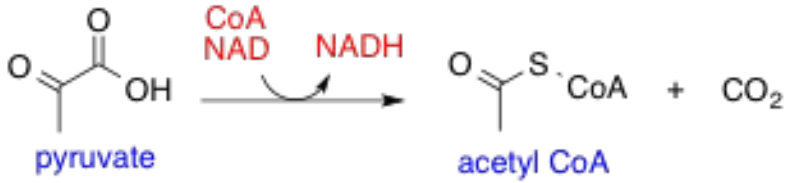
Week 14: Whole-cell models; Digital evolution

- Genome-scale metabolic models
  - Reconstruction
  - Flux balance analysis
- Artificial life

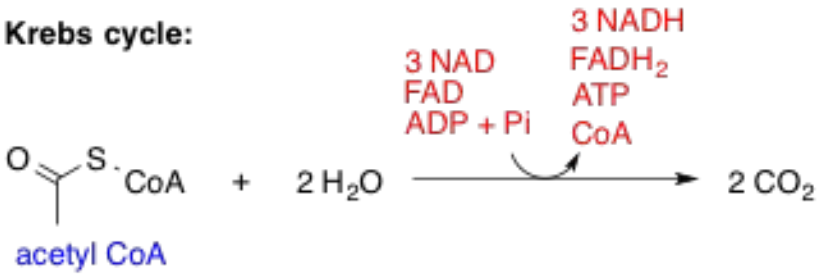
Metabolic networks

Glycolysis:

Oxidative decarboxylation:

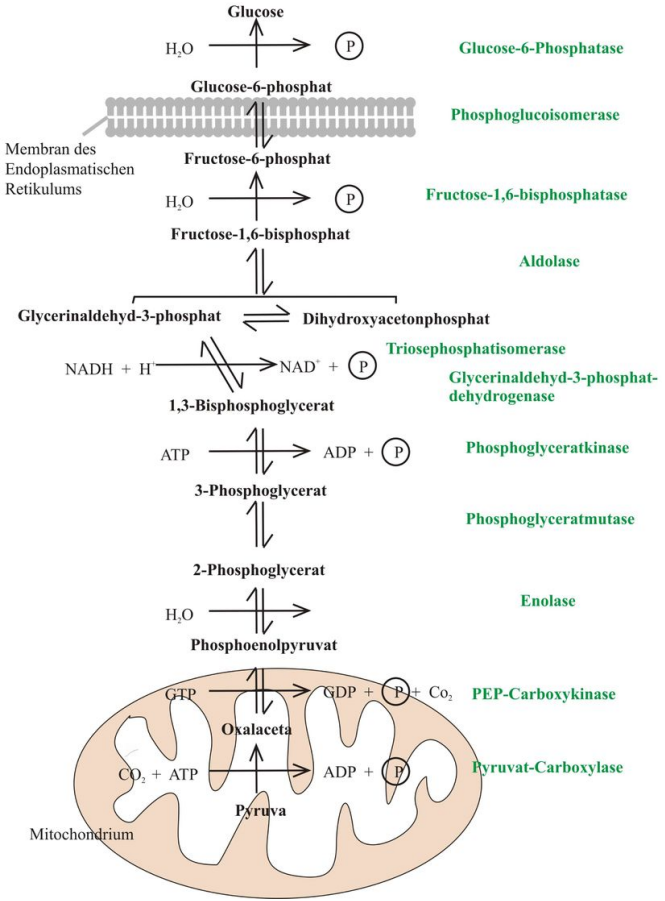


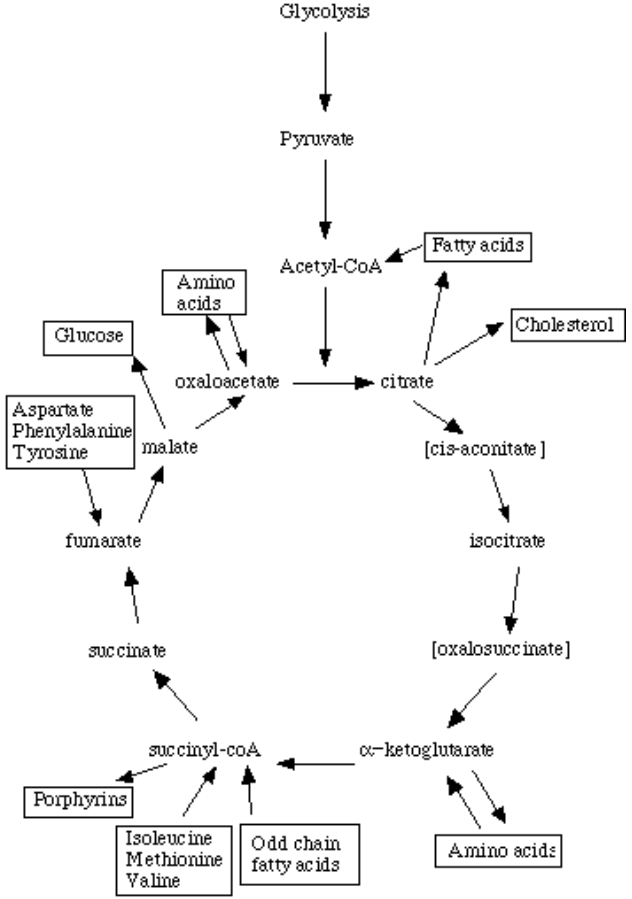
Krebs cycle:



