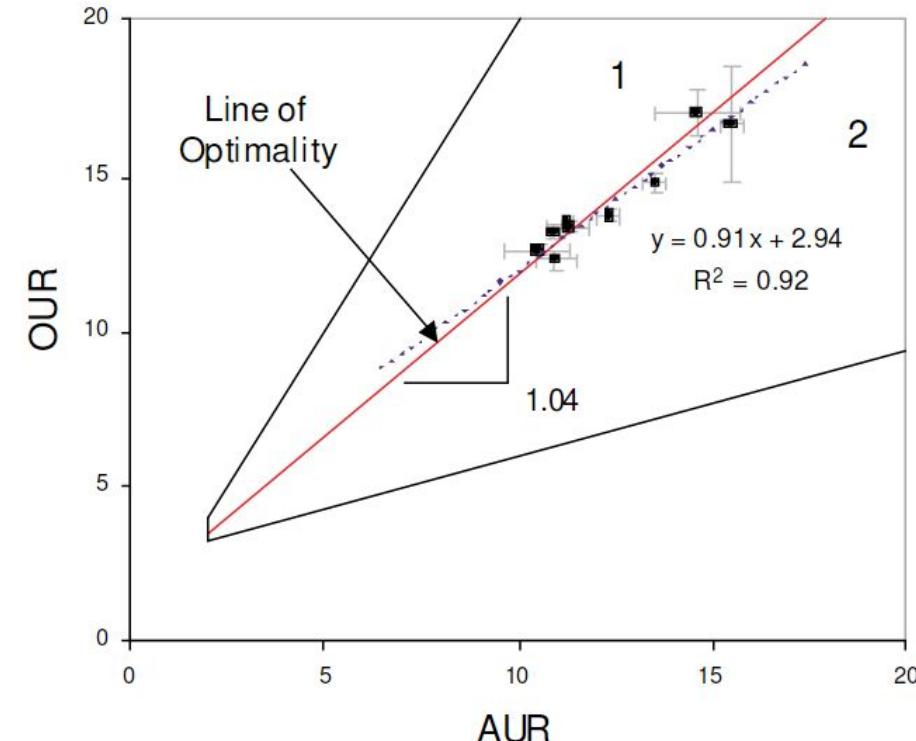
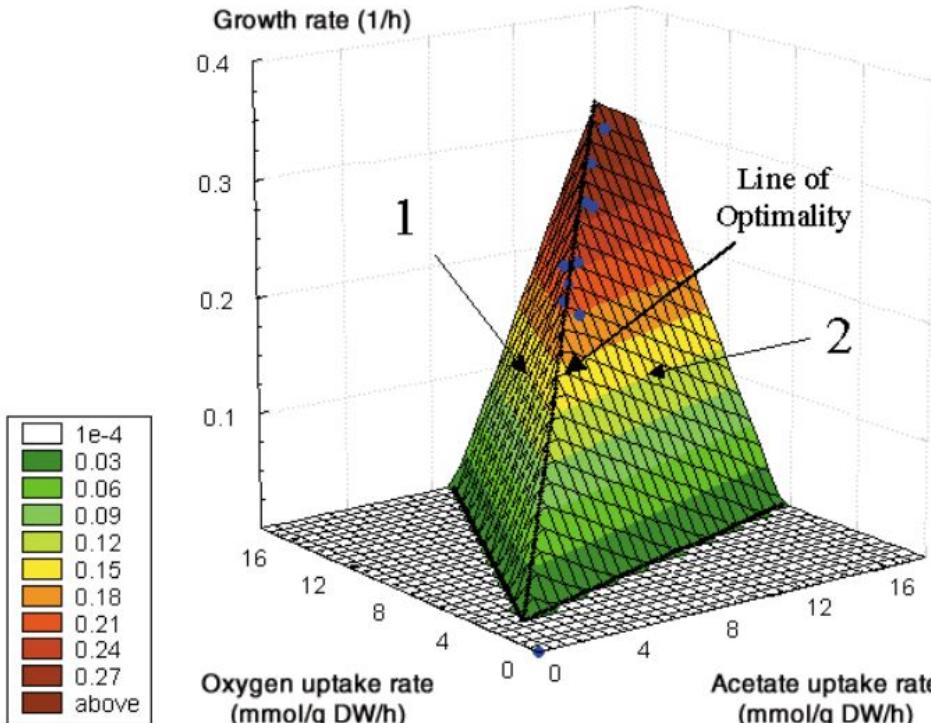
$$Lb_i \leq v_i \leq Ub_i$$

3. Check the value of the  $v_{\text{biomass}}$  and exchanges fluxes.



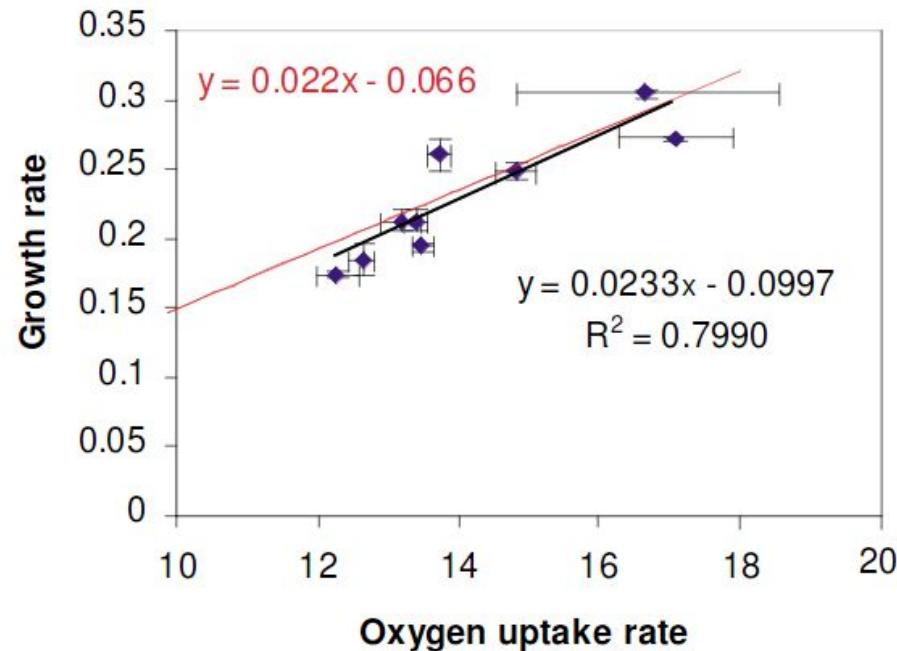
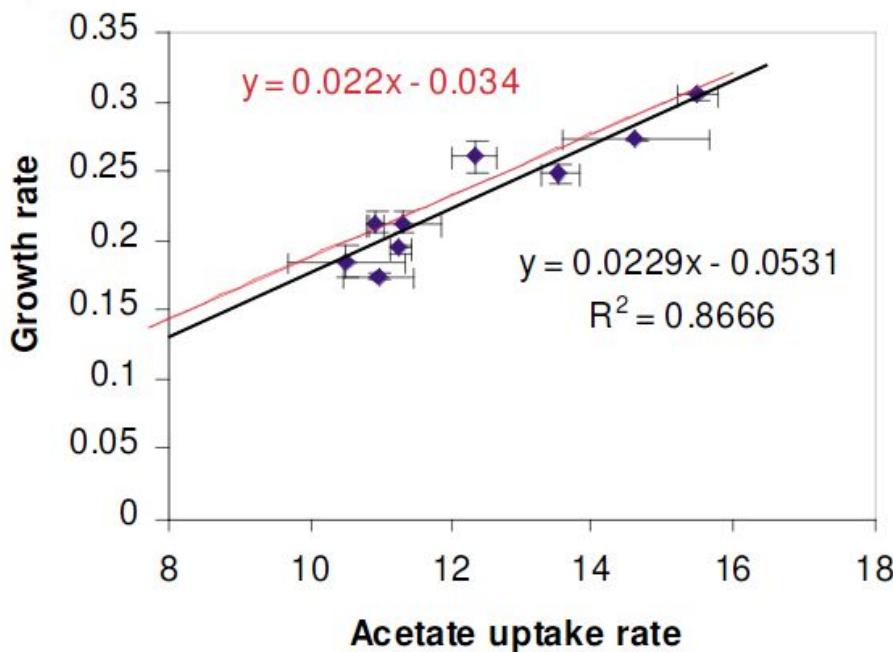
# Predicting growth rate for different carbon sources

