Flux Balance Analysis

Learning Objectives

Each student should be able to:

- Explain flux balance analysis (FBA).
- Explain mass balanced linear equations.
- Explain the role of the biomass reaction.
- Explain how to create a stoichiometric matrix.
- Explain the purpose of loopless FBA
- Explain the purpose of parsimoneous FBA
- Demonstrate the ability to use Flux Balance Analysis to calculate metabolic fluxes with a COBRA model.

Flux Balance Analysis

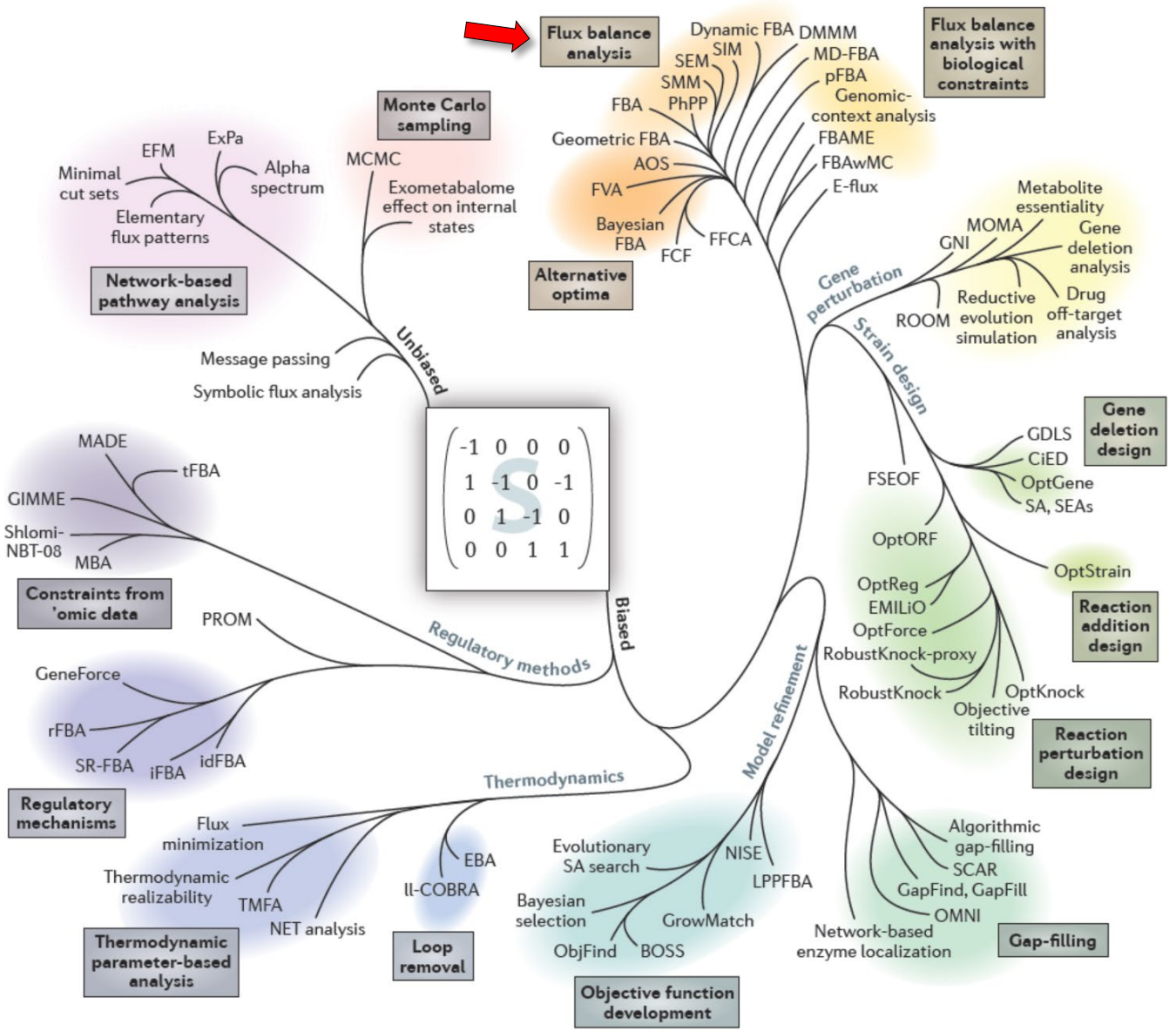
- Overview
- Mathematical Representation of Reactions & Constraints
- Biomass Reaction
- Flux Balance Analysis
- Loopless Flux Balance Analysis
- Parsimonious Flux Balance Analysis

Flux Balance Analysis

Orth, J. D., I. Thiele, et al. (2010). "What is flux balance analysis?" Nature biotechnology 28(3): 245-248.