Wikipedia

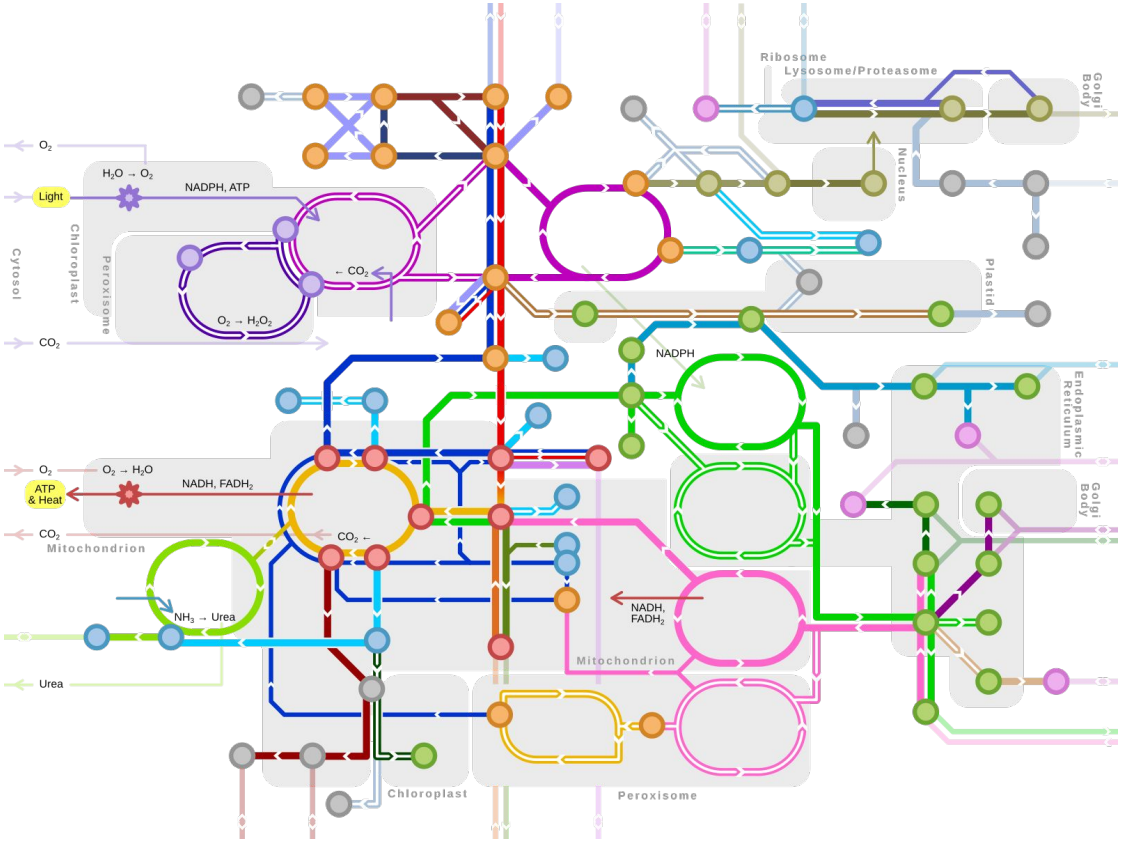
Metabolic networks





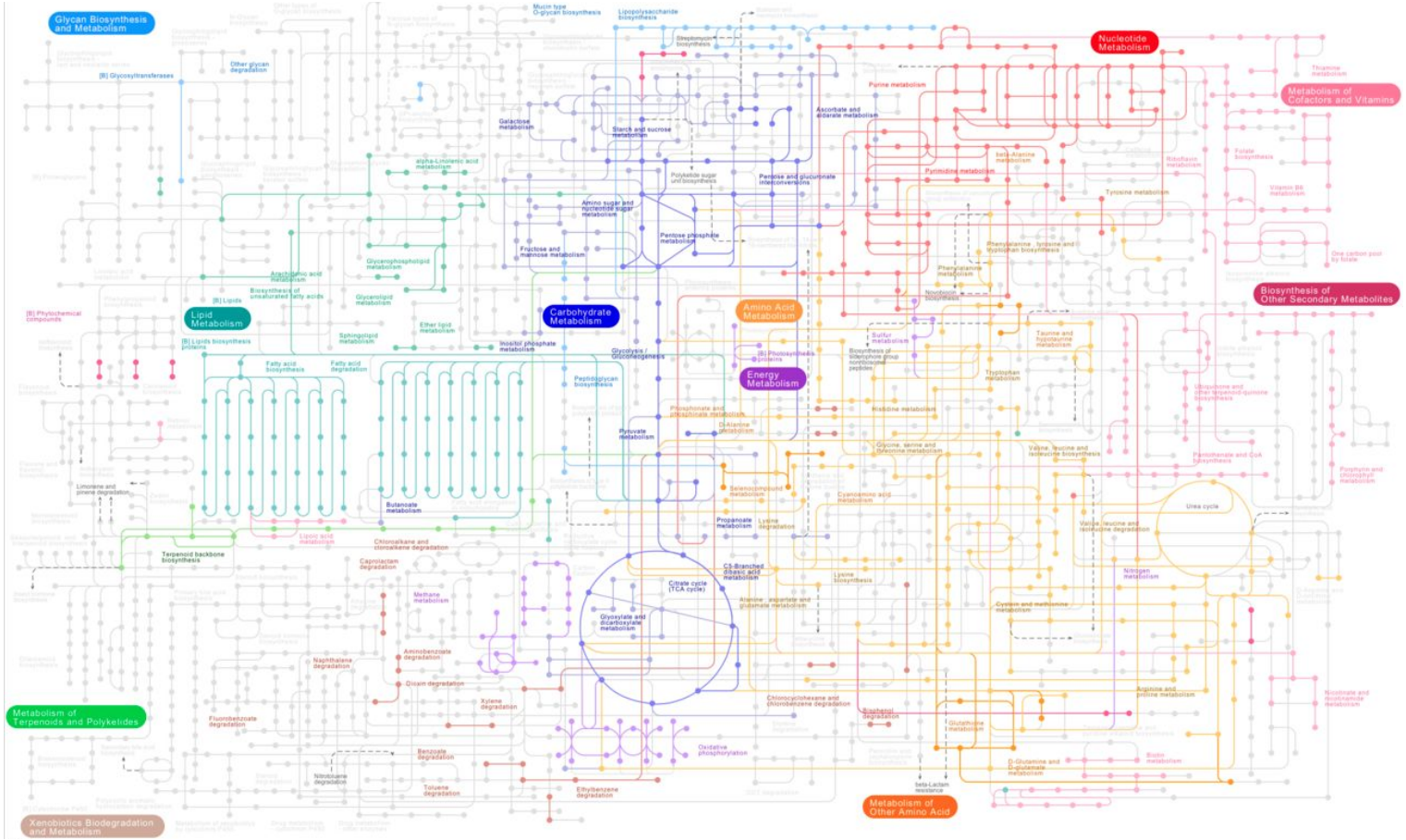
Wikipedia

Metabolic networks



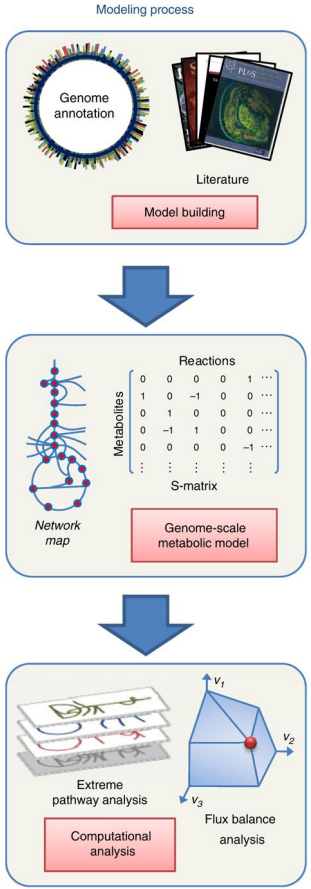
Wikipedia

Metabolic networks



KEGG

Genome-scale metabolic network reconstruction & model



Genome-scale metabolic network **reconstruction**:

- A collection of biochemical transformation derived from the genome annotation and the literature of a particular organism.
- Formed based on an organism-specific knowledge base.
- A network reconstruction is unique to an organism.

Genome-scale metabolic network **model**:

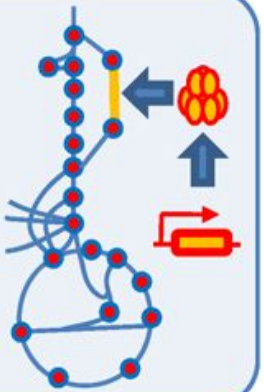
- Derived from a *reconstruction* by converting it into a mathematical form (i.e., an in silico model) and by assessing its phenotypic properties computationally.

Oberhardt (2009) Mol. Sys. biol.
Thiele & Palsson (2010) Nat. Protoc.

Genome-scale metabolic network reconstruction & model

3. Directing hypothesis-driven discovery

A Metabolic GENRE aided in determining pathway usage and discovering a novel citramalate synthase gene in *G. sulfurreducens*. GENREs have also helped study the effects of transposons on downstream genes, and identify transc riptional timing patterns in *S. cerevisiae*.



2. Guidance of metabolic engineering

Metabolic GENREs guided efforts to engineer malate and succinate producing strains of *S. cerevisiae* and *M. succiniciproducens*. GENREs have also helped determine ways to increase the respiration rate of *G. sulfurreducens* and scale-up vaccine production against *N. meningitides*.

Oberhardt (2009) Mol. Sys. biol.
Thiele & Palsson (2010) Nat. Protoc.