Comparing predictions for E. coli with wet-lab experimental measurements



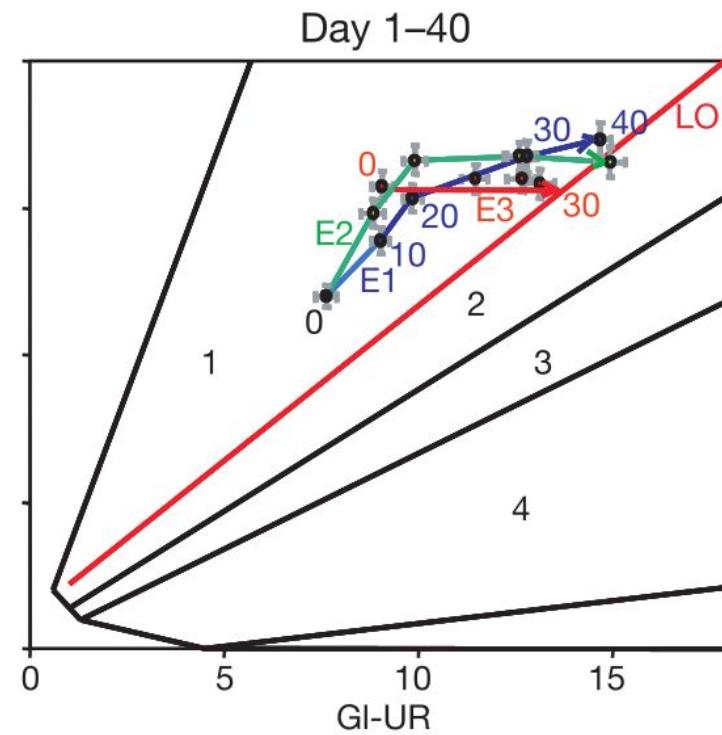
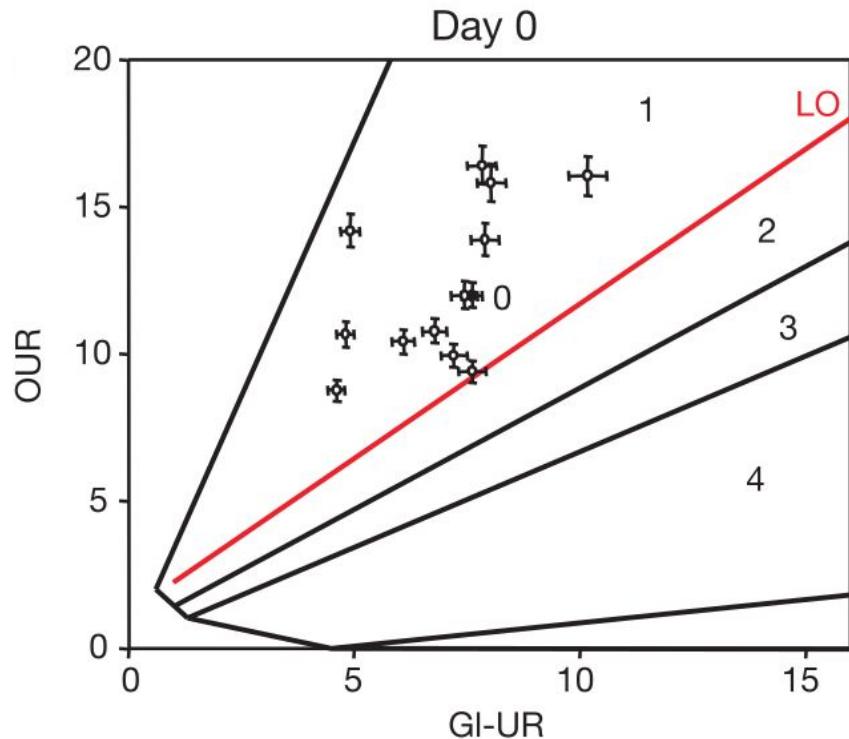
# Predicting growth rate for different carbon sources

Comparing predictions for E. coli with wet-lab experimental measurements



# Prediction of microbial evolution by FBA (*E. coli*)

Comparing predictions for *E. coli* with wet-lab experimental measurements

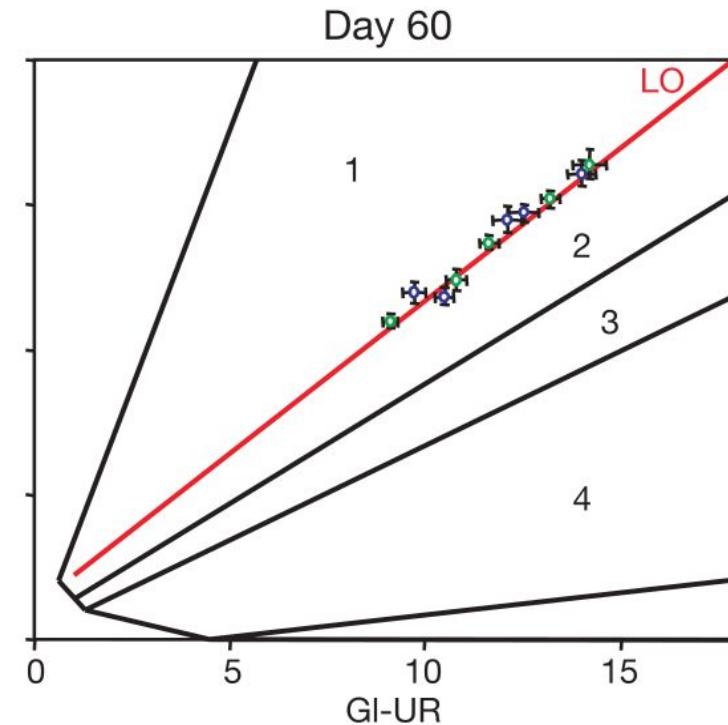
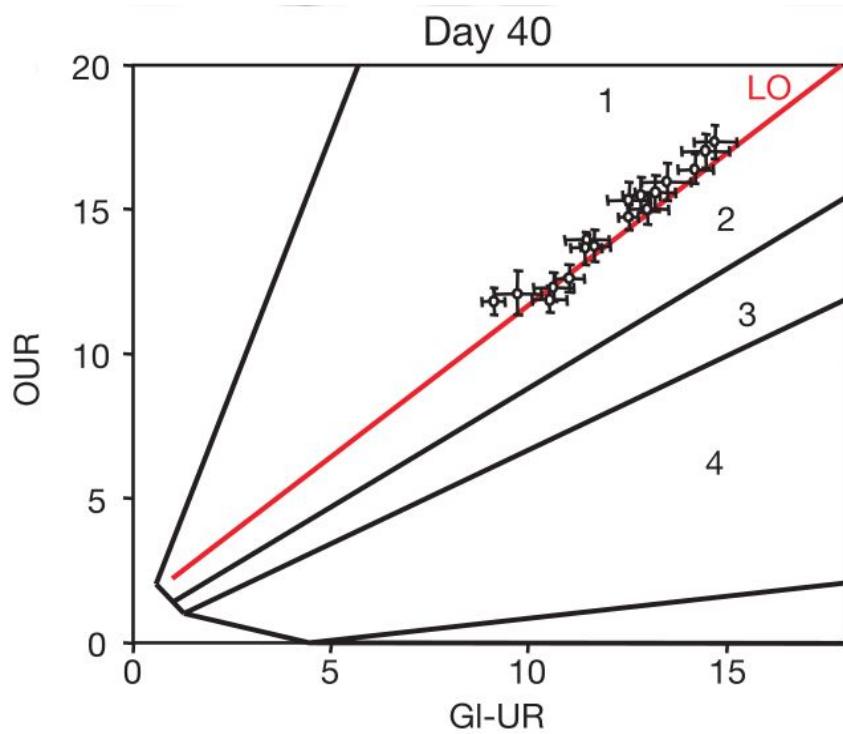


Rafael U. Ibarra<sup>\*†</sup>, Jeremy S. Edwards<sup>†‡</sup> & Bernhard O. Palsson<sup>\*</sup>

NATURE | VOL 420 | 14 NOVEMBER 2002

# Prediction of microbial evolution by FBA (*E. coli*)

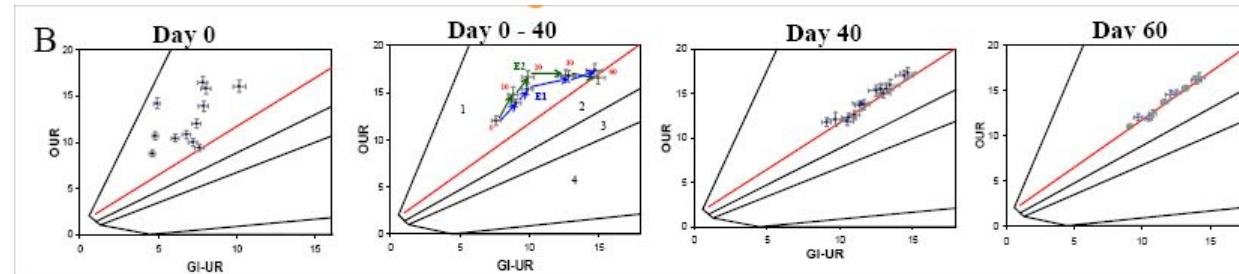
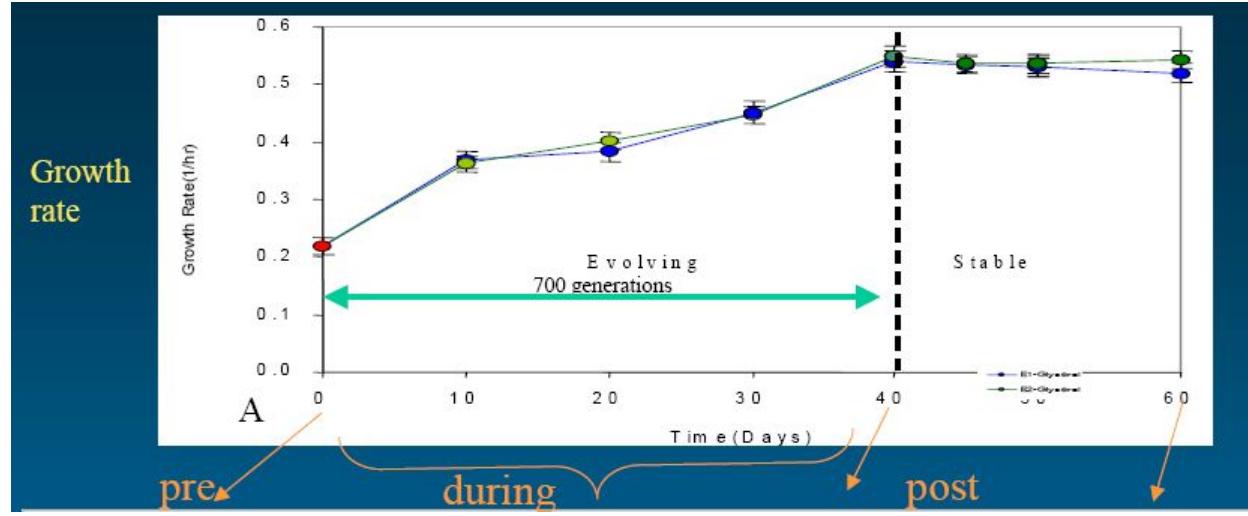
Comparing predictions for *E. coli* with wet-lab experimental measurements



Rafael U. Ibarra<sup>\*†</sup>, Jeremy S. Edwards<sup>†‡</sup> & Bernhard O. Palsson<sup>\*</sup>