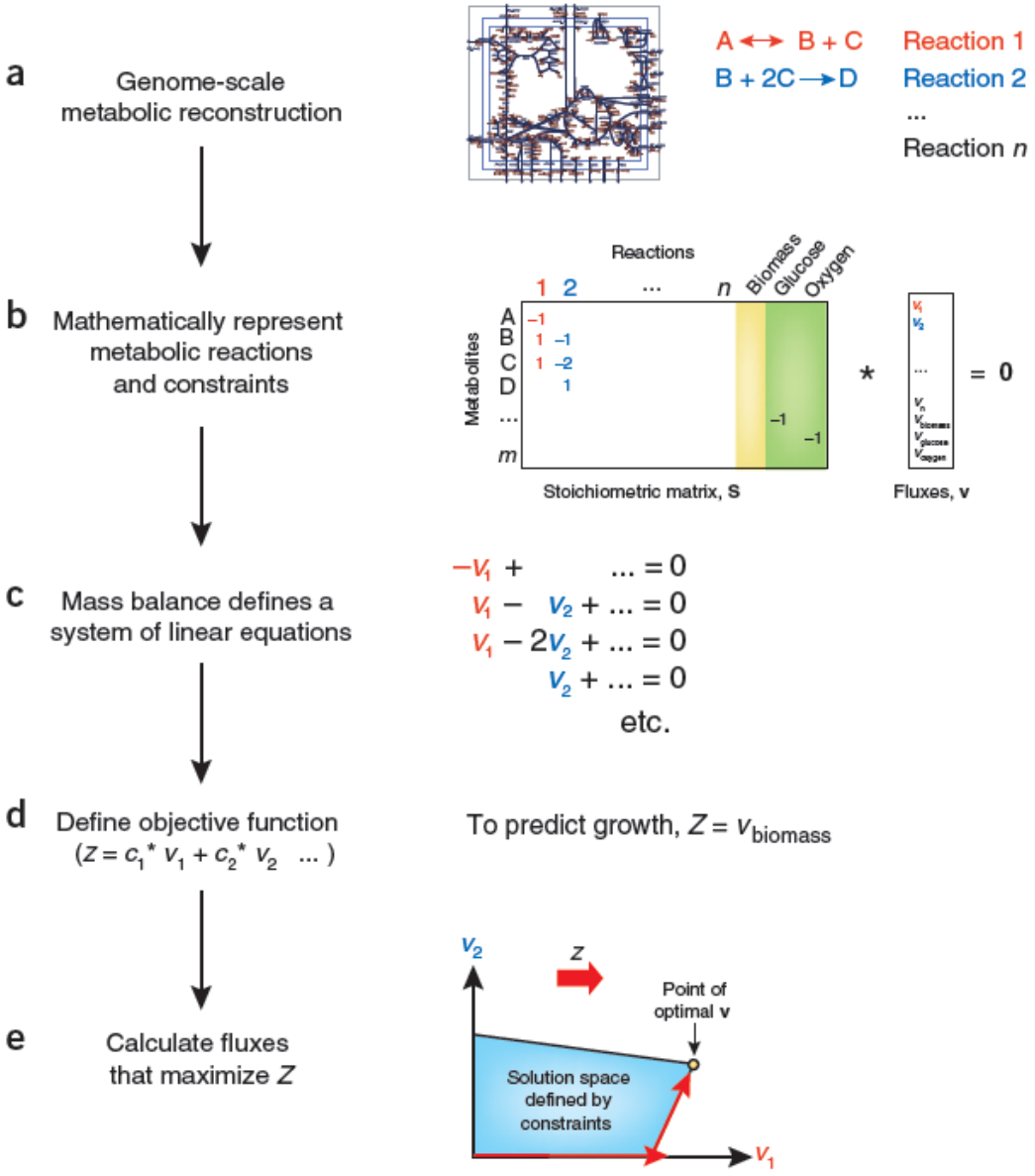
- Through the use of genome-scale metabolic network reconstructions, Flux Balance Analysis (FBA) can be used to calculate the flow of metabolites (flux) through a metabolic network. This capability makes it possible to predict the growth rate of an organism and/or the rate of production of a given metabolite.
- FBA has limitations! It does not use kinetic parameters, thus it cannot predict metabolite concentrations. It is also only capable of determining fluxes at steady state. Typically, FBA does not account for regulatory effects such as activation of enzymes by protein kinases or regulation of gene expression. Therefore, its predictions may not always be accurate.

Methods in Constraint-based Reconstruction and Analysis



Lewis, N. E., H. Nagarajan, et al. (2012). "Constraining the metabolic genotype-phenotype relationship using a phylogeny of in silico methods." Nature reviews. Microbiology 10(4): 291-305.

Formulation of Flux Balance Analysis



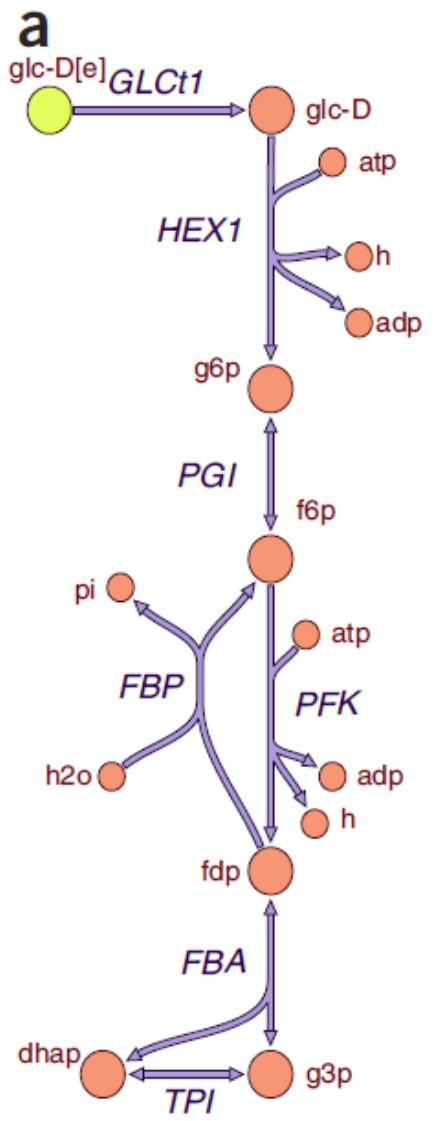
Orth, J. D., I. Thiele, et al. (2010). "What is flux balance analysis?" Nature biotechnology 28(3): 245-248.

