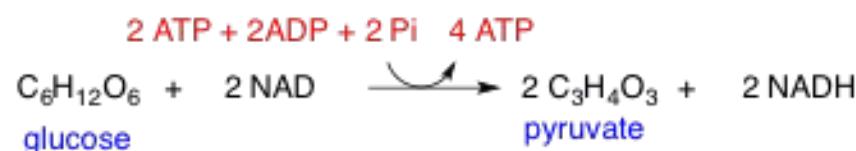


Whole-cell models; Digital evolution

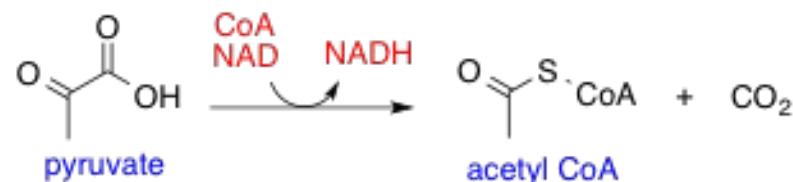
- Flux balance analysis
- Artificial life

Metabolic networks

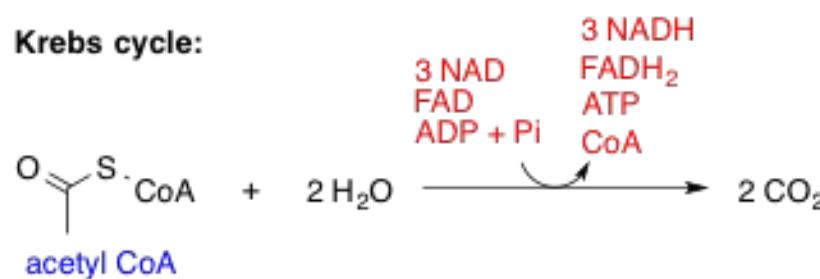
Glycolysis:



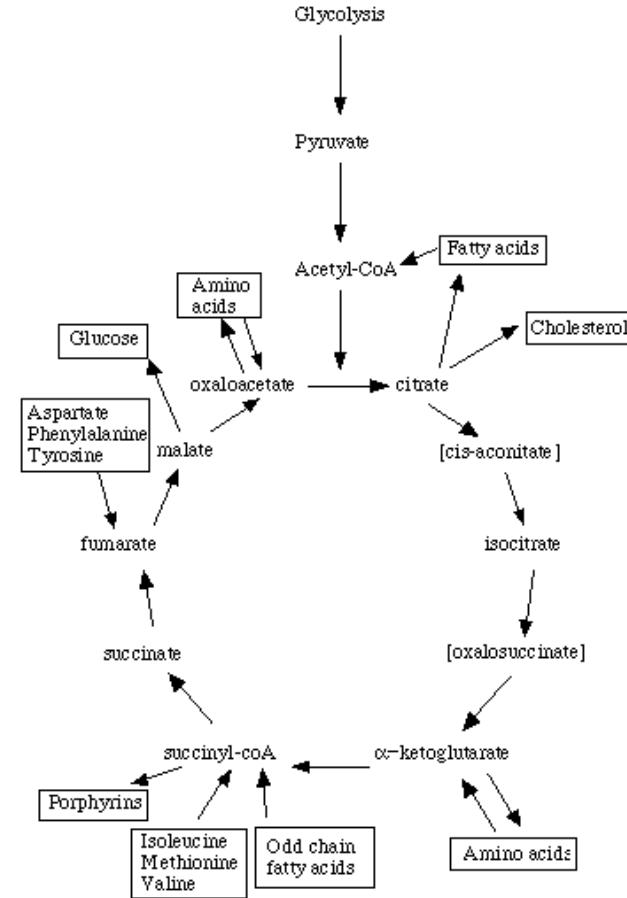
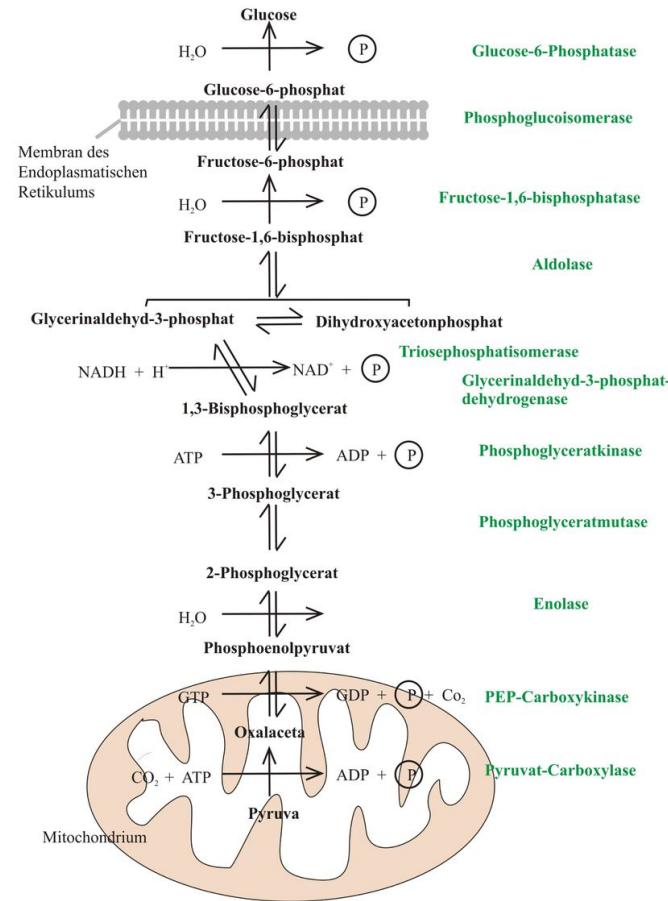
Oxidative decarboxylation:



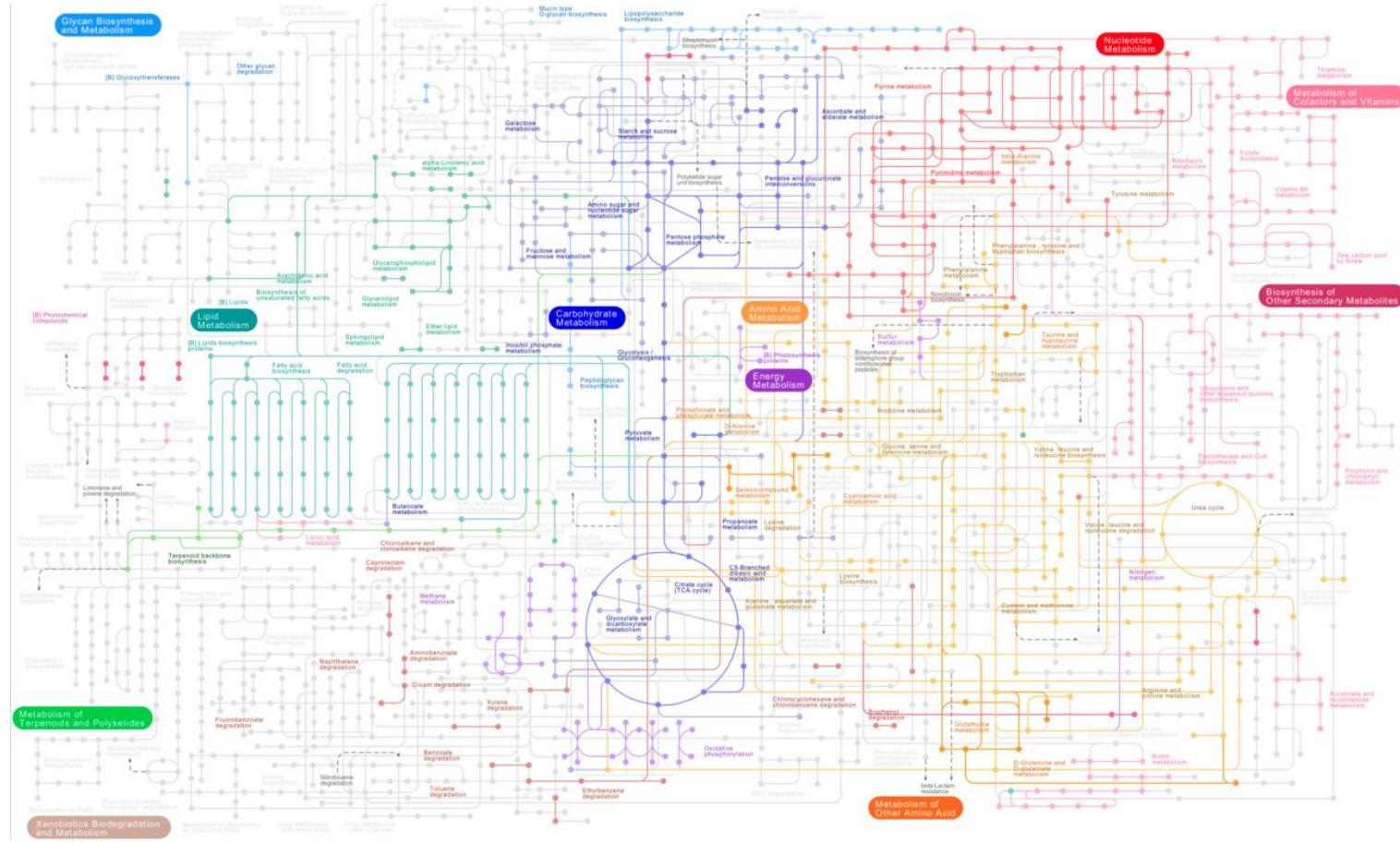
Krebs cycle:



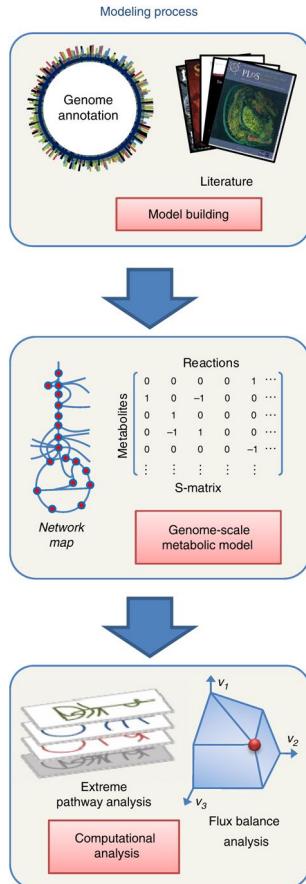
Metabolic networks



Metabolic networks



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.

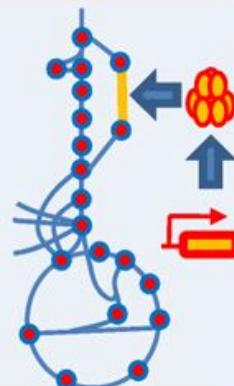
Thiele & Palsson (2010) Nat. Protoc.

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 transcriptional 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. meningitidis*.



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 database³² <http://theseed.uchicago.edu/FIG/index.cgi>

Biochemical databases

KEGG⁴¹ <http://www.genome.jp/kegg/>

BRENDA⁴² <http://www.brenda-enzymes.info/>

Transport DB⁸⁹ <http://www.membranetransport.org/>

PubChem⁸⁶ <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

Ecocyc⁴³ <http://ecocyc.org/>

PyloriGene³⁷ <http://genolist.pasteur.fr/PyloriGene>

Gene Cards <http://www.genecards.org/>

Protein localization databases

PSORT⁴⁷ <http://www.psorth.org/psortb/>

PA-SUB⁴⁸ <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/>

Constructing a genome-scale metabolic model

Refinement of draft construction

2. Refinement of reconstruction

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

Constructing a genome-scale metabolic model

Refinement of draft construction

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

Draft reconstruction

Gene alias	Locus name	EntrezGene function	EcoCyc function	EC number	Reaction
glk	b2388	Glucokinase	Glucokinase	2.7.1.2	$\beta\text{-D-glucose} + \text{ATP} \rightarrow \beta\text{-D-Glucose-6-Phosphate} + \text{ADP}$

Curated reconstruction

Rxn Abb	Rxn Name	Reaction	Pathway	GPR	EC	CS Notes	References	Location
GLK	Glucokinase	[c]:1Glc + 1 ATP-> 1 G6P + 1 ADP + 1 H ⁺	Glycolysis	b2388	2.7.1.2	4	Protein structure has been crystallized[1]; Cloned and sequenced[2]; biochemical activity measured [2]; [c] = cytosol	[1] Lunin <i>et al.</i> 2004. J. Bacteriol. 186(20):6915–27; [2] Meyer <i>et al.</i> 1997. J. Bacteriol. 179(4):1298–306

Genomics ORF annotation

Gene: b2388 → Gene → Protein → Reaction

Transcriptomics mRNA levels
Proteomics: protein levels
Fluxomics: flux measurements

Substrates

Neutral	Charged	Stoichiometry	Directionality
Glc: <chem>C6H12O6</chem>	ATP: <chem>C10H16N5O13P3^0</chem>	G6p: <chem>C6H13O9P0</chem>	ADP: <chem>C10H15N5O10P2^0</chem>
<chem>C6H12O6^-</chem>	<chem>C10H12N5O13P3^-</chem>	<chem>C6H11O9P2^-</chem>	<chem>C10H12N5O10P2^-</chem>

Chemical structures:

- Glc: O=C[C@H](O)[C@H](O)[C@H](O)[C@H](O)[C@H](O)C
- ATP: CC1=NC=NC1OP(=O)([O-])[O-]
- G6p: OC1=CC(O)=C(O[C@@H]1OP(=O)([O-])[O-])COP(=O)([O-])[O-]
- ADP: CC1=NC=NC1OP(=O)([O-])[O-]

Stoichiometry

$$\text{C}_{16}\text{H}_{24}\text{O}_{19}\text{P}_3, 4\text{e}^- \quad \text{1 Glc} + \text{1 ATP} \quad == \quad \text{C}_{16}\text{H}_{23}\text{O}_{18}\text{P}_3, 5\text{e}^- \quad \text{1 G6p} + \text{1 ADP} + \text{1 H}^+$$

Directionality

$$\text{1 Glc} + \text{1 ATP} \rightarrow \text{1 G6p} + \text{1 ADP} + \text{1 H}^+$$

Location

$$\text{cytosol: 1 GLc} + \text{1 ATP} \rightarrow \text{1 G6p} + \text{1 ADP} + \text{1 H}^+$$