Constructing a genome-scale metabolic model

Draft construction

Genome databases

Comprehensive Microbial Resource (CMR)	http://cmr.jcvi.org/cgi-bin/CMR/CmrHomePage.cgi
Genomes OnLine Database (GOLD)	http://www.genomesonline.org/
TIGR	http://www.tigr.org/db.shtml
NCBI Entrez Gene	http://www.ncbi.nlm.nih.gov/sites/entrez
SEED database[32]	http://theseed.uchicago.edu/FIG/index.cgi

Biochemical databases

KEGG[41]	http://www.genome.jp/kegg/
BRENDA[42]	http://www.brenda-enzymes.info/
Transport DB[89]	http://www.membranetransport.org/
PubChem[86]	http://pubchem.ncbi.nlm.nih.gov/
Transport Classification Database (TCDB)	http://www.tcdb.org/
pK_a Plugin	http://www.chemaxon.com/product/pka.html
pK_a DB	http://www.acdlabs.com/products/phys_chem_lab/pka/

1. Draft reconstruction

1| Obtain genome annotation.
2| Identify candidate metabolic functions.
3| Obtain candidate metabolic reactions.
4| Assemble draft reconstruction.
5| Collect experimental data.

Organism-specific databases

Ecocyc[43]	http://ecocyc.org/
PyloriGene[37]	http://genolist.pasteur.fr/PyloriGene
Gene Cards	http://www.genecards.org/

Protein localization databases

PSORT[47]	http://www.psort.org/psortb/
PA-SUB[48]	http://www.cs.ualberta.ca/~bioinfo/PA/Sub/

Bio-numbers

CyberCell Database (CCDB)[88]	http://redpoll.pharmacy.ualberta.ca/CCDB/cgi-bin/STAT_NEW.cgi
B10NUMB3R5	http://bionumbers.hms.harvard.edu/

Thiele & Palsson (2010) Nat. Protoc.

Constructing a genome-scale metabolic model

Refinement of draft construction

2. Refinement of reconstruction

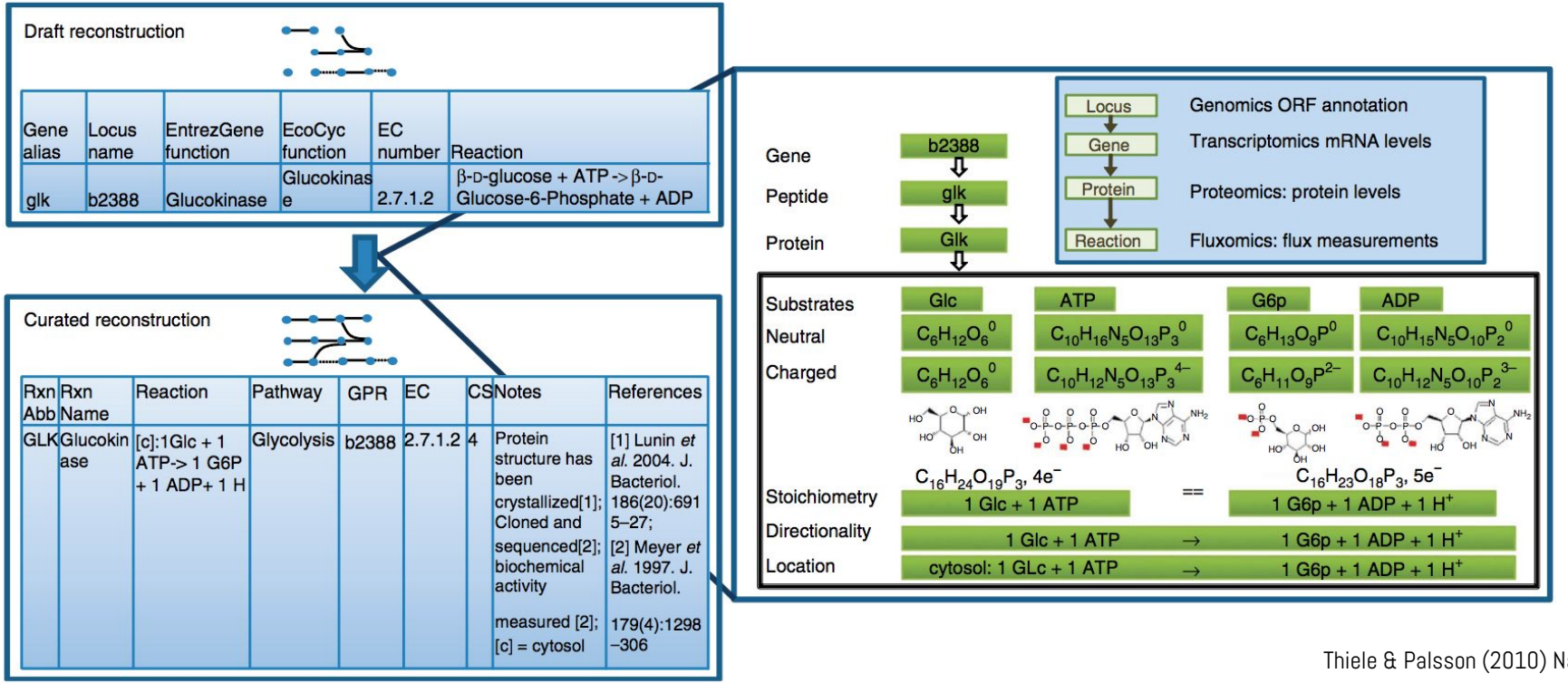
6| Determine and verify substrate and cofactor usage.
7| Obtain neutral formula for each metabolite.
8| Determine the charged formula.
9| Calculate reaction stoichiometry.
10| Determine reaction directionality.
11| Add information for gene and reaction localization.
12| Add subsystems information.
13| Verify gene–protein-reaction association.
14| Add metabolite identifier.
15| Determine and add confidence score.
16| Add references and notes.
17| Flag information from other organisms.
18| Repeat Steps 6 to 17 for all genes.
19| Add spontaneous reactions to the reconstruction.
20| Add extracellular and periplasmic transport reactions.
21| Add exchange reactions.
22| Add intracellular transport reactions.
23| Draw metabolic map (optional).
24–32| Determine biomass composition.
33| Add biomass reaction.
34| Add ATP-maintenance reaction (ATPM).
35| Add demand reactions.
36| Add sink reactions.
37| Determine growth medium requirements.