Flux Balance Analysis

- Overview
- Mathematical Representation of Reactions & Constraints
- Biomass Reaction
- Flux Balance Analysis
- Loopless Flux Balance Analysis
- Parsimonious Flux Balance Analysis

Creating A Stoichiometric Matrix

The stoichiometric matrix, S, is the centerpiece of a mathematical representation of genome-scale metabolic networks. This matrix represents each reaction as a column and each metabolite as a row, where each numerical element is the corresponding stoichiometric coefficient.



b

	GLCt1	HEX1	PGI	PFK	FBP	FBA	TPI	EX_glc
glc-D[e]	−1	0	0	0	0	0	0	−1
glc-D	1	−1	0	0	0	0	0	0
atp	0	−1	0	−1	0	0	0	0
H	0	1	0	1	0	0	0	0
adp	0	1	0	1	0	0	0	0
g6p	0	1	−1	0	0	0	0	0
f6p	0	0	1	−1	1	0	0	0
fdp	0	0	0	1	−1	−1	0	0
pi	0	0	0	0	1	0	0	0
h2o	0	0	0	0	−1	0	0	0
g3p	0	0	0	0	0	1	1	0
dhap	0	0	0	0	0	1	−1	0

= S

c

UB =	∞	∞	∞	∞	∞	∞	∞	∞
LB =	−∞	0	−∞	0	0	−∞	−∞	−2

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

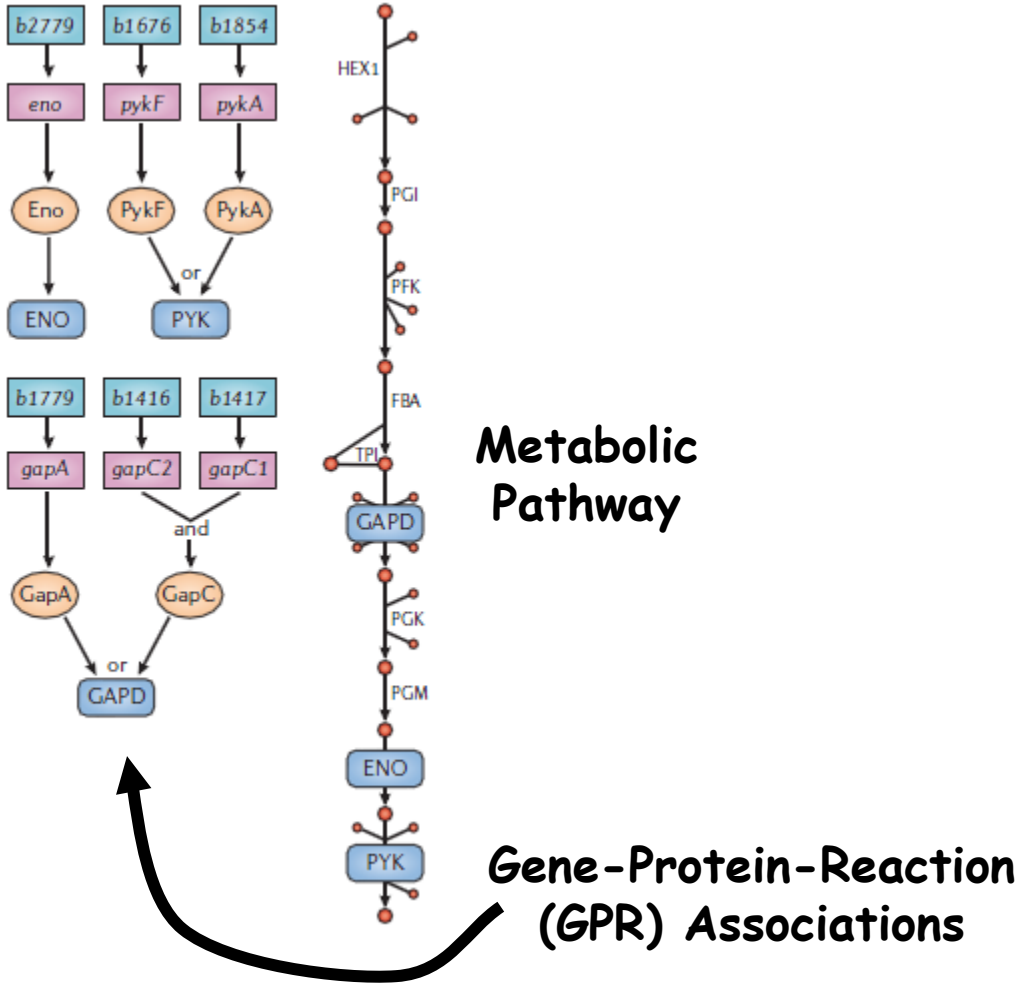
Genome-scale Metabolic Reconstruction

BIGG Database

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

Stoichiometric Matrix

	HEX1	PGI	PFK	FBA	TPI	GAPD	PGK	PGM	ENO	PYK
ATP	−1	0	−1	0	0	0	1	0	0	1
GLC	−1	0	0	0	0	0	0	0	0	0
ADP	1	0	1	0	0	0	−1	0	0	−1
G6P	1	−1	0	0	0	0	0	0	0	0
H	1	0	1	0	0	1	0	0	0	−1
F6P	0	1	−1	0	0	0	0	0	0	0
FDP	0	0	1	−1	0	0	0	0	0	0
DHAP	0	0	0	1	−1	0	0	0	0	0
G3P	0	0	0	1	1	−1	0	0	0	0
NAD	0	0	0	0	0	−1	0	0	0	0