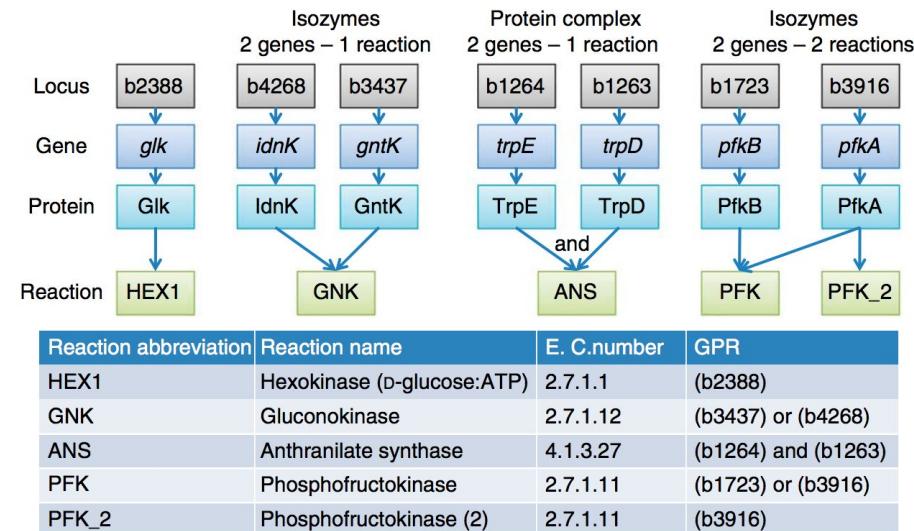
Constructing a genome-scale metabolic model

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



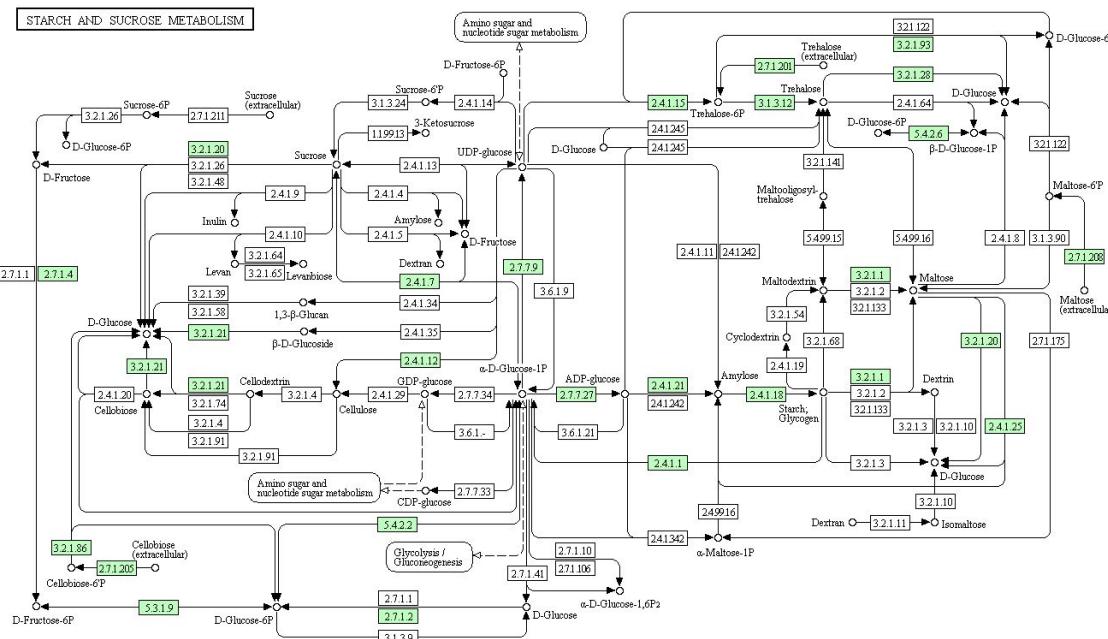
Constructing a genome-scale metabolic model

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
Lipoproteins	3.4
Peptidoglycan	2.5
Glycogen	2.5
Polyamines	0.4
Other	3.5
Total	100.00

Identification of missing functions

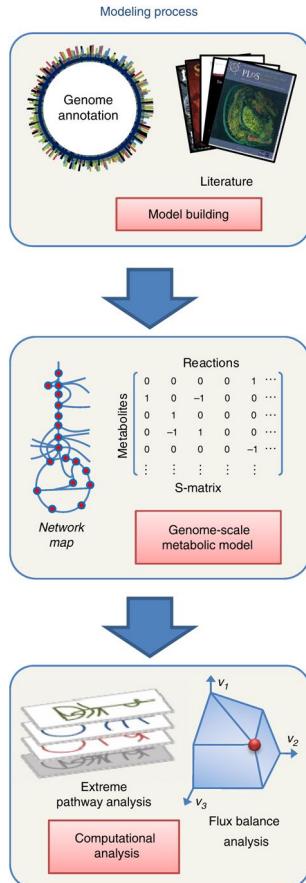


Constructing a genome-scale metabolic model

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 <i>et al.</i> 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	

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.

Flux balance analysis (FBA)

FBA: metabolic network → linear programming optimization problem.

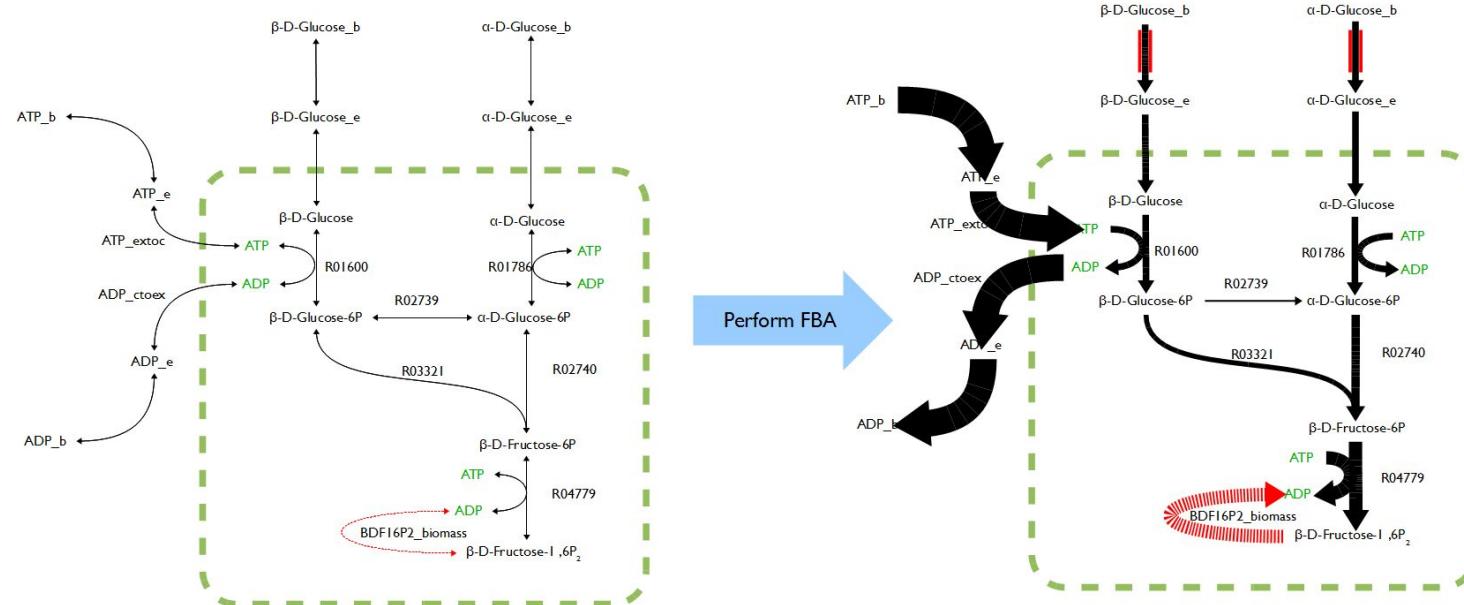
The main constraints in FBA: steady-state mass conservation of metabolites.