NATURE | VOL 420 | 14 NOVEMBER 2002

# Prediction of microbial evolution by FBA (*E. coli*)

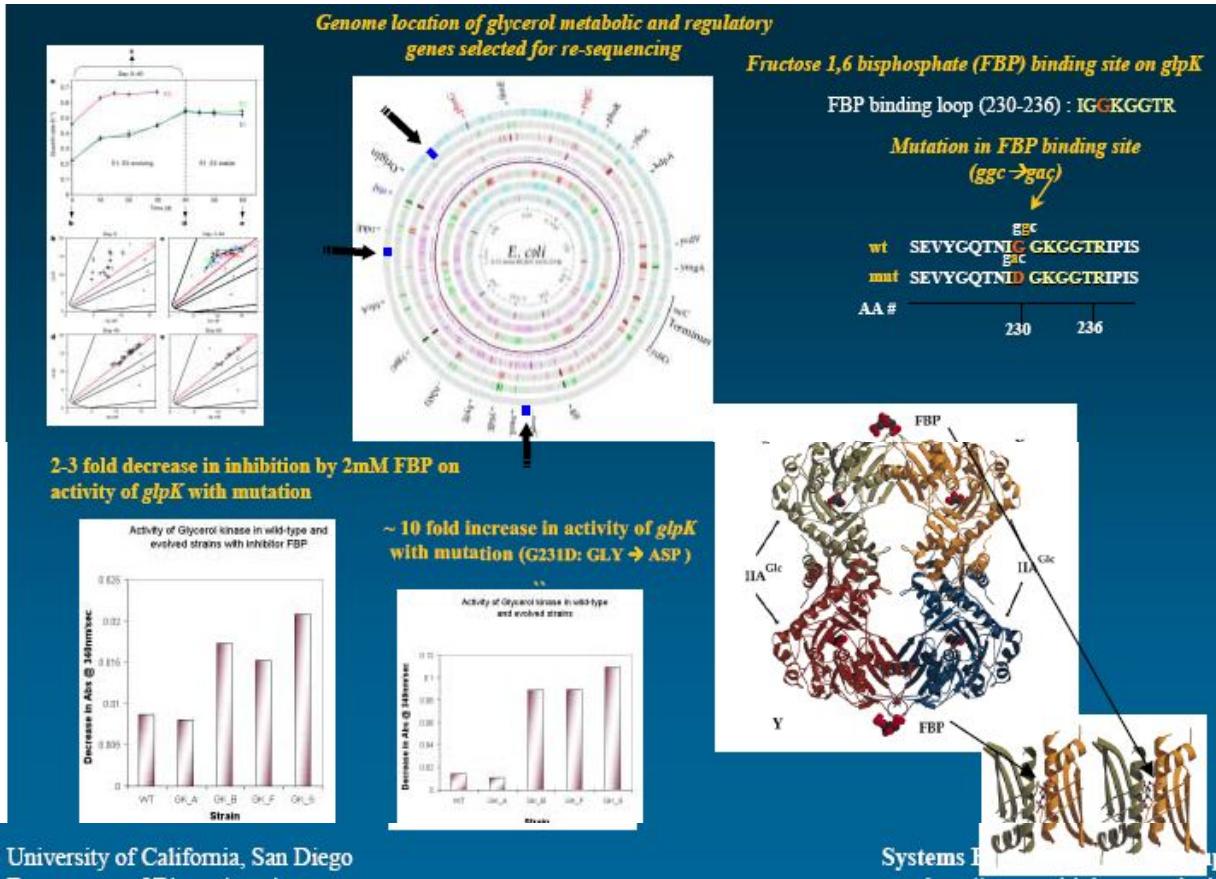


Rafael U. Ibarra<sup>\*†</sup>, Jeremy S. Edwards<sup>†‡</sup> & Bernhard O. Palsson<sup>\*</sup>

NATURE | VOL 420 | 14 NOVEMBER 2002

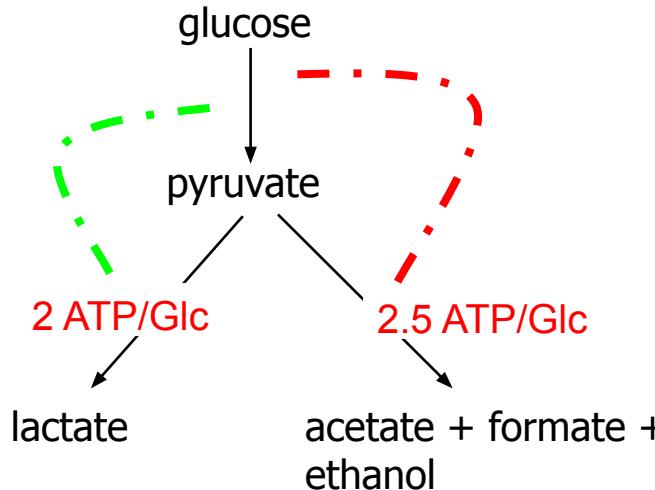
# Prediction of microbial evolution by FBA (E. coli)

Comparing predictions for E. coli with wet-lab experimental measurements



# FBA fails to predict *L. plantarum* growth

FBA does not consider many constraint such regulation and thus overestimate the metabolic state

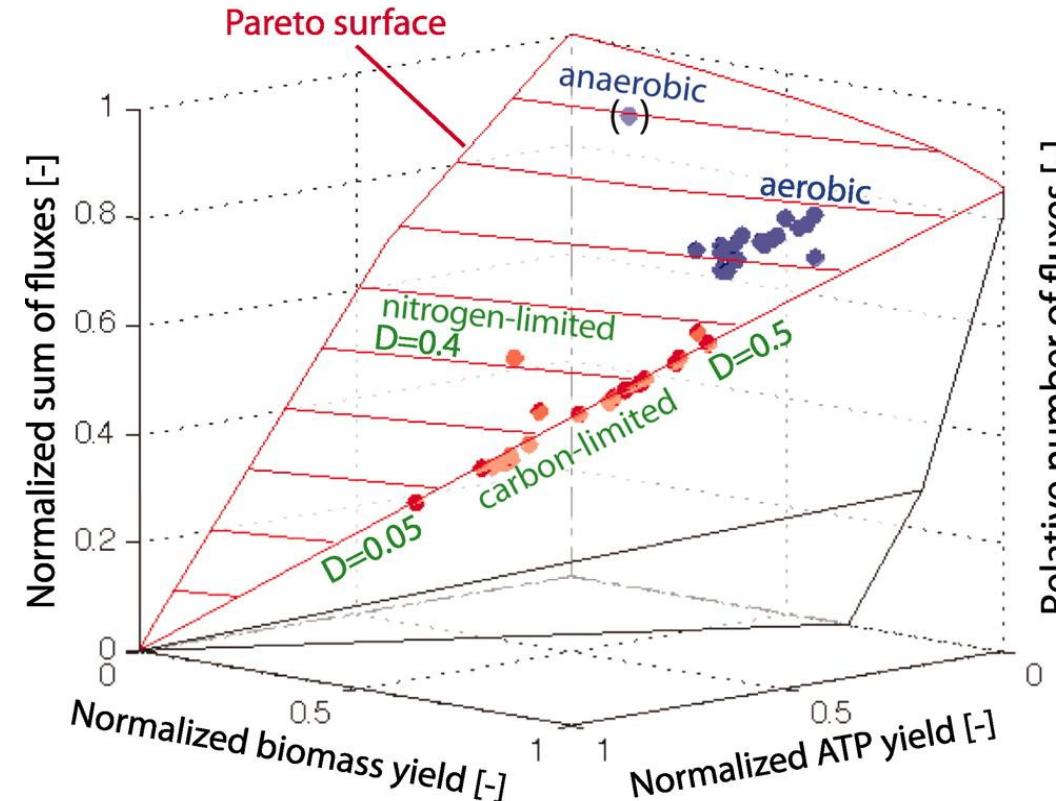


D = 0.32 h <sup>-1</sup>	simulation	experiment
<b>lactate</b>	0	13.8
<b>Pyruvate</b>	0	0.1
<b>formate</b>	8.0	0.9
<b>acetate</b>	15.2	1.9
<b>acetoin</b>	0	0
<b>EtOH</b>	3.4	0.45
<b>citric acid</b>	-0.7	-0.7
<b>succinate</b>	0	0.9
<b>Glucose</b>	-8.4	-8.4

FBA predicts mixed acid fermentation with 40% too high biomass formation  
← thus *L. plantarum* is not using its metabolism in the “most efficient” way

# More complex strategies: multi-objective optimization

*B. subtilis* growth strategy can be predicted using a multiobjective optimization (three targets)



## Part 3. Summary

- FBA can predict microbial growth in different media conditions.
- It correctly predicts **growth rate, uptakes and secretions** rates in laboratory conditions.
- It predicts the maximum capacity of network using the stoichiometry, constraints, and linear programming and can be used for engineering microorganism.
- The maximization of the the growth rate as the optimization target has several **limitations** in predicting metabolic states in natural environments.
- Cell composition and growth strategies:
  - vary from one organism to another
  - depend on the growing medium
  - depend on the growth rate
  - relative composition of the macromolecules changes at different growth rate

# Flux Balance Analysis II

## *In-silico gene knockouts*

### Tutorial Part 3

3.1 - The gene protein reaction rules

3.2 - Gene knockouts

# Flux Balance Analysis: *in-silico* gene knockouts

## *In-silico* gene knock-out

### Steps:

- Find the reactions ( $J_{KO(B)}$ ). that requires the *gene A* using GPRs.
- Set the bounds of inactivated reactions in  $J_{KO(B)}$  to  $(0, 0)$
- Solve the FBA linear problem:

$$\text{Max: } v_{\text{biomass}}$$

s.t.