PI	0	0	0	0	0	−1	0	0	0	0
13DPG	0	0	0	0	0	1	−1	0	0	0
NADH	0	0	0	0	0	1	0	0	0	0
3PG	0	0	0	0	0	0	1	−1	0	0
2PG	0	0	0	0	0	0	0	1	−1	0
PEP	0	0	0	0	0	0	0	0	1	−1
H_2O	0	0	0	0	0	0	0	0	1	0
PYR	0	0	0	0	0	0	0	0	0	1



Metabolic Pathway

Gene-Protein-Reaction (GPR) Associations

Reed, J. L., I. Famili, et al. (2006). "Towards multidimensional genome annotation." Nature reviews. Genetics 7(2): 130-141.

How can we use the Stoichiometric Matrix?

- The stoichiometric matrix, **S**, is a linear transformation of the flux vector, **v**, to a vector of time derivatives of the concentration vector, **x**.

$$\frac{d\mathbf{x}}{dt} = \mathbf{S} \cdot \mathbf{v}$$

- The concentration vector, **x**, represents the concentration of each of the metabolites.
- If we assume that a cell will be in a particular phenotype for a time much larger than the changing time of metabolites then we can also assume that the concentration pools for the metabolites will be non-changing thus setting $dx/dt = 0$. This is the steady state assumption of flux balance analysis.

$$\frac{d\mathbf{x}}{dt} = 0 = \mathbf{S} \cdot \mathbf{v}$$

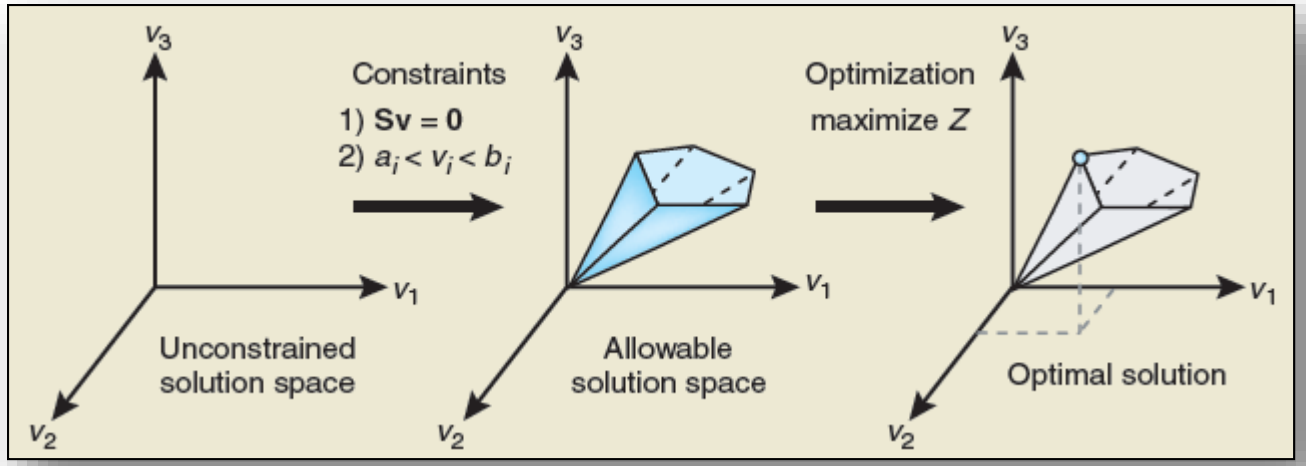
- Since there are normally many more reactions (columns) than metabolites (rows), more unknown variables than equations, then there is no unique solutions (could be a large number of solutions).
- Need to find a way to constrain the solution space!