- 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.

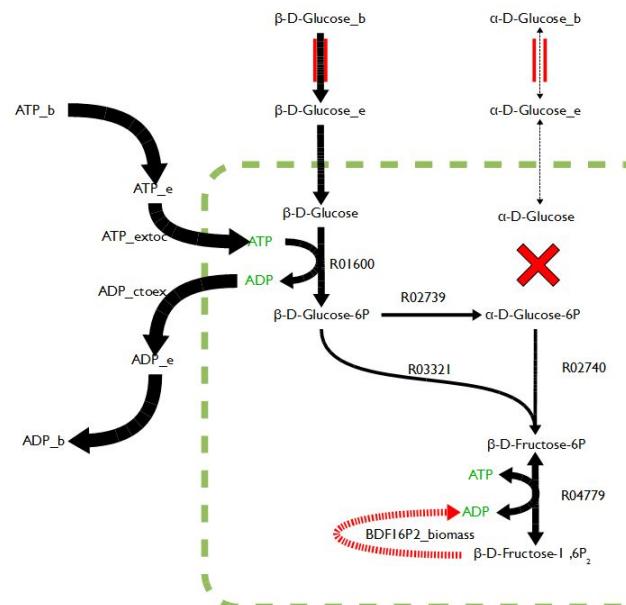
- The predicted flux through each reaction is proportional to the width of the line.
- **Red springy arrow:** Objective function; **Red bars:** Constraints on α -D-glucose and β -D-glucose import.



Flux balance analysis

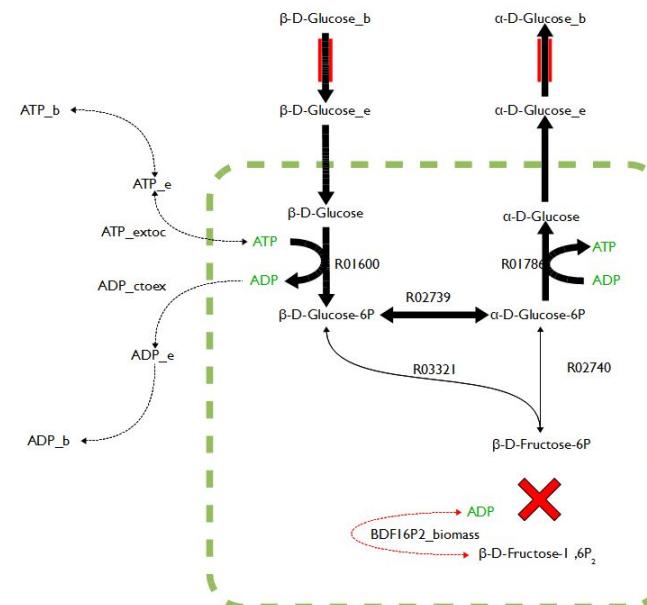
Non-lethal gene deletion in a metabolic network.

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



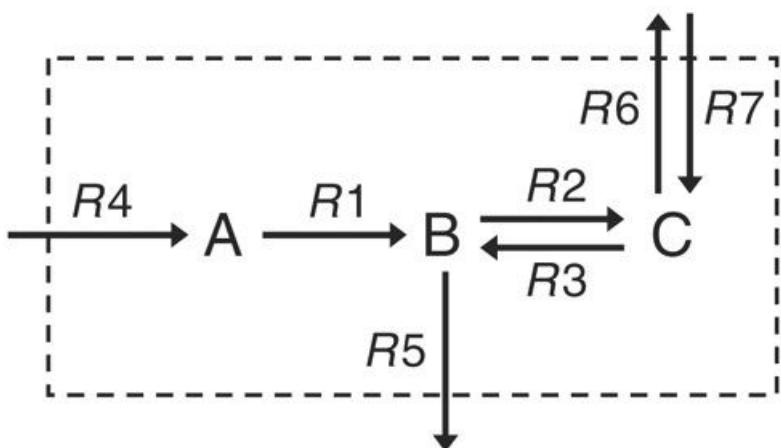
Lethal gene deletion in a metabolic network.

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



Flux balance analysis

1. Reaction network formalism



Chemical reactions	
Internal	Exchange
$R1: -1 A \rightarrow 1 B$	$R4: 1 A$
$R2: -1 B \rightarrow 1 C$	$R5: -1 B$
$R3: -1 C \rightarrow 1 B$	$R6: -1 C$
	$R7: 1 C$

S =

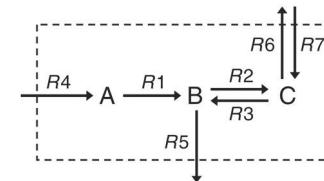
	$R1$	$R2$	$R3$	$R4$	$R5$	$R6$	$R7$
A	-1	0	0	1	0	0	0
B	1	-1	1	0	-1	0	0
C	0	1	-1	0	0	-1	1

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



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

Flux balance analysis

2. FBA formulation

Dynamic mass balance

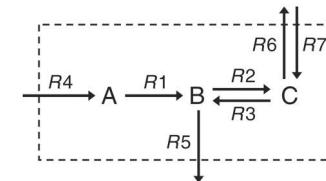
$$\frac{dC}{dt} = \mathbf{S}\mathbf{v}$$

C : Concentration

t : Time

\mathbf{S} : Stoichiometric matrix

\mathbf{v} : Flux vector



Steady-state assumption

$$\mathbf{S}\mathbf{v} = 0$$

LP formulation

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

Constraints:

$$\begin{matrix} & R1 & R2 & R3 & R4 & R5 & R6 & R7 \\ A & -1 & 0 & 0 & 1 & 0 & 0 & 0 \\ B & 1 & -1 & 1 & 0 & -1 & 0 & 0 \\ C & 0 & 1 & -1 & 0 & 0 & -1 & 1 \end{matrix} \begin{bmatrix} v_1 \\ \vdots \\ v_7 \end{bmatrix} = 0 \quad 0 \leq v_1, \dots, v_7 \leq 10$$

Flux balance analysis: Objective function

Objective function: Physiologically-meaningful or design-based objective for the interrogation or exploitation of a given system.

Examples:

- Maximizing...
 - biomass or cell growth
 - maximizing ATP production
 - maximizing the rate of synthesis of a particular product
- Minimizing...
 - ATP production
 - nutrient uptake (both to determine conditions of optimal metabolic energy efficiency)

Flux balance analysis: Constraints

No. of equations (one per reactant) << no. of unknown variables (reaction fluxes).