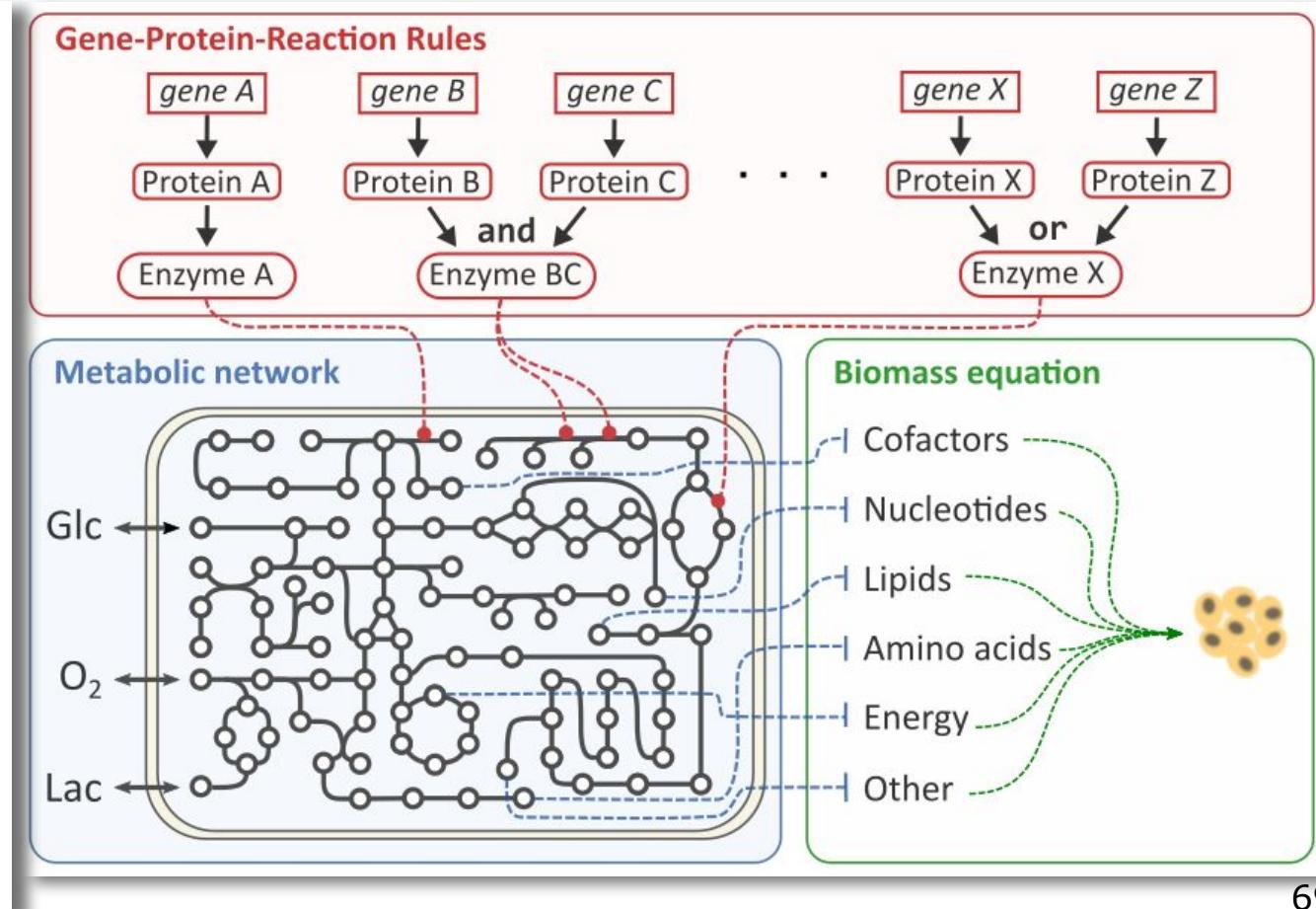
$$N \cdot v = 0$$

$$v_i \geq 0 \text{ for all } i \text{ in } J_{\text{Irrev}}$$

$$Lb_i \leq v_i \leq Ub_i$$

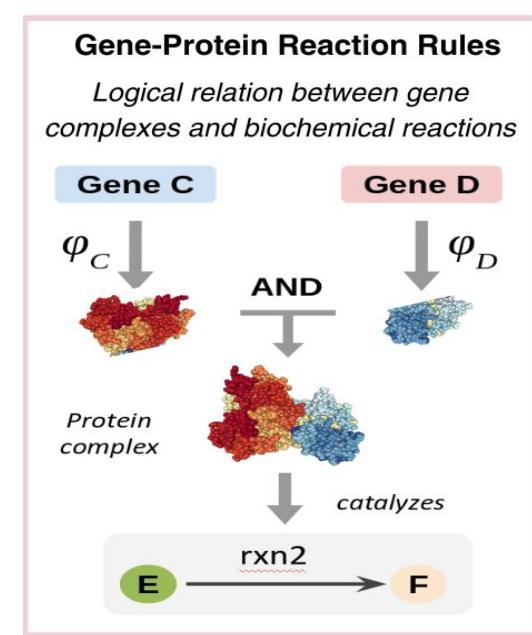
$$Lb_i = Ub_i = 0 \text{ for all } i \text{ in } J_{KO(b)}$$

- Check  $v_{\text{biomass}} \geq \text{threshold}$



# Gene Protein Reaction rules: examples from *E. coli* iJO1366

<b>PGI</b> ←	<b>b4025</b>
<b>PGK</b> ←	<b>b2926</b>
<b>NDPK1</b> ←	<b>b0474 or b2518</b>
<b>NDPK2</b> ←	<b>b0474 or b2518</b>
<b>MPTS</b> ←	<b>b0784 and b0785</b>
<b>G3PD6</b> ←	<b>b2241 and b2242 and b2243</b>
<b>G3PD7</b> ←	<b>b2241 and b2242 and b2243</b>
<b>ASPK</b> ←	<b>b0002 or b3940 or b4024</b>
<b>MECDPDH5</b> ←	<b>( b2895 and b2515 ) or ( b0684 and b2515 )</b>
<b>RNDR1</b> ←	<b>( b2234 and b2235 and b2582 ) or ( b2234 and b2235 and b3781 )</b>
<b>MPTG</b> ←	<b>( b3396 and b1069 ) or ( b0149 and b1069 ) or ( b1069 and b2519 )</b>
<b>MEOHtex</b> ←	<b>b1377 or b0929 or b2215 or b0241</b>



# Flux Balance Analysis: *in-silico* gene knockouts

## *In-silico* gene knock-out

### Steps:

- Find the reactions ( $J_{KO(B)}$ ) that requires the *gene B* using GPRs.
- Set the bounds of inactivated reactions in  $J_{KO(B)}$  to  $(0, 0)$
- Solve the FBA linear problem:

$$\text{Max: } v_{\text{biomass}}$$

s.t.