Reactions

Metabolites

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

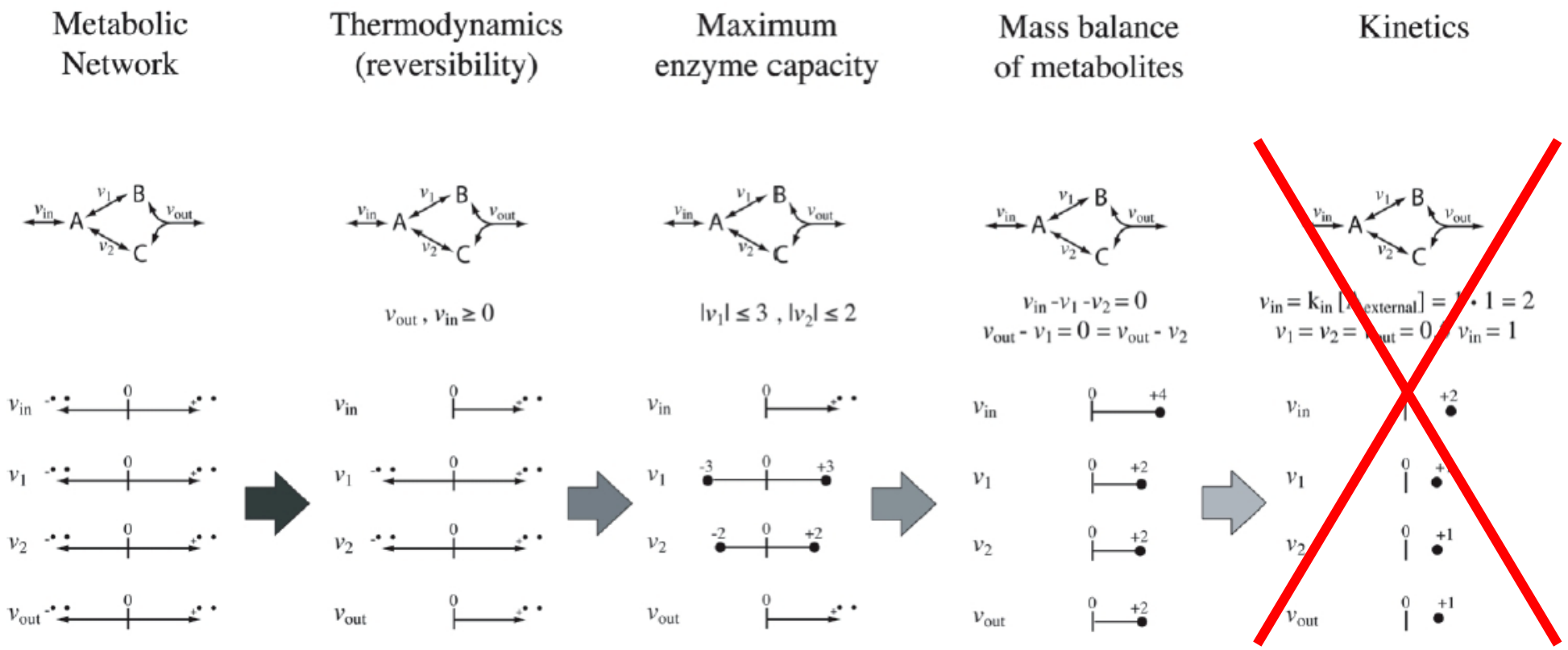
The Conceptual Basis of Constraint-based Modeling



With no constraints, the flux distribution of a biological network may lie at any point in a solution space. When mass balance constraints imposed by the stoichiometric matrix S (label 1) and capacity constraints imposed by the lower and upper bounds (a_i and b_i) (label 2) are applied to a network, it defines an allowable solution space. The network may acquire any flux distribution within this space, but points outside this space are denied by the constraints. Through optimization of an objective function using linear programming, FBA can identify a single optimal flux distribution that lies on the edge of the allowable solution space.

Orth, J. D., I. Thiele, et al. (2010). "What is flux balance analysis?" Nature biotechnology 28(3): 245-248.

Role of Constraints



REI601M, Introduction to Systems Biology, Dr. Innes Thiele,2012, https://systemsbiology.hi.is/wiki/REI601M

Flux Optimization

(Linear Programming or Linear Optimization Problem)

Maximize the objective function

$$Z = \sum_i c_i v_i = \mathbf{c} \cdot \mathbf{v}$$

with the following constraints

$$\frac{d\mathbf{x}}{dt} = \mathbf{S} \cdot \mathbf{v} = \mathbf{0}$$

$$\alpha_j \le v_j \le \beta_j$$

where

$\mathbf{x}$ = concentration vector
$\mathbf{v}$ = flux vector
$\mathbf{c}$ = objective function weights
$\mathbf{S}$ = Stoichiometric matrix
α_j = Lower bound of flux
β_j = upper bound of flux

The goal is to create and objective function that is biologically meaningful. These could include;

1. Cellular growth (maximization)
2. Particular metabolite engineering (maximization)
3. Energy consumption (minimization)

For the case of cellular growth as the objective function (Biomass Function)

1. "It has been shown that under rich growth conditions (i.e. no lack of phosphate and nitrogen), *E. coli* grows in a stoichiometrically optimal manner." (Schilling 2001, Edwards 1994)
2. "It is reasonable to hypothesize that unicellular organisms have evolved toward maximal growth performance." (Segre, 2002.)