- An *under-determined* set of linear equations.
- Therefore, optimize fluxes given cellular objective given a bunch of constraints.
- **Principal constraint:** mass balance
- **Additional constraints:**
 - physico-chemical constraints
 - spatial or topological constraints
 - condition dependent environmental constraints
 - regulatory constraints
- All constraints together represent a set of linear equations.

Flux balance analysis

II. FBA formulation

Dynamic mass balance

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

C : Concentration

t : Time

\mathbf{S} : Stoichiometric matrix

\mathbf{v} : Flux vector

Steady-state assumption

$$\mathbf{Sv} = 0$$

LP formulation

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

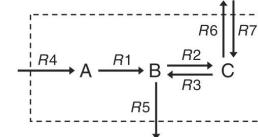
Constraints:

$$A \begin{bmatrix} R1 & R2 & R3 & R4 & R5 & R6 & R7 \end{bmatrix} \begin{bmatrix} v_1 \\ \vdots \\ v_7 \end{bmatrix} = 0 \quad 0 \leq v_1, \dots, v_7 \leq 10$$

where A is the stoichiometric matrix:

$$A = \begin{bmatrix} -1 & 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & -1 & 1 & 0 & -1 & 0 & 0 \\ 0 & 1 & -1 & 0 & 0 & -1 & 1 \end{bmatrix}$$

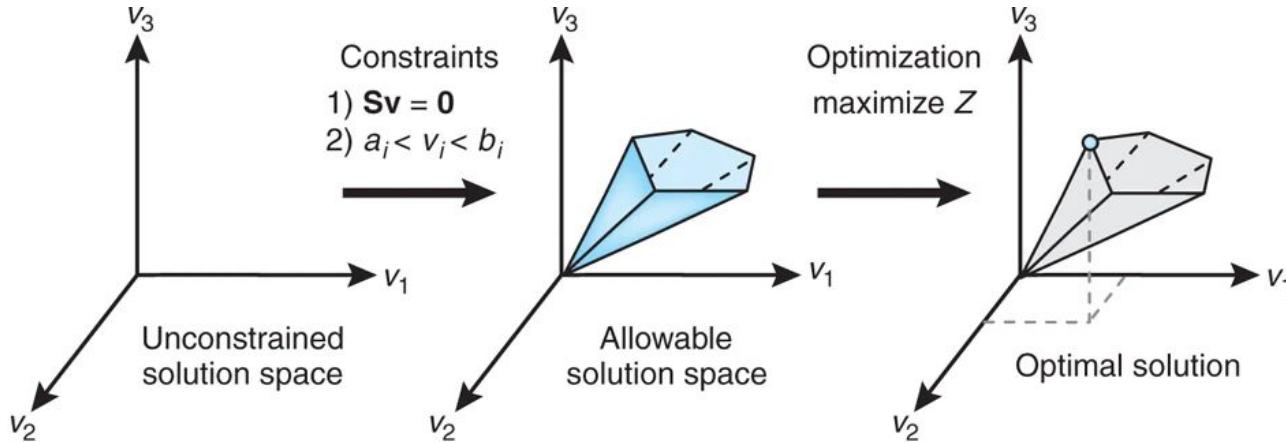
I. Reaction network formalism



Chemical reactions	
Internal	Exchange
R1: -1 A → 1 B	R4: 1 A
R2: -1 B → 1 C	R5: -1 B
R3: -1 C → 1 B	R6: -1 C
	R7: 1 C

$$S = \begin{bmatrix} R1 & R2 & R3 & R4 & R5 & R6 & R7 \\ \hline A & -1 & 0 & 0 & 1 & 0 & 0 & 0 \\ B & 1 & -1 & 1 & 0 & -1 & 0 & 0 \\ C & 0 & 1 & -1 & 0 & 0 & -1 & 1 \end{bmatrix}$$

Constraint-based modeling

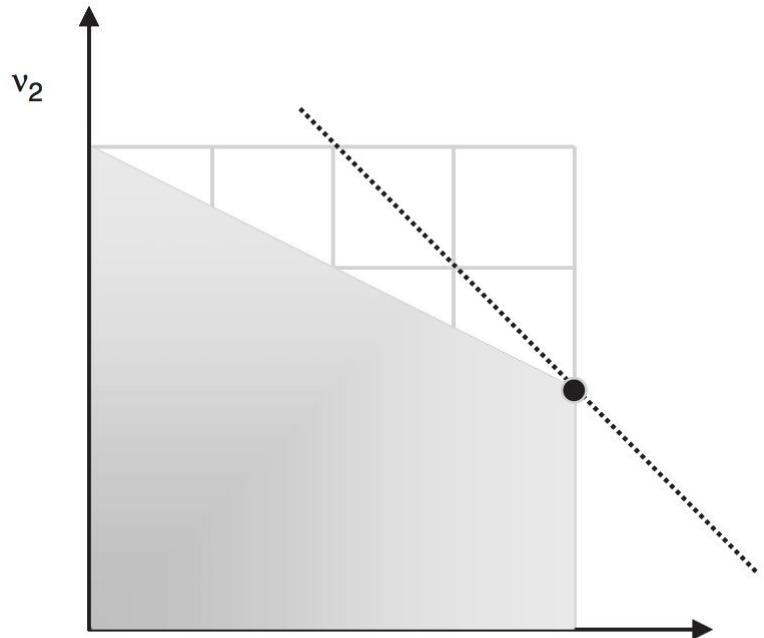


1. No constraints: flux may lie at any point in solution space.
2. Mass balance constraints (imposed by the stoichiometry) and capacity constraints (imposed by the lower and upper bounds: a_i & b_i): defines allowable solution space.
 - a. Any flux distribution within this space is allowable; Points outside this space are denied
3. Optimization of an objective function: A single optimal flux distribution that lies on the edge of the allowable solution space.

Constraint-based modeling

Linear programming

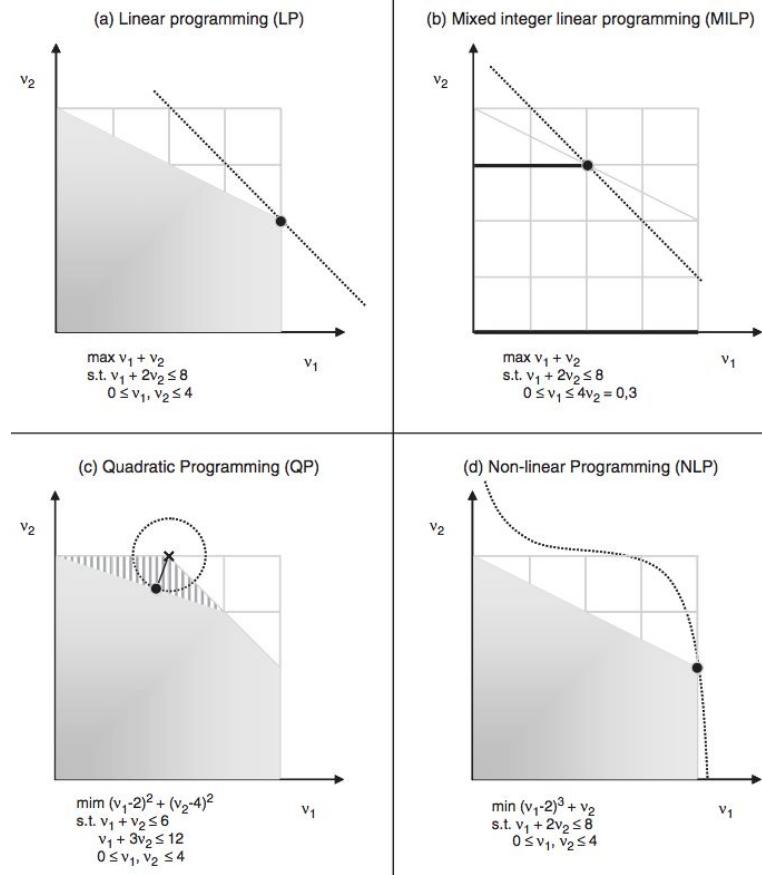
- Feasible solution space:
 - shaded area and solid lines
 - defined by:
 - flux capacities,
 - stoichiometric relationships, and
 - design specification (e.g. gene deletions).
- Objective function: dotted line
- Optimal solution: circular dot