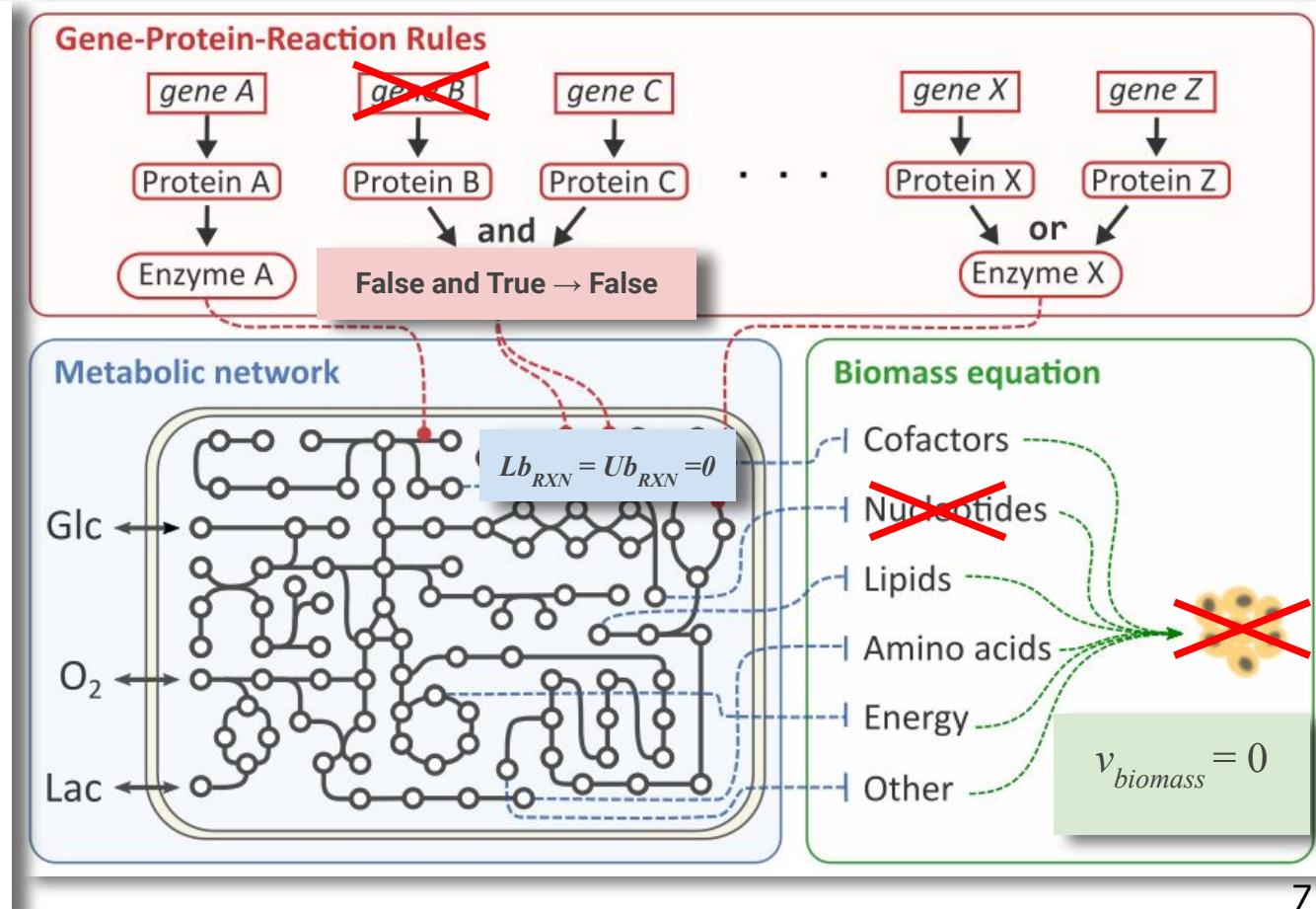
$$N \cdot v = 0$$

$$v_i \geq 0 \text{ for all } i \text{ in } J_{\text{Irrev}}$$

$$Lb_i \leq v_i \leq Ub_i$$

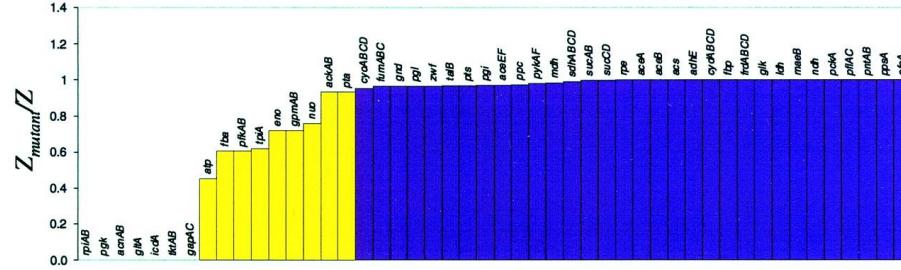
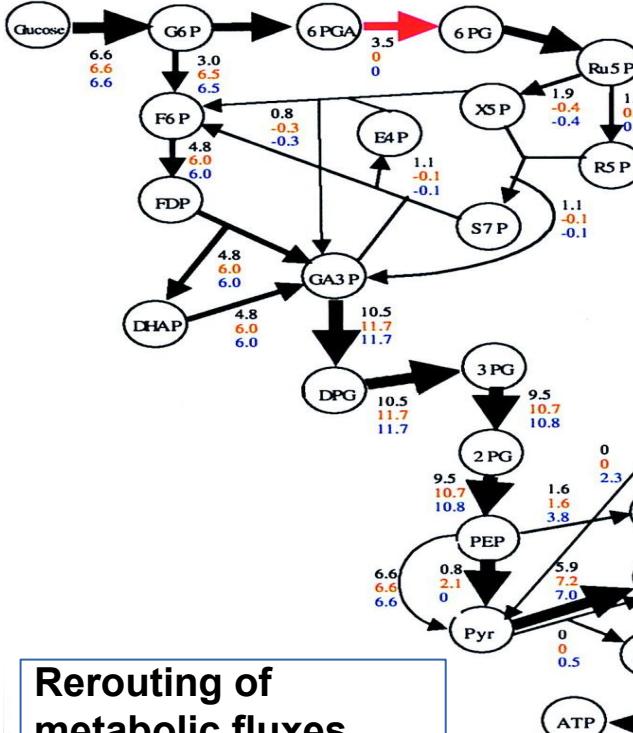
$$Lb_i = Ub_i = 0 \text{ for all } i \text{ in } J_{KO(B)}$$

- Check  $v_{\text{biomass}} \geq \text{threshold}$



# Flux Balance Analysis: *in-silico* gene knockouts

Predicting essential genes in *E. coli*

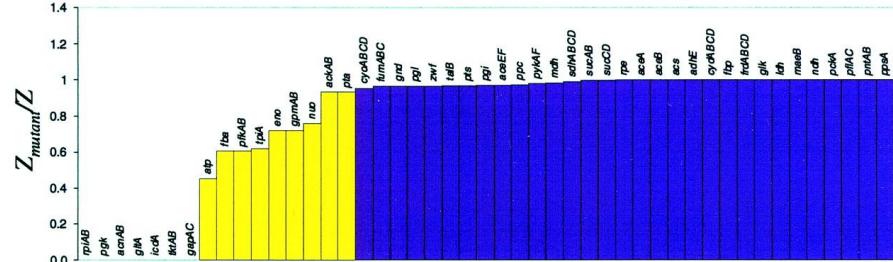
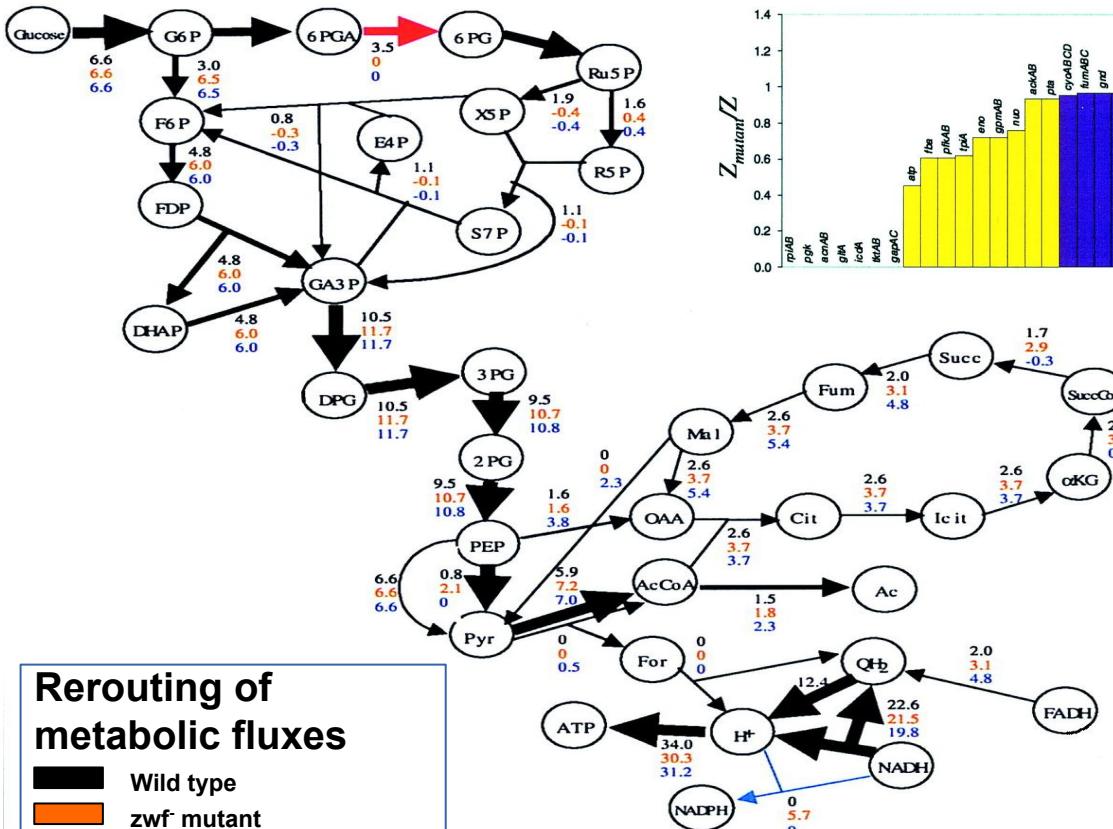


Rerouting of metabolic fluxes

- Wild type
- zwf mutant
- zwf pnt<sup>+</sup> double mutant

# Flux Balance Analysis: *in-silico* gene knockouts

Predicting essential genes in *E. coli*



**86% (68 of 79 cases) of the *in-silico* predictions were consistent with the experimental data.**

## Rerouting of metabolic fluxes

- Wild type
- zwf mutant
- zwf pnt<sup>-</sup> double mutant

Edwards JS, Palsson BO. PNAS (2000)

Why does it fails in 11 cases?  
Any explanation?  
Two types of errors

- **False Positives**
- **False Negatives**

# Part 4. Summary

- FBA can be used to **predicts essential genes** while performing *in-silico* knock-outs.
- It allow to test if the cell is able to produce all biomass components after a genetic perturbation
- Predictions are expected to be between different media conditions.
- **86% (68 of 79 cases)** of the *in-silico* predictions were consistent with the experimental data.
- **Wrong predictions** can be classified as:
  - **false positives:** *in-silico* predicted → non-essential / *in-vivo* found as essential
  - **false negatives:** *in-silico* predicted → essential / *in-vivo* found as non-essential
- **Limit of predictions:**
  - **false positives:** the model is overestimating what the cell can do:
    - not considering kinetics, regulation or other constraints
    - Wrong GPRs or reactions stoichiometry..
  - **false negatives:** the model underestimate cell capacities
    - missing information in the model
    - Wrong GPRs or biomass definition
- There are other methods that improves FBA *in-silico* knock-outs: **MOMA, ROOM**

- Edwards, J. S. & Palsson, B. O. The *Escherichia coli* MG1655 in silico metabolic genotype: its definition, characteristics, and capabilities. *Proceedings of the National Academy of Sciences of the United States of America* 97, 5528–33 (2000).
- Edwards, J. S., Ibarra, R. U. & Palsson, B. O. In silico predictions of *Escherichia coli* metabolic capabilities are consistent with experimental data. *Nature biotechnology* 19, 125–30 (2001).
- Segrè, D., Vitkup, D. & Church, G. M. Analysis of optimality in natural and perturbed metabolic networks. *Proceedings of the National Academy of Sciences of the United States of America* 99, 15112–7 (2002).
- Klamt, S. & Stelling, J. Two approaches for metabolic pathway analysis? *Trends in biotechnology* 21, 64–9 (2003).
- Lewis, N. E. et al. Omic data from evolved *E. coli* are consistent with computed optimal growth from genome-scale models. *Molecular systems biology* 6, 390 (2010).
- Lewis, N. E., Nagarajan, H. & Palsson, B. O. Constraining the metabolic genotype-phenotype relationship using a phylogeny of in silico methods. *Nature reviews. Microbiology* 10, 291–305 (2012).
- Schuetz, R., Zamboni, N., Zampieri, M., Heinemann, M. & Sauer, U. Multidimensional optimality of microbial metabolism. *Science* 336, 601–604 (2012).

# Metabolic modeling of Human cells

*Integrating omic data in metabolic models to provide  
biological context*

# Genome-Scale Model of Human Metabolism

Table 1

## Statistics of currently published generic human GEMs.

Generic GEMs	Genes	Metabolites <sup>a</sup>	Reactions <sup>a</sup>	Features
RECON1	1496	1509	3744	Manually reconstructed from bibliomics data
EHMN	2322	2671	2823	Manually reconstructed from bibliomics data
RECON2	1789	2626	7440	Merging EHMN and HepatoNet1 with RECON1
RECON 2.2	1675	5324	7785	Reconstructed by integrating previous versions, with emphasis on mass and charge balance
HMR1.0	1512	3397	4144	Reconstructed based on RECON1, EHMN, HumanCyc and KEGG
HMR2.0	3765	3160	8181	Reconstructed based on HMR1, with additional emphasis on lipid metabolism by integrating iAdipocytes1809, KEGG, Lipidomics Gateway
Recon3D	2248	5835	10600	Reconstructed based on RECON2 and includes mapping to 3D structure of proteins through PDB ids
Human1	3625	10138 (4164)	13417	Integrated and extensively curated the most recent human metabolic models to construct a consensus GEM, Human1

- Several options available ( all derived from RECON1)
- Human1 is most recent version

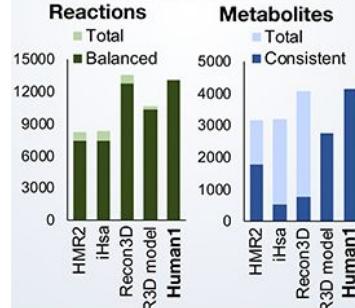
Modified from Swainston, N., et al (2016).  
Metabolomics, 12(7), 109.