Thiele & Palsson (2010) Nat. Protoc.

Constructing a genome-scale metabolic model

Refinement of draft construction

Mass & charge balancing; Filling-in H^+ & water; adjusting metabolites to a particular pH



Thiele & Palsson (2010) Nat. Protoc.

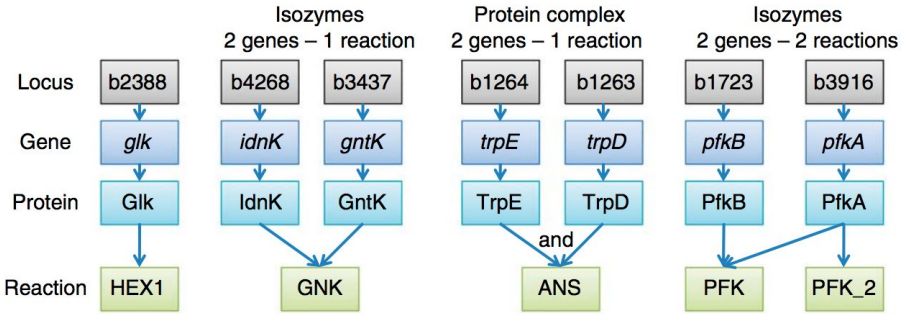
Constructing a genome-scale metabolic model

Refinement of draft construction

Subcellular localization

Compartment	Commonly used symbol[#]	Achaea	Bacteria	Eukaryotic pathogens[a]	Fungi[b]	Photosynthetic eukarya[c]	Baker's yeast	Human
Extracellular space	[e]	■	■		■	■	■	■
Periplasm	[p]		■					
Cytoplasm	[c]	■	■	■	■	■	■	■
Nucleus	[n]			■			■	■
Mitochondrion	[m]			■	■		■	■
Chloroplast	[h]					■		
Lysosome[*]	[l]							
Vacuole	[v]						■	■
Golgi apparatus	[g]						■	■
Endoplasmatic	[r]			■			■	■

Gene–protein-reaction associations



Reaction abbreviation	Reaction name	E. C.number	GPR
HEX1	Hexokinase (D-glucose:ATP)	2.7.1.1	(b2388)
GNK	Gluconokinase	2.7.1.12	(b3437) or (b4268)
ANS	Anthranilate synthase	4.1.3.27	(b1264) and (b1263)
PFK	Phosphofructokinase	2.7.1.11	(b1723) or (b3916)
PFK_2	Phosphofructokinase (2)	2.7.1.11	(b3916)

Thiele & Palsson (2010) Nat. Protoc.

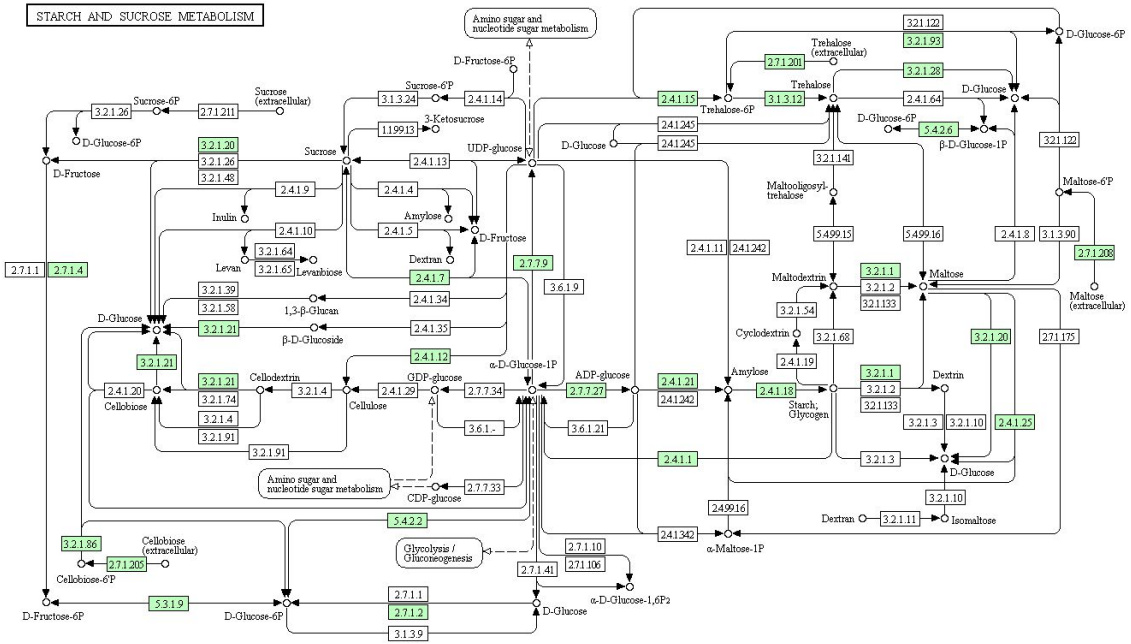
Constructing a genome-scale metabolic model