$$\begin{aligned} & \max v_1 + v_2 \\ \text{s.t. } & v_1 + 2v_2 \leq 8 \\ & 0 \leq v_1, v_2 \leq 4 \end{aligned}$$

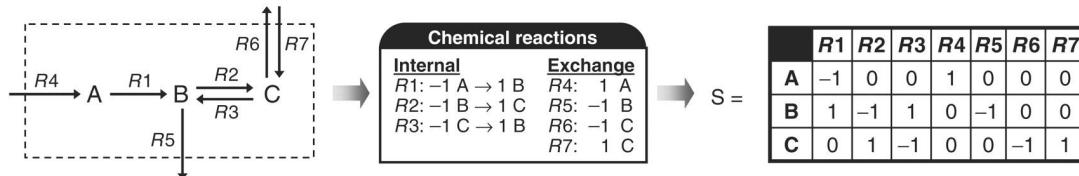
Constraint-based modeling

- Mixed Integer LP (MILP):
 - Integer variables are involved in a linear programming problem (e.g. binary variable formulation for gene deletion).
- Quadratic programming (QP):
 - Quadratic objective function subject to linear constraints.
 - This technique is generally used for finding the closest point to a specified point.
- Nonlinear programming (NLP):
 - Nonlinear objectives or constraints.
 - Generally difficult to solve for global optimal solution because of its non-convexity.



Flux balance analysis

I. Reaction network formalism



II. FBA formulation

Dynamic mass balance

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

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

Steady-state assumption

$$\mathbf{Sv} = 0$$

LP formulation

Objective: $\max Z = v_5$

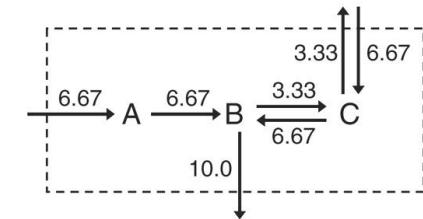
Constraints:

$$A \begin{bmatrix} R1 & R2 & R3 & R4 & R5 & R6 & R7 \end{bmatrix} \begin{bmatrix} v_1 \\ \vdots \\ v_7 \end{bmatrix} = 0 \quad 0 \leq v_1, \dots, v_7 \leq 10$$

III. Hypothetical flux distribution at steady-state

$$Z = 10$$

$$\mathbf{v} = [6.67 \ 3.33 \ 6.67 \ 6.67 \ 10.0 \ 3.33 \ 6.67]^T$$



Genome-scale metabolic network reconstruction & model

Organism	Strain	Genes	Version	GR	Mets	Rxns	Comp
<i>Bacillus subtilis</i>		4,225	model_v3	844	988	1,020	2 (c,e)
<i>Escherichia coli</i>	K12 MG1655	4,405	iAF1260	1,260	1,039	2,077	3 (c,e,p)
<i>Helicobacter pylori</i>	26695	1,632	iIT341	341	485	476	2 (c,e)
<i>Pseudomonas putida</i>	KT2440	5,350	iNJ746	746	911	950	3 (c,p,e)
<i>Pseudomonas putida</i>	KT2440	5,350	iJP815	815	886	877	2 (c,e)
<i>Pseudomonas aeruginosa</i>	PA01	5,640	iM01056	1,056	760	883	2 (c,e)
<i>Mycoplasma genitalium</i>	G-37	521	iPS189	189	274	262	2 (c,e)
<i>Lactobacillus plantarum</i>	WCFS1	3,009		721	531	643	2 (c,e)
<i>Streptomyces coelicolor</i>	A3(2)	8,042		700	500	700	2 (c,e)
<i>Leishmania major</i>	Friedlin	8,370	iAC560	560	1,101	1,112	8 (a,f,y,c,e,m,r,n)
<i>Saccharomyces cerevisiae</i>	Sc288	6,183	iMM904	904	713	1,412	8 (c,e,m,x,n,r,v,g)
<i>Homo sapiens</i>		28,783	Recon 1	1,496	2,766	3,311	8 (c,e,m,x,n,r,v,g)

Constructing a genome-scale metabolic model

Conversion of reconstruction into a model

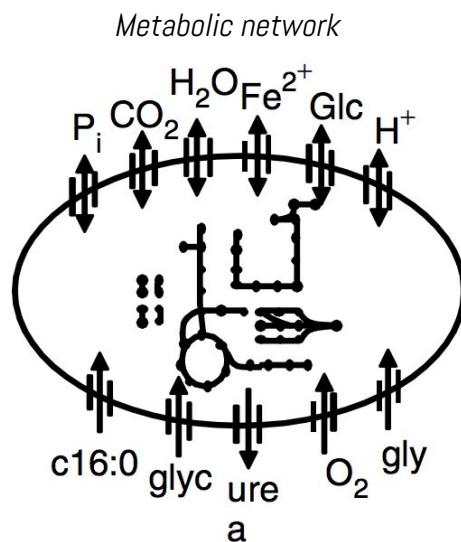
3. Conversion of reconstruction into computable format

- 38| Initialize the COBRA toolbox.
- 39| Load reconstruction into Matlab.
- 40| Verify S matrix.
- 41| Set objective function.
- 42| Set simulation constraints.

Constructing a genome-scale metabolic model

Conversion of reconstruction into a model

Mathematical representation



Stoichiometric matrix

Reaction ↓

Metabolite →

$$S = \begin{bmatrix} -1 & 0 & -1 & 0 & -1 & 0 & 0 & 0 \\ 1 & -1 & 0 & 0 & 0 & -1 & 0 & 0 \\ 0 & 1 & 1 & -1 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \end{bmatrix}$$

$S_{(\text{metabolite, reaction})}$

By definition:

- Substrates have negative coefficients (i.e., they are consumed)
- Products have positive coefficients (i.e., they are produced)