### Full stoichiometric consistency

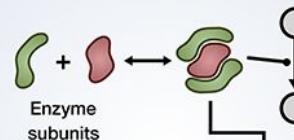
- Balance reactions by mass & charge
- Conserve mass of all metabolites



**memote**

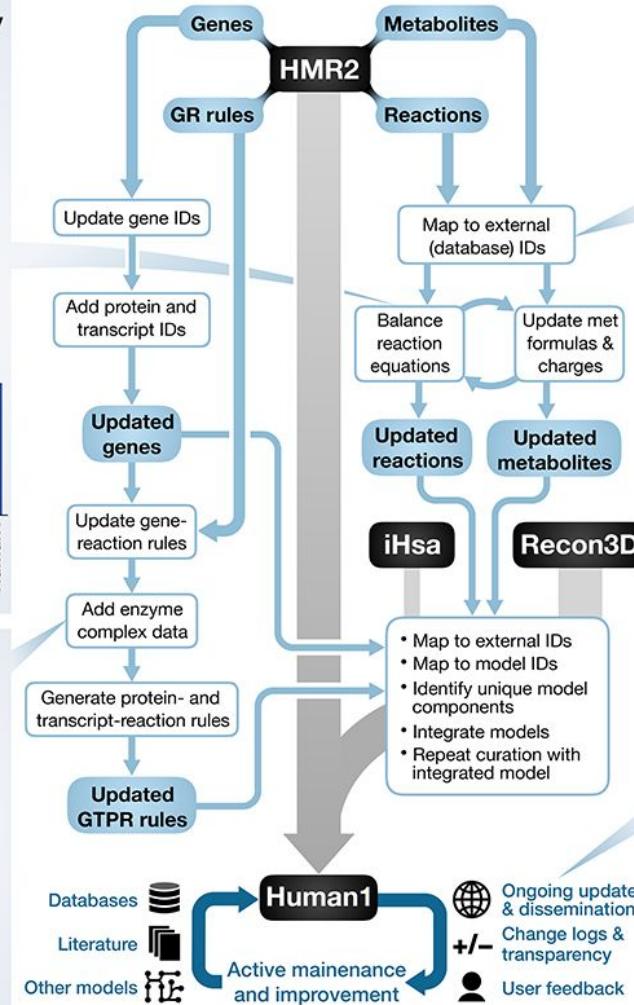


### Enzyme complex information



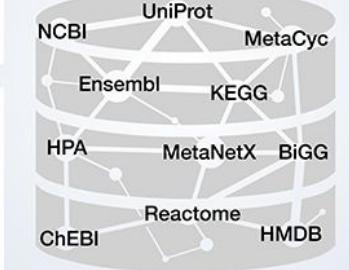
CORUM  
enzyme  
complex  
database

$$\text{Proteins} \begin{pmatrix} 0 & 1 & 0 & 1 \\ 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \end{pmatrix}$$

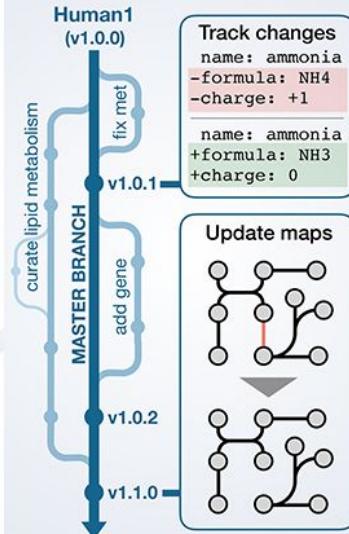


### Mapping to database IDs

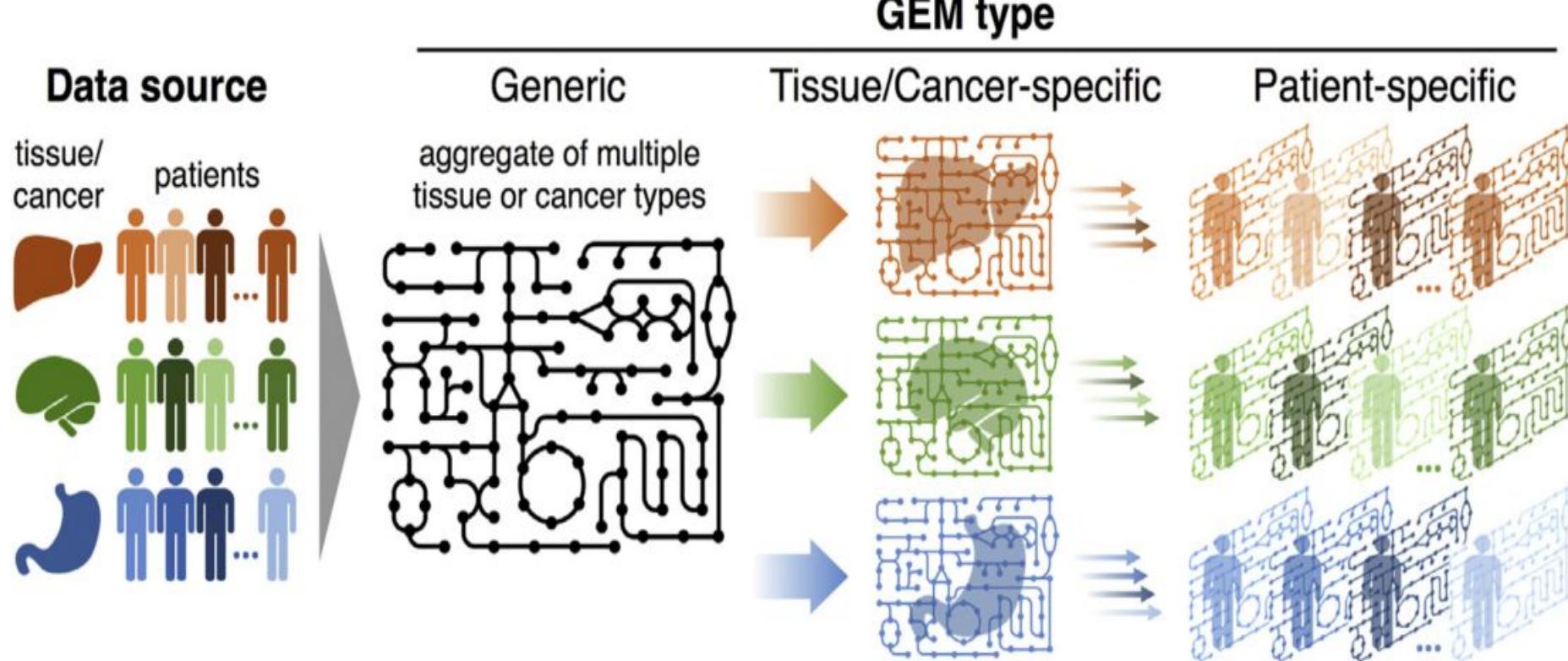
Reactions Genes Metabolites



### Ongoing transparent curation

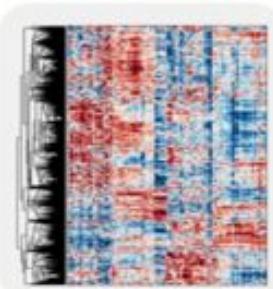


# Metabolic modeling in humans

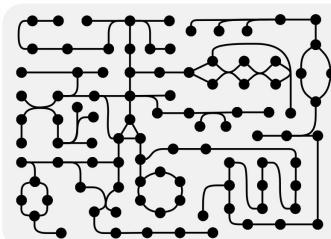


# Context-specific model reconstruction

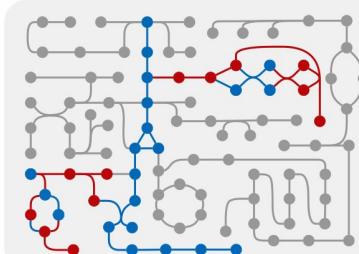
Omic context



References model



Model Extraction Method



Validation

Simulations

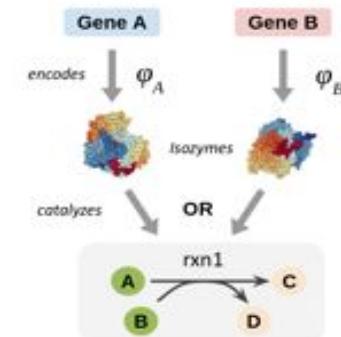
Context-specific model

1-Gene Thresholding

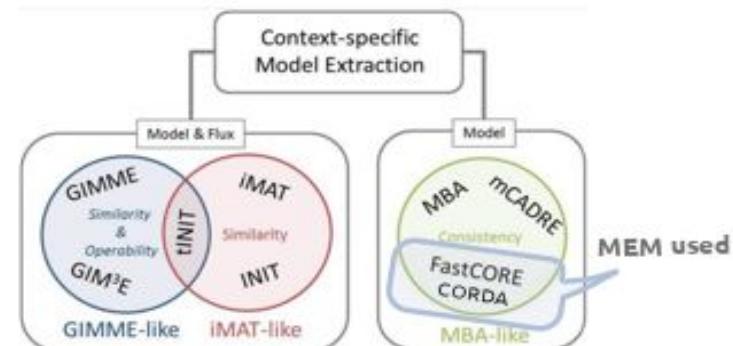


$$\varphi_D \begin{cases} \varphi_D < T \rightarrow 0 \text{ (inactive)} \\ \varphi_D \geq T \rightarrow 1 \text{ (active)} \end{cases}$$

2- Reaction Mapping



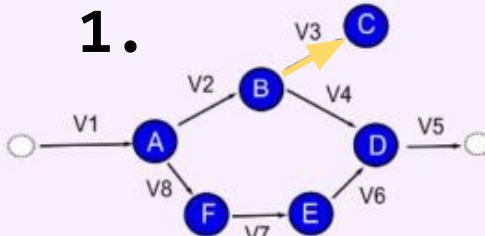
3- Model Extraction



# Model Extraction Methods

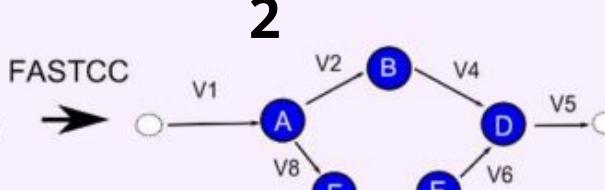
Find a flux-consistent sub network based on *omic*-derived constraint

1.