Refinement of draft construction

Chemical composition of a cell

Cellular component	Cellular content %(wt/wt)
Protein	55
RNA	20.5
DNA	3.1
Lipids	9.1
Lipopolysaccharides	3.4
Peptidoglycan	2.5
Glycogen	2.5
Polyamines	0.4
Other	3.5
Total	100.00

Identification of missing functions



Thiele & Palsson (2010) Nat. Protoc.

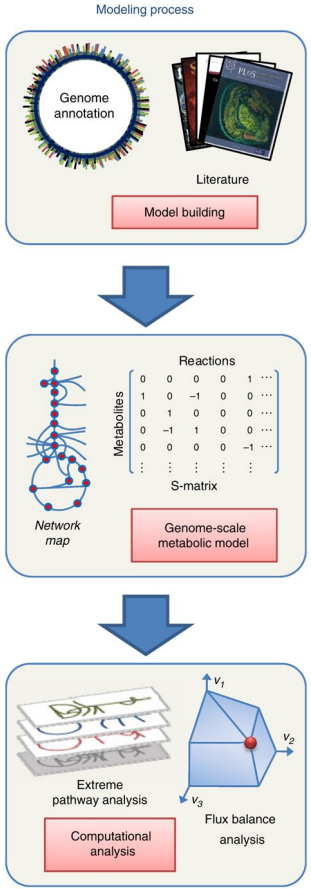
Constructing a genome-scale metabolic model

Refinement of draft construction

Evidence type	Confidence score	Examples
Biochemical data	4	Direct evidence for gene product function and biochemical reaction: protein purification, biochemical assays, experimentally solved protein structures and comparative gene-expression studies (e.g., Chhabra *et al.*[95])
Genetic data	3	Direct and indirect evidence for gene function: knockout characterization, knock-in characterization and overexpression
Physiological data	2	Indirect evidence for biochemical reactions based on physiological data: secretion products or defined medium components serve as evidence for transport and metabolic reactions
Sequence data	2	Evidence for gene function: genome annotation and SEED annotation[32]
Modeling data	1	No evidence is available, but reaction is required for modeling. The included function is a hypothesis and needs experimental verification. The reaction mechanism may be different from the included reaction(s)
Not evaluated	0	

Thiele & Palsson (2010) Nat. Protoc.

Genome-scale metabolic network reconstruction & model



Genome-scale metabolic network **reconstruction**:

- A collection of biochemical transformation derived from the genome annotation and the literature of a particular organism.
- Formed based on an organism-specific knowledge base.
- A network reconstruction is unique to an organism.

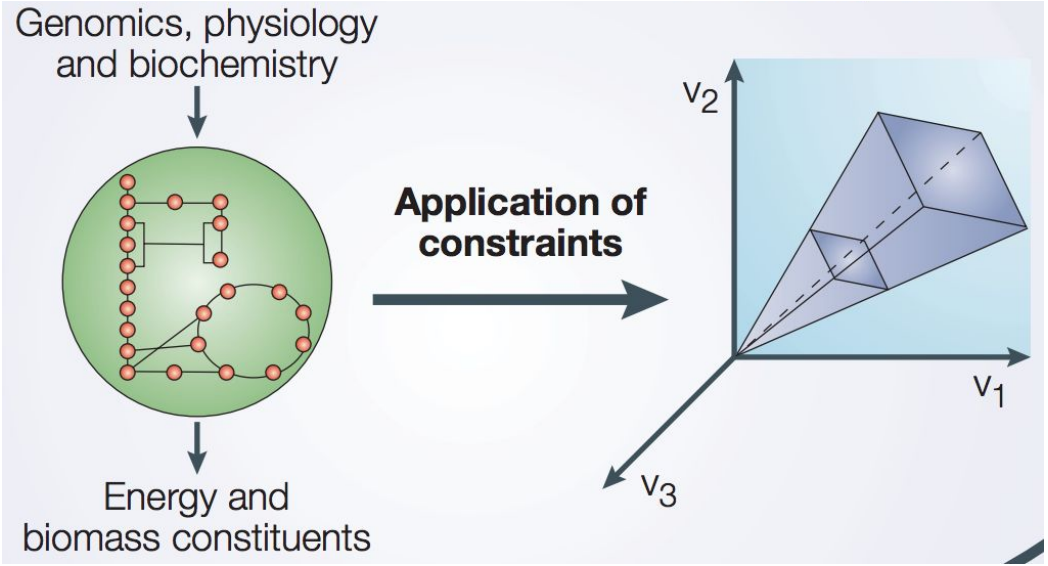
Genome-scale metabolic network **model**:

- Derived from a *reconstruction* by converting it into a mathematical form (i.e., an in silico model) and by assessing its phenotypic properties computationally.