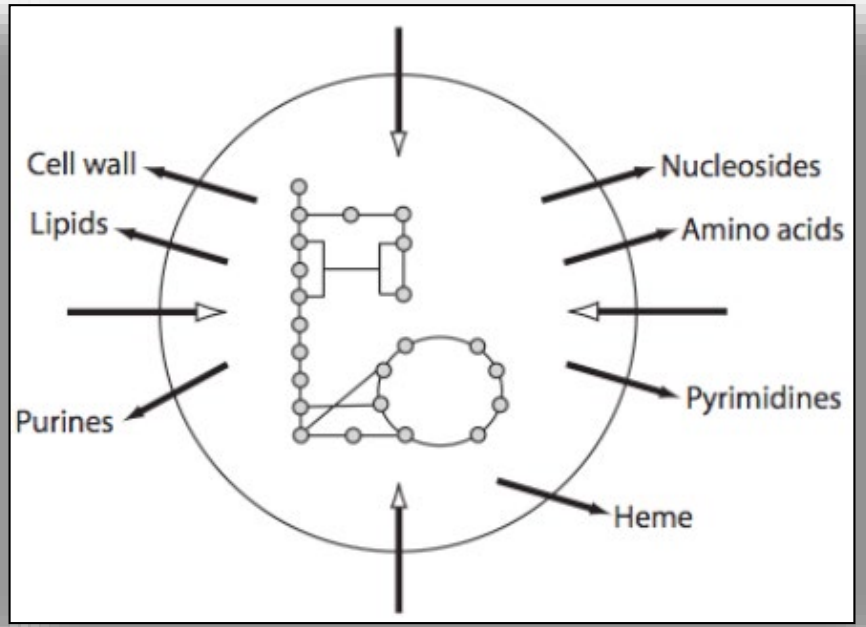
Flux Balance Analysis

- Overview
- Mathematical Representation of Reactions & Constraints
- Biomass Reaction
- Flux Balance Analysis
- Loopless Flux Balance Analysis
- Parsimonious Flux Balance Analysis

Biomass Precursors

- The biomass reaction accounts for all the fractional contributions from biosynthetic precursors and key cofactors to create 1g of biomass.
- These factional contributions need to be determined experimentally for cells growing in **log phase**.
- It may not be possible to obtain a detailed biomass composition for the target organism. In this case, one can estimate the relative fraction of each precursor from existing databases.

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



Thiele, I. and B. O. Palsson (2010). "A protocol for generating a high-quality genome-scale metabolic reconstruction." Nature protocols 5(1): 93-121.

E.coli Precursor Metabolites

Nicotinamide coenzymes
Glycerol-3-phosphate -> Phospholipids

Sugar nucleotides

Amino sugars

Serine Family
Serine -> Tryptophan
-> Ethanolamine
-> 1-C units
Glycine -> Purine nucleotides
Cysteine

Aromatic Family
Tyrosine
Tryptophan
Phenylalanine

Chorismate

Vitamins and cofactors
Ubiquinone
Menaquinone
Folates

Pyruvate family
Alanine
Valine
Leucine
Isoleucine
Isoprenoids

Fatty Acids
Murein
Leucine

Aspartate family
Asparagine
Threonine
Methionine -> Spermidine
Aspartate -> Nicotinamide coenzymes
-> Pyrimidine nucleotides
Lysine

Glutamate family
Glutamate -> Hemes
Glutamine
Arginine -> Polyamines
Proline

Vitamins and cofactors
Folates
Riboflavin
Coenzyme A
Adenosylcobalamine
Nicotinamide

Purine nucleotides
Pyrimidine nucleotides

Phosphoribosyl pyrophosphate

Tryptophan
Histidine

2-Keto-3-deoxyoctanate

Heptose in LPS

Heme

M. Schaechter, et al, Microbe, ASM Press, 2006, p. 116

Maintenance Energy Requirements

- To simulate growth, the energy required to maintain the cell growth must be accounted for.
- Two forms of energy are required; growth associated maintenance (GAM) energy and non-growth associated maintenance (NGAM) energy (e.g. turgor pressure).
- GAM reaction accounts for the energy (ATP) necessary to replicate a cell. It is represented in the model by

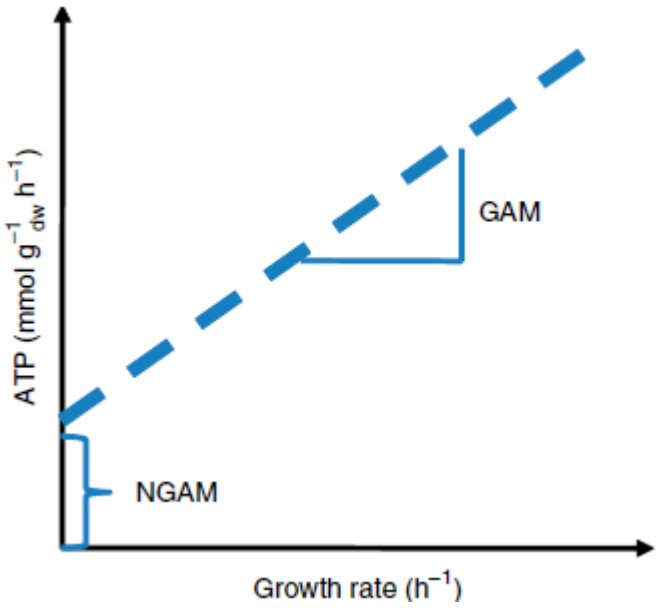
$$x\ ATP + x\ H_2O \rightarrow x\ ADP + x\ P_i + x\ H^+$$

Where x is the number of required phosphate bonds (59.81 in core model). This will be included in the biomass reaction

- The NGAM reaction (ATPM) is given by

$$1\ ATP + 1 H_2O \rightarrow 1\ ADP + 1\ P_i + 1\ H^+$$

where the flux through this reaction is constrained by experimental data to 8.39 mmol $g_{DW}^{-1}h^{-1}$