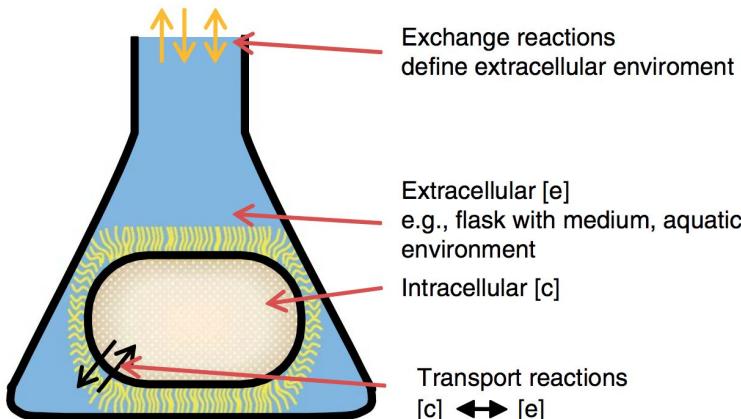
v is a vector of reaction fluxes

Conservation of mass: All steady states can be described by $Sv = 0$

Constructing a genome-scale metabolic model

Conversion of reconstruction into a model

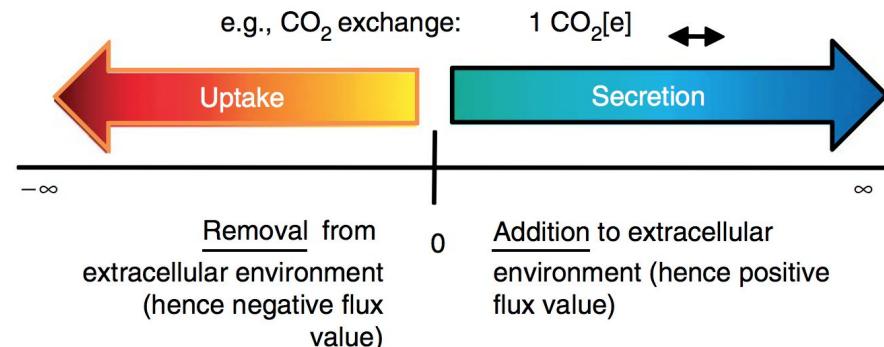
Definition of system boundaries



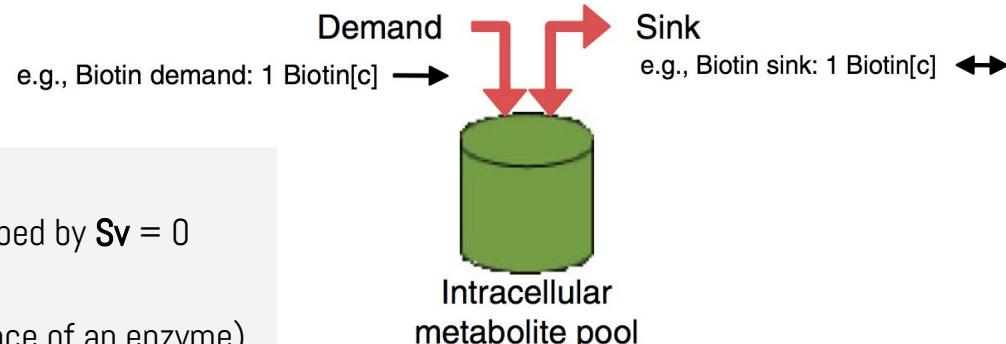
Constraints:

- Mass conservation: all steady states can be described by $\mathbf{Sv} = 0$
 - Thermodynamics (reaction directionality)
 - Enzyme capacity or regulation (i.e., presence/absence of an enzyme)

Exchange reactions



Demand/sink reactions



Constructing a genome-scale metabolic model

Network evaluation ≈ Debugging

4. Network evaluation

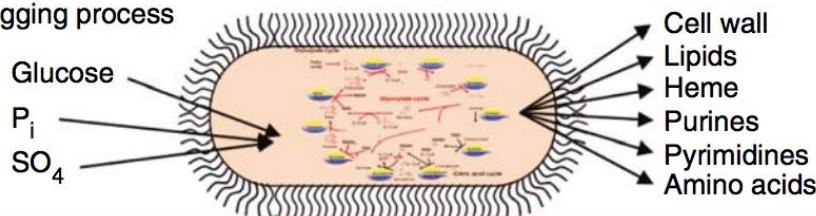
- 43–44| Test if network is mass-and charge balanced.
- 45| Identify metabolic dead-ends.
- 46–48| Perform gap analysis.
- 49| Add missing exchange reactions to model.
- 50| Set exchange constraints for a simulation condition.
- 51–58| Test for stoichiometrically balanced cycles.
- 59| Re-compute gap list.
- 60–65| Test if biomass precursors can be produced in standard medium.
- 66| Test if biomass precursors can be produced in other growth media.
- 67–75| Test if the model can produce known secretion products.
- 76–78| Check for blocked reactions.
- 79–80| Compute single gene deletion phenotypes.
- 81–82| Test for known incapabilities of the organism.
- 83| Compare predicted physiological properties with known properties.
- 84–87| Test if the model can grow fast enough.
- 88–94| Test if the model grows too fast.

Constructing a genome-scale metabolic model

Network evaluation ≈ Debugging

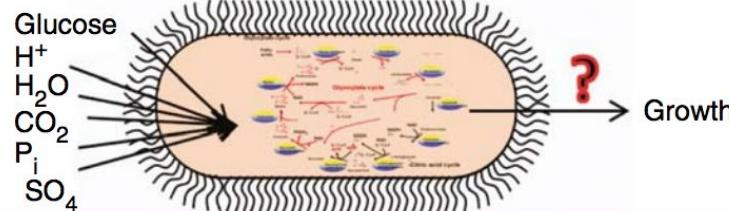
Analysis of biomass precursors synthesis

- Biomass precursors = cellular growth requirements
- Pathways to synthesize precursors must be complete (i.e., functional) for the network to simulate growth
- Testing synthesis of each separate biomass precursor is part of the debugging process



Analysis of growth in minimal medium

- Minimal medium is defined for many organisms and can be found in primary literature
- Contains at least 1 C-, N-, S- and P-source
- Auxotrophs may require the presence of addition metabolites



Test for growth on known carbon sources

- Exchange reactions define medium and environment
- Transport reactions allow network to consume carbon sources
- Biodegradative pathways that are required for carbon utilization

Secretion capability

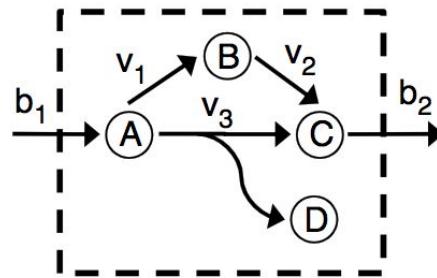
- Transport and exchange reactions are required in reconstruction to enable secretion
- Secretion may only occur under certain circumstances (e.g., D-lactic acid formation under anoxic conditions)
- Comparison with known secretion pattern of multiple metabolites (e.g., secretion of a certain ratio of CO₂ and acetate)

Constructing a genome-scale metabolic model

Network evaluation \asymp Debugging

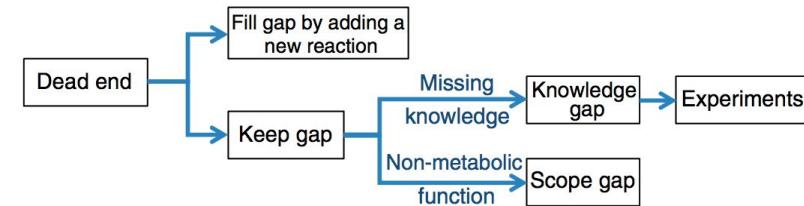
Identifying gaps

Connectivity based (topology):