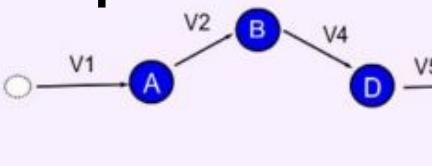
global model

2



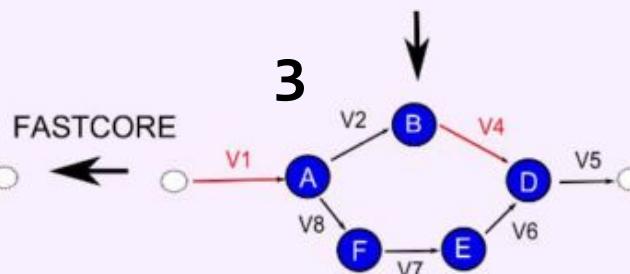
consistent global model N

4



context-specific  
reconstruction

3



model N and core set C  
(shown in red)

Inconsistent (blocked) reactions  
(due to *dead end* C)

High confidence reactions  
(Experimental evidence: *omics*)

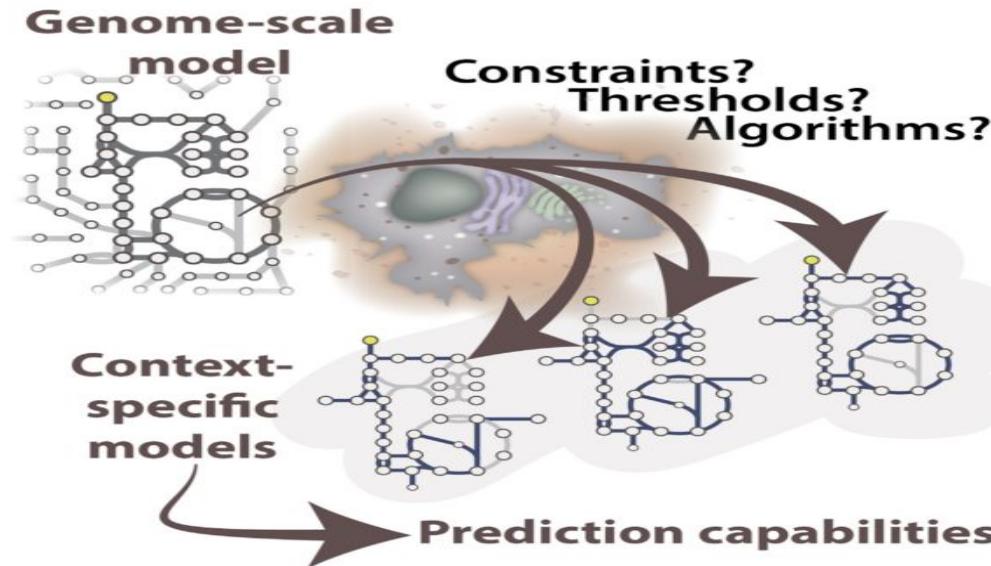
Med/Low confidence reactions  
(Not supported by *omic* data)

## Model extraction steps

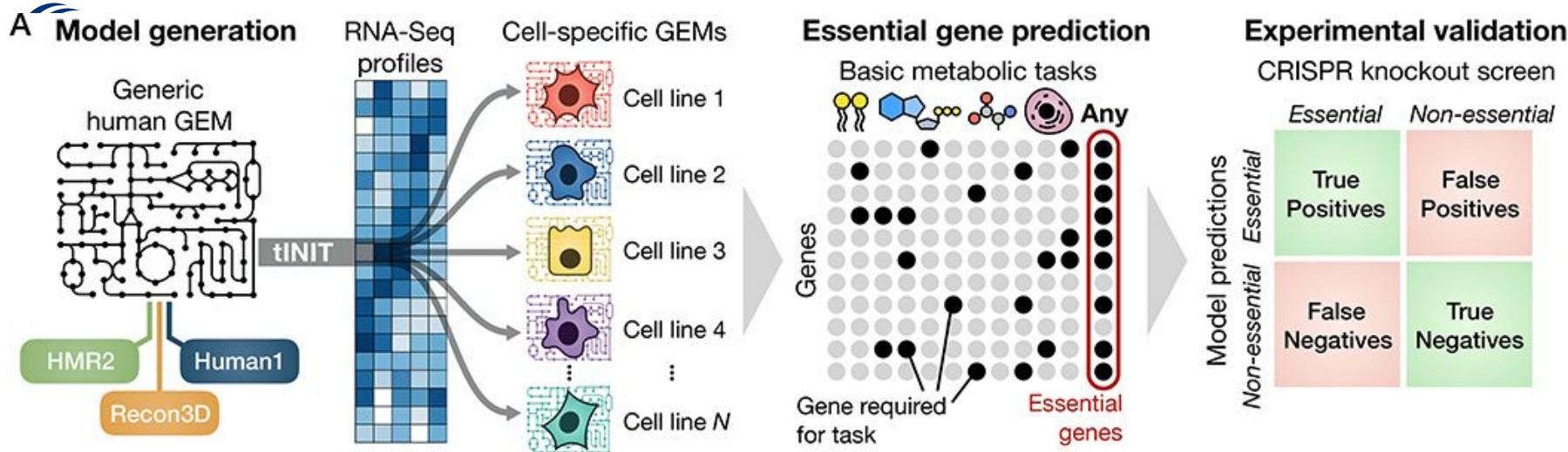
1. Select a global model (Recon2.2)
2. Remove model inconsistencies
3. Weight reactions presence confidence
4. Extract context-specific metabolic network

# Algorithm and parameters

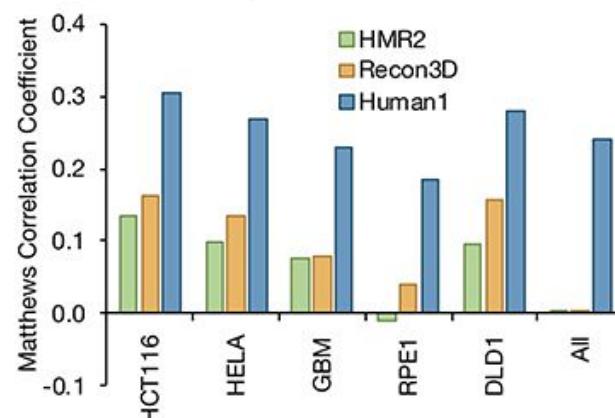
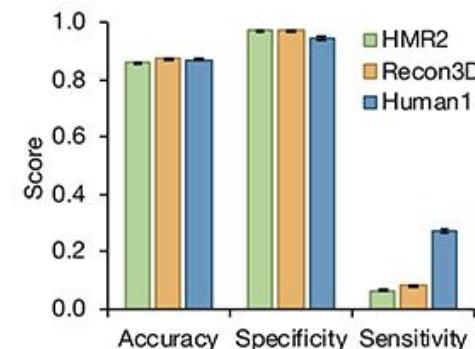
CSM depend on key decisions on methodology and data processing



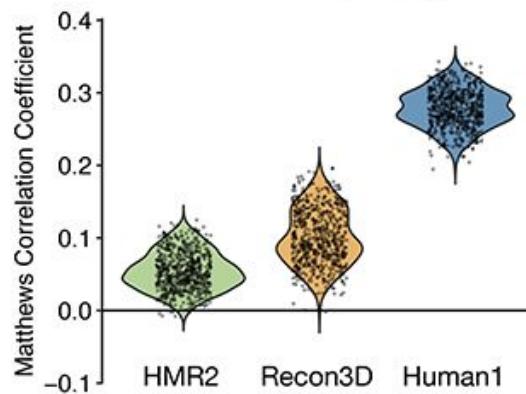
- No strong evidence that one MEM universally gives the most accurate models
- Each method has different underlying assumptions that affect the resulting model
- **Gene thresholding** seems to be the most determinant decision



**B Validation results: Hart et al. dataset (5 cell lines)**



**C Validation results: DepMap (621 cell lines)**



# Modeling Human Metabolism

## Summary

- Modeling is iterative → models are updated
- Integrating expression → context-specific models
- Used Metabolic tasks and omic-context to guide the exploration of a model
- Prediction and validation: *In-silico* gene knockout experiments

## Applications in precision medicine

- Synthetic lethals
- Onco-metabolites / Anti-metabolites
- Prediction of biomarkers
- Side effects (toxicity)

## Technical challenges

- Define growth condition
- Define a meaningful cell objective
- Kinetic and regulatory constraints
- Integrate omic data: transcriptomics, metabolomics, proteomics, fluxomic