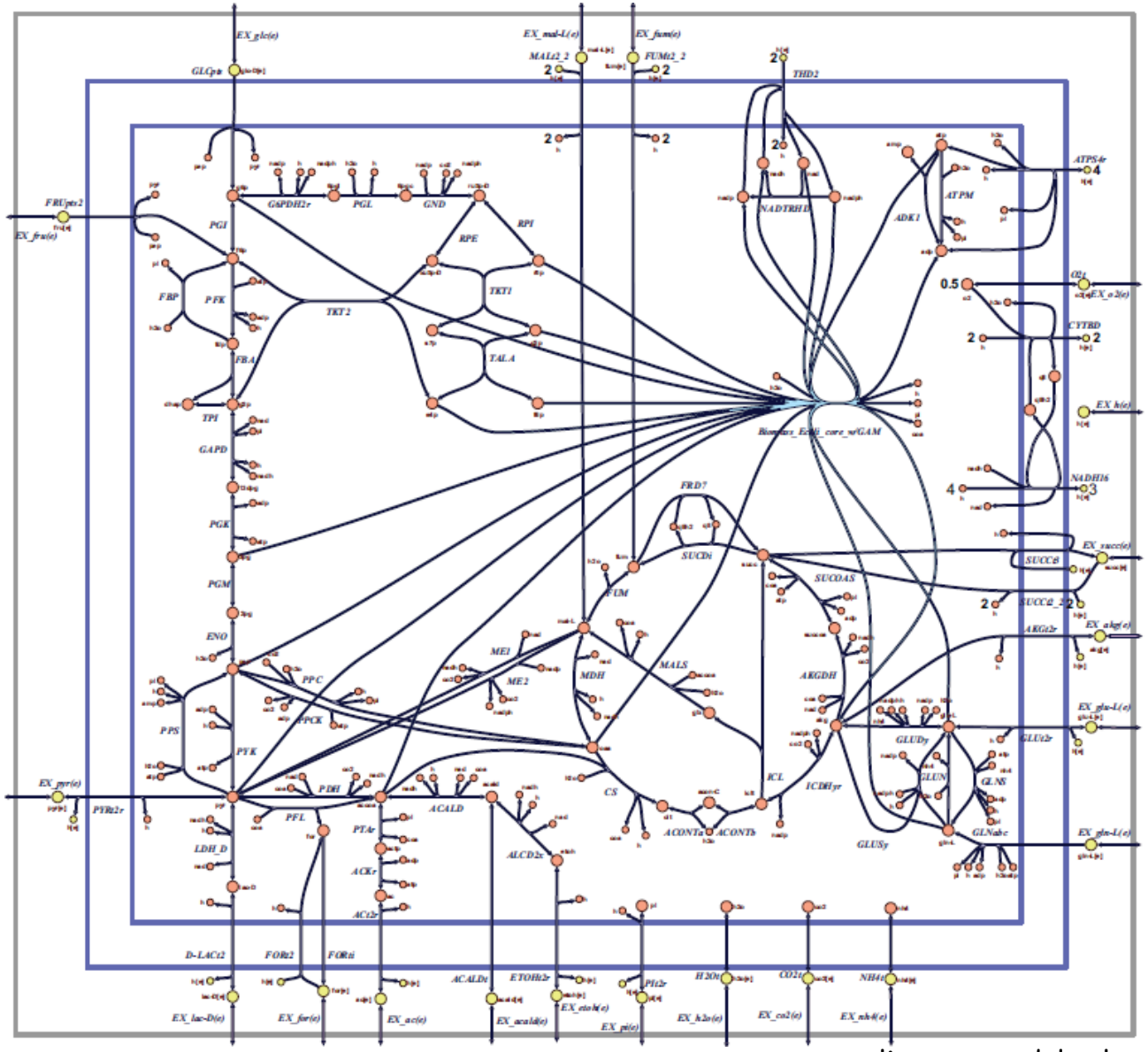


Thiele, I. and B. O. Palsson (2010). "A protocol for generating a high-quality genome-scale metabolic reconstruction." Nature protocols 5(1): 93-121.

Biomass Reaction For *E.coli* Core Model

(1.496) 3pg + (3.7478) accoa + (59.8100) atp + (0.3610) e4p + (0.0709) f6p + (0.1290) g3p + (0.2050) g6p + (0.2557) gln-L + (4.9414) glu-L + (59.8100) h2o + (3.5470) nad + (13.0279) nadph + (1.7867) oaa + (0.5191) pep + (2.8328) pyr + (0.8977) r5p --> (59.8100) adp + (4.1182) akg + (3.7478) coa + (59.8100) h + (3.5470) nadh + (13.0279) nadp + (59.8100) pi