Oberhardt (2009) Mol. Sys. biol.
Thiele & Palsson (2010) Nat. Protoc.

Constructing a genome-scale metabolic model

Network construction **Allowable solution space**



Price (2004) Nat. Rev. Micro.

Flux balance analysis (FBA)

FBA: metabolic network ⟶ linear programming optimization problem.

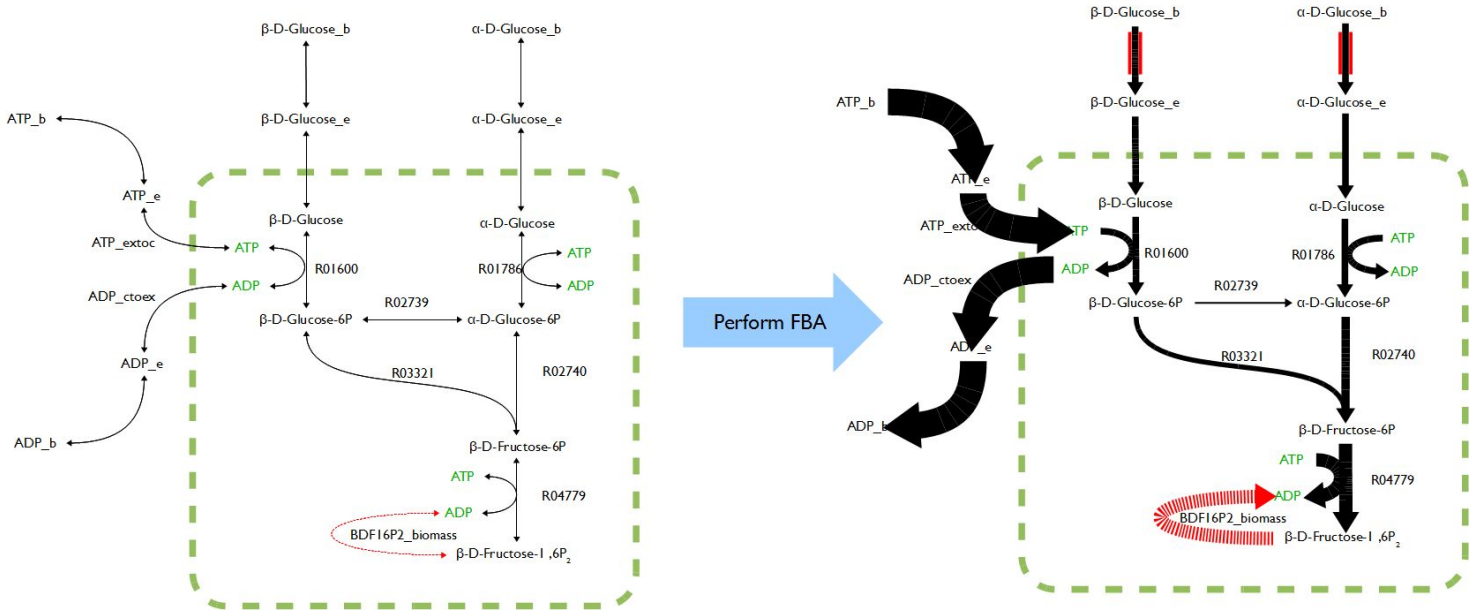
The main constraints in FBA: steady-state mass conservation of metabolites.

- Relies on balancing metabolic fluxes
- Is based on the fundamental law of mass conservation
- Is performed under steady-state conditions (an example of constraint…)
- Requires information only about:
  a. the stoichiometry of metabolic pathways,
  b. metabolic demands, and
  c. a few strain specific parameters
- Does NOT require enzymatic kinetic data

Flux balance analysis

The results of FBA on a metabolic network of the top six reactions of glycolysis.

- The predicted flux through each reaction is proportional to the width of the line.
- **Red springy arrow**: Objective function; **Red bars**: Constraints on **α**-D-glucose and **β**-D-glucose import.