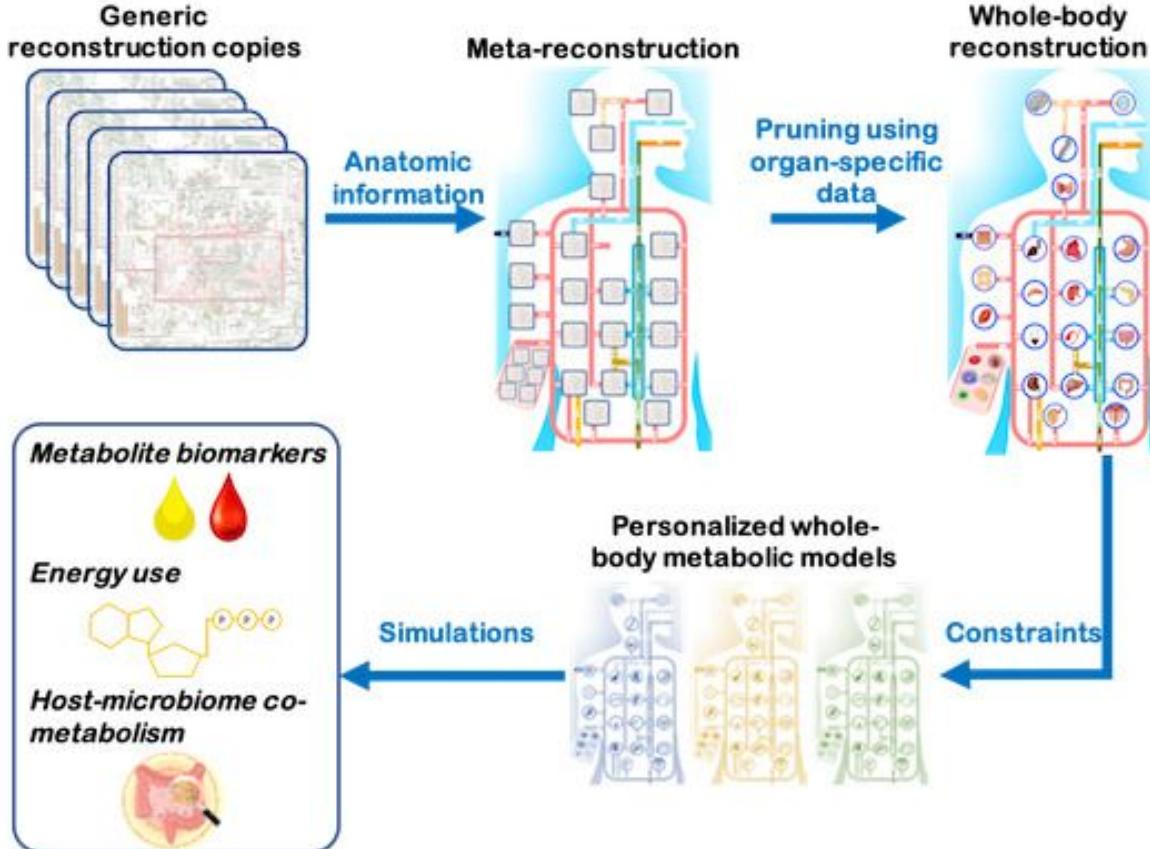


**Barcelona  
Supercomputing  
Center**

Centro Nacional de Supercomputación

# Perspectives

# The future is here: whole body models



Metabolic network reconstruction approach that used **organ-specific information** from literature and omics data to generate **two sex-specific whole-body metabolic (WBM)** reconstructions. These reconstructions capture the metabolism of **26 organs and six blood cell types**. Each WBM reconstruction represents whole-body organ-resolved metabolism with over **80,000 biochemical reactions** in an anatomically and physiologically consistent manner.

# The future is here: microbial community modeling

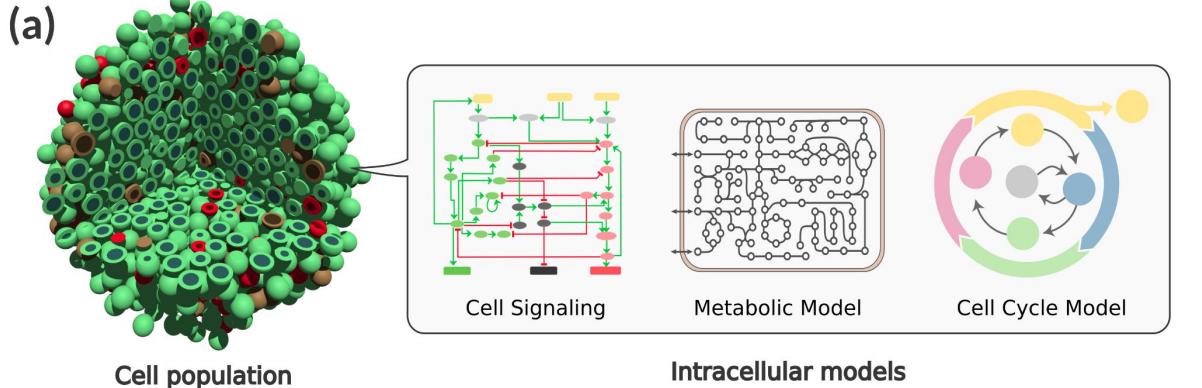
Generation of genome-scale metabolic reconstructions for 773 members of the human gut microbiota

Stefanía Magnúsdóttir<sup>1,2</sup>, Almut Heinken<sup>1,2</sup>, Laura Kutt<sup>1</sup>, Dmitry A Ravcheev<sup>1</sup>, Eugen Bauer<sup>1</sup>, Alberto Noronha<sup>1</sup>, Kacy Greenhalgh<sup>1</sup>, Christian Jäger<sup>1</sup>, Joanna Baginska<sup>1</sup>, Paul Wilmes<sup>1</sup>, Ronan M T Fleming<sup>1</sup> & Ines Thiele<sup>1</sup>

An extended reconstruction of human gut microbiota metabolism of dietary compounds

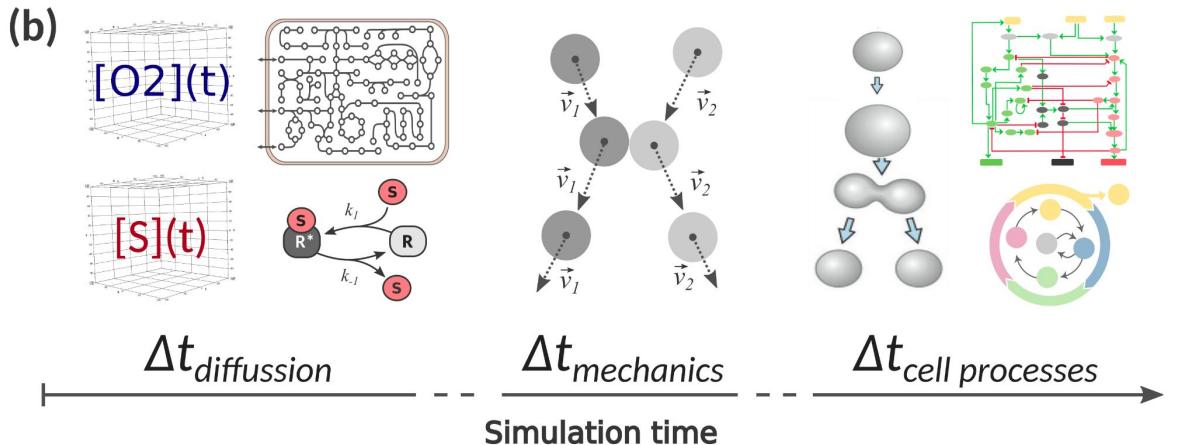
Telmo Blasco  <sup>1,2</sup>, Sergio Pérez-Burillo  <sup>3,7</sup>, Francesco Balzerani<sup>1,2,7</sup>, Daniel Hinojosa-Nogueira  <sup>3,7</sup>, Alberto Lerma-Aguilera  <sup>4,5,7</sup>, Silvia Pastoriza  <sup>3</sup>, Xabier Cendoya  <sup>1,2</sup>, Ángel Rubio<sup>1,2</sup>, María José Gosálbez<sup>4,5</sup>, Nuria Jiménez-Hernández<sup>4,5</sup>, M. Pilar Francino  <sup>4,5</sup>✉, Iñigo Apaolaza  <sup>1,2</sup>✉, José Ángel Rufián-Henares  <sup>3,6</sup>✉ & Francisco J. Planes  <sup>1,2</sup>✉

# What are multi-scale modes?



## Hybrid modeling

- Integration of *omic* data
- Populations heterogeneity
- Models as components



## Multi-scale simulations

- Different time-scales
- Explicit microenvironment
- Process coupling

# Metabolic modeling with COBREXA

*Hands-on Tutorial*



 **cobrapy-tutorial** Public

[Pin](#) [Unwatch 1](#)

---

[main](#) ▼ [1 branch](#) [0 tags](#) [Go to file](#) [Add file](#) [Code](#) ▼

---

 migp11	Update README.md	46b8e38 on Nov 23, 2022	35 commits
 data	adding escher dependency	3 years ago	
 img	all except slides	3 years ago	
 out	all except slides	3 years ago	
 .gitignore	Initial commit	3 years ago	
 01_Intro_cobrapy.ipynb	fixing small things	last year	
 02_Flux_Balance_Analysis.ipynb	fixing small things	last year	
 03_In-silico_gene_knockouts.ipynb	fixing small things	last year	
 LICENSE	Initial commit	3 years ago	
 README.md	Update README.md	last year	
 requirements.in	fixing small things	last year	
 requirements.txt	Update requirements.txt	last year	
 runtime.txt	Updating requirments	last year	



\_intro Last Checkpoint: 05/10/2023 (autosaved)

### Kernel not found

Could not find a kernel matching Julia 1.9.0-beta3. Please select a kernel:

Julia (8 threads) 1.9.0

Continue Without Kernel

Set Kernel

stallation. In

```
port Pkg  
g.add("COBREXA")
```

```
ing COBREXA
```

```
Updating registry at `~/.julia/registries/General.toml`  
Resolving package versions...
```

### Kernel not found

Could not find a kernel matching Julia 1.9.0-beta3. Please select a kernel:

Julia 1.6.7

Continue Without Kernel

Set Kernel

```
_coli_core.json"
```



Installation. In Julia REPL this can be done easier using the packaging mode (typing ]add COBREXA )

```
In [ ]: import Pkg  
Pkg.add("COBREXA")  
  
using COBREXA
```

Let's first get a simple model to have a look at

```
In [ ]: import Downloads  
Downloads.download(  
    "http://bigg.ucsd.edu/static/models/e_coli_core.json",  
    "e_coli_core.json",  
)
```

Load the model

```
In [ ]: ecoli = load_model("e_coli_core.json");
```

have a look at what the model contains. First, metabolites

```
In [ ]: metabolites(ecoli)
```

reactions



# ESCHER

*Build, share, and embed visualizations of metabolic pathways*

## Filter by organism

All

Map

Model (Optional)

Tool

Central metabolism (iJO1366)

iJO1366

Builder

## Options

- Scroll to zoom (instead of scroll to pan)
- Never ask before reloading

Load map

<https://escher.github.io/#/>