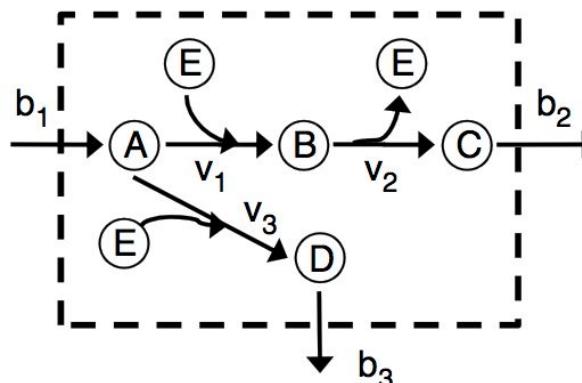


	v ₁	v ₂	v ₃	b ₁	b ₂
A	-1	0	-1	1	0
B	1	-1	0	0	0
C	0	1	1	0	-1
D	0	0	1	0	0

Dead end



Functionality based (computation)

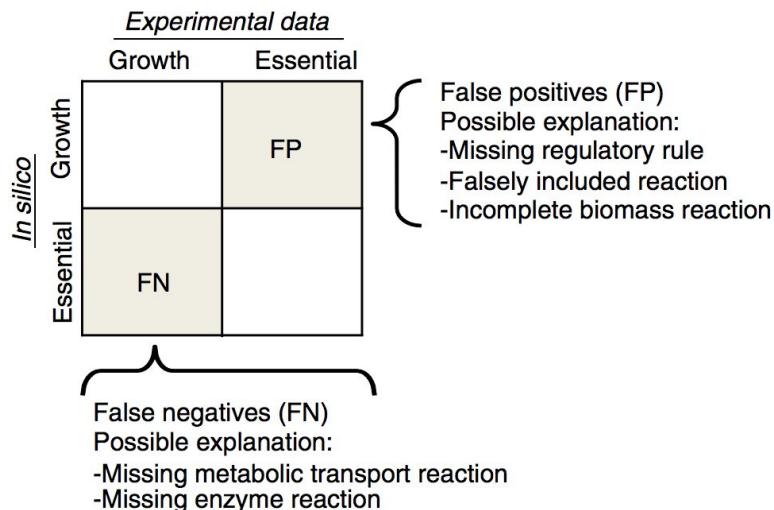


	v ₁	v ₂	v ₃	b ₁	b ₂	b ₃
A	-1	0	-1	1	0	0
B	1	-1	0	0	0	0
C	0	1	0	0	-1	0
D	0	0	1	0	0	-1
E	-1	1	-1	0	0	0

Constructing a genome-scale metabolic model

Network evaluation ≈ Debugging

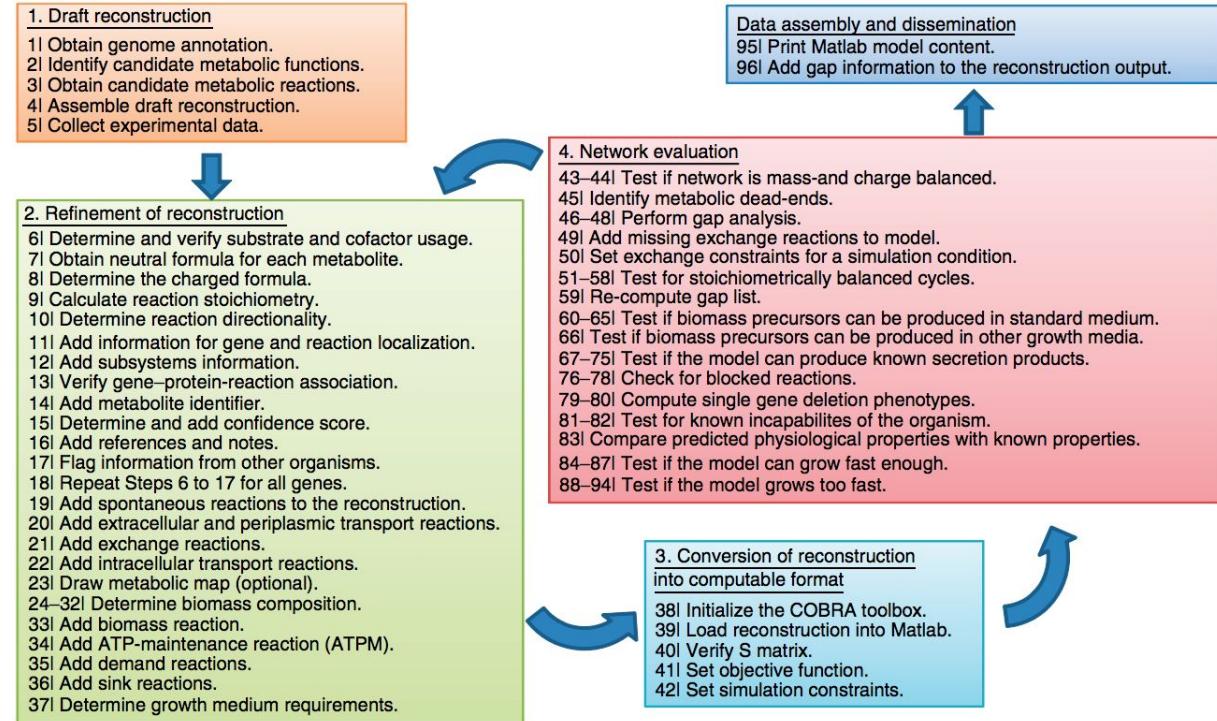
Gene essentiality



Constructing a genome-scale metabolic model

Procedure to iteratively reconstruct metabolic networks.

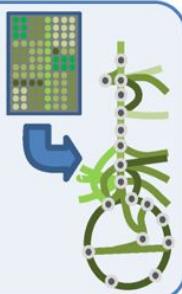
- Iterate stages 2–4 are continuously...
- ...until model predictions are similar to the phenotypic characteristics of the target organism and/or all experimental data for comparison are exhausted.



Genome-scale metabolic network reconstruction & model

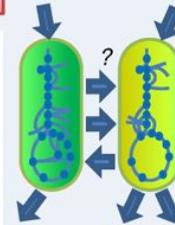
1. Contextualization of HT data

Several studies have overlaid gene microarray data on a metabolic GENRE to determine condition-dependent cell phenotypes. Metabolic GENREs have also been used to interpret metabolomic data, ¹³C flux data, and to link multiple data types.



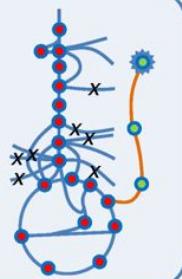
4. Interrogation of multi-species relationships

A dual-species metabolic model was built to study interactions between the syntrophic bacteria, *D. vulgaris* and *M. maripaludis*. Metabolic models have also been used in comparisons of multiple species, such as an analysis of pathway differences between four halophilic bacteria.



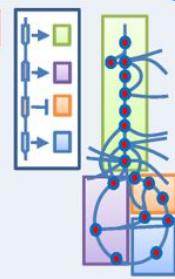
2. Guidance of metabolic engineering

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. meningitidis*.



5. Network property discovery

Metabolic GENREs have been used to study metabolite connectivity, and pathway redundancy *in silico*. Pathway-analysis tools have also spawned techniques such as flux coupling analysis, which has helped identify novel drug targets in *M. tuberculosis*.



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 transcriptional timing patterns in *S. cerevisiae*.