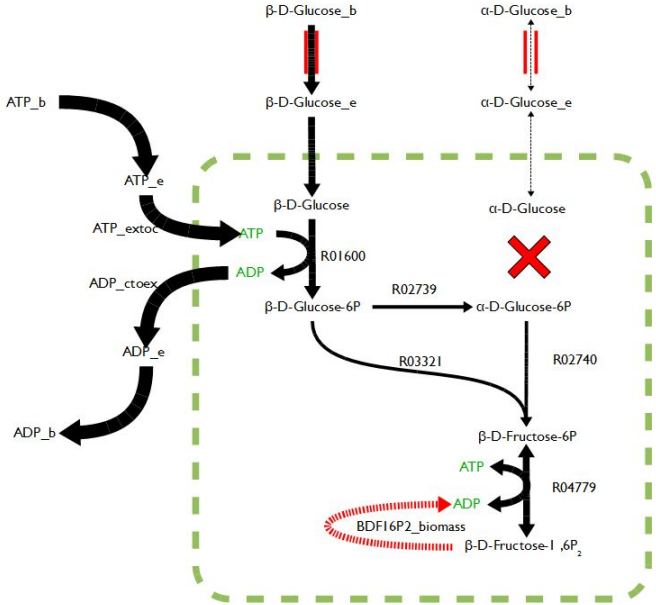


Wikipedia

Flux balance analysis

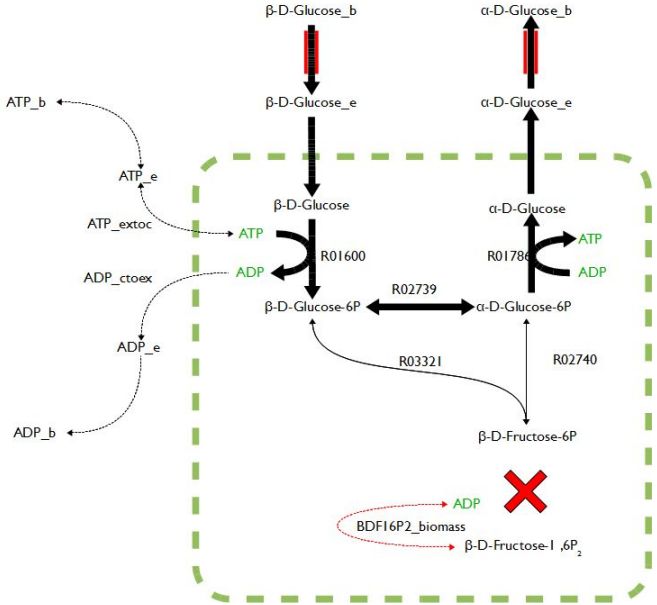
Non-lethal gene deletion in a metabolic network.

- Flux through the objective function is halved but is still present.



Lethal gene deletion in a metabolic network.

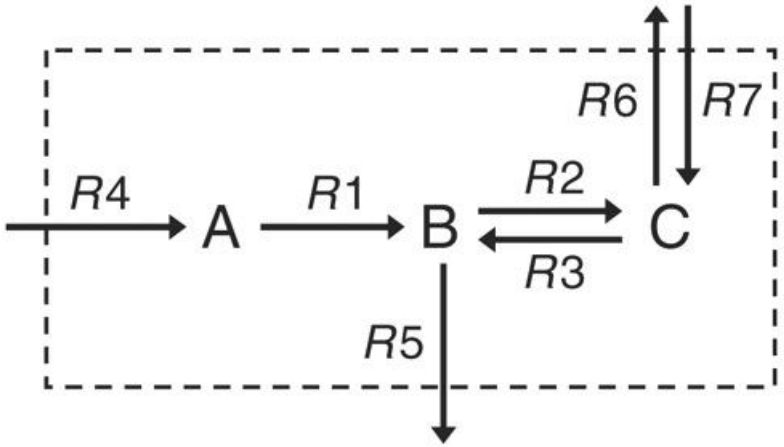
- No flux through the objective function → pathway is no longer functional.



Wikipedia

Flux balance analysis

1. Reaction network formalism



Chemical reactions

Internal	Exchange
*R*1: −1 A → 1 B	*R*4: 1 A
*R*2: −1 B → 1 C	*R*5: −1 B
*R*3: −1 C → 1 B	*R*6: −1 C
	*R*7: 1 C

S =

	*R*1	*R*2	*R*3	*R*4	*R*5	*R*6	*R*7
A	−1	0	0	1	0	0	0
B	1	−1	1	0	−1	0	0
C	0	1	−1	0	0	−1	1

Lee (2006) Brief. Bioinfo.

Flux balance analysis

1. Reaction network formalism

S =

	R1	R2	R3	R4	R5	R6	R7
A	–1	0	0	1	0	0	0
B	1	–1	1	0	–1	0	0
C	0	1	–1	0	0	–1	1

$$\begin{array}{c} \\ A \\ B \\ C \end{array} \begin{array}{c} \begin{array}{ccccccc} R1 & R2 & R3 & R4 & R5 & R6 & R7 \end{array} \\ \begin{bmatrix} -1 & 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & -1 & 1 & 0 & -1 & 0 & 0 \\ 0 & 1 & -1 & 0 & 0 & -1 & 1 \end{bmatrix} \end{array} \begin{bmatrix} v_1 \\ \vdots \\ v_7 \end{bmatrix}$$

Lee (2006) Brief. Bioinfo.

Flux balance analysis

2. FBA formulation

Dynamic mass balance

$$\frac{dC}{dt} = \mathbf{Sv}$$

C : Concentration
t : Time
$\mathbf{S}$: Stoichiometric matrix
$\boldsymbol{v}$: Flux vector

Steady-state assumption $\quad \mathbf{Sv = 0}$

LP formulation

Objective: max $Z = \mathbf{c} \cdot \mathbf{v}$

Constraints:

$$\begin{array}{c} \\ A \\ B \\ C \end{array} \begin{array}{c} \begin{array}{ccccccc} R1 & R2 & R3 & R4 & R5 & R6 & R7 \end{array} \\ \begin{bmatrix} -1 & 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & -1 & 1 & 0 & -1 & 0 & 0 \\ 0 & 1 & -1 & 0 & 0 & -1 & 1 \end{bmatrix} \end{array} \begin{bmatrix} v_1 \\ \vdots \\ v_7 \end{bmatrix} = \mathbf{0} \qquad 0 \le v_1, \ldots, v_7 \le 10$$

Lee (2006) Brief. Bioinfo.