* Key Cofactors



ecoli_core_models.xls

iaf1260 Biomass Objective Function

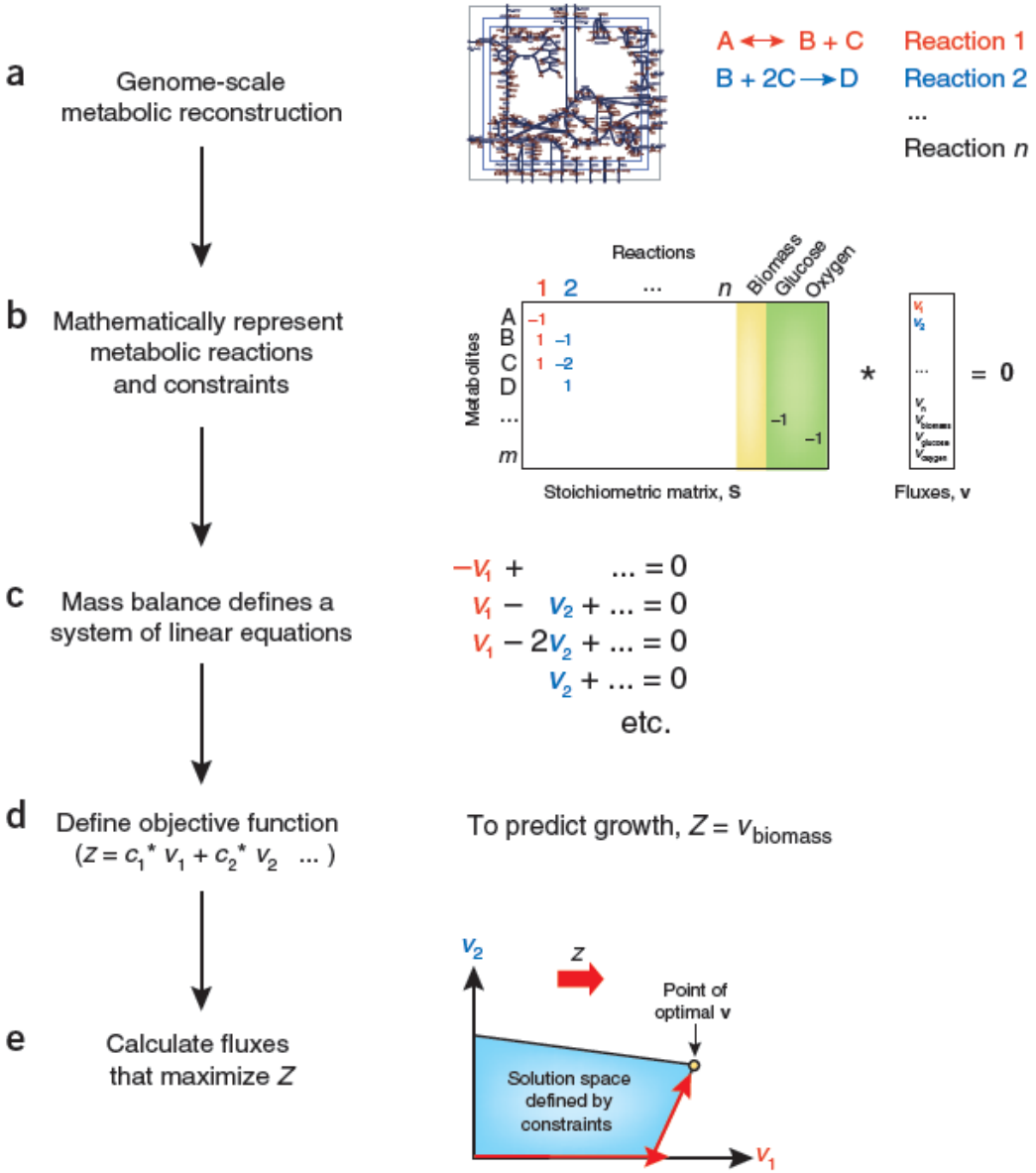
(Ec_biomass_iAF1260_core_59p81M)

Z = 0.000223 10fthf[c] + 0.000223 2ohph[c] + 0.5137 ala-L[c] + 0.000223 amet[c] + 0.2958 arg-L[c] + 0.2411 asn-L[c] + 0.2411 asp-L[c] + 59.984 atp[c] + 0.004737 ca2[c] + 0.004737 cl[c] + 0.000576 coa[c] + 0.003158 cobalt2[c] + 0.1335 ctp[c] + 0.003158 cu2[c] + 0.09158 cys-L[c] + 0.02617 datp[c] + 0.02702 dctp[c] + 0.02702 dgtp[c] + 0.02617 dttp[c] + 0.000223 fad[c] + 0.007106 fe2[c] + 0.007106 fe3[c] + 0.2632 gln-L[c] + 0.2632 glu-L[c] + 0.6126 gly[c] + 0.2151 gtp[c] + 54.462 h2o[c] + 0.09474 his-L[c] + 0.2905 ile-L[c] + 0.1776 k[c] + 0.01945 kdo2lipid4[e] + 0.4505 leu-L[c] + 0.3432 lys-L[c] + 0.1537 met-L[c] + 0.007895 mg2[c] + 0.000223 mlthf[c] + 0.003158 mn2[c] + 0.003158 mobd[c] + 0.01389 murein5px4p[p] + 0.001831 nad[c] + 0.000447 nadp[c] + 0.011843 nh4[c] + 0.02233 pe160[c] + 0.04148 pe160[p] + 0.02632 pe161[c] + 0.04889 pe161[p] + 0.1759 phe-L[c] + 0.000223 pheme[c] + 0.2211 pro-L[c] + 0.000223 pydx5p[c] + 0.000223 ribflv[c] + 0.2158 ser-L[c] + 0.000223 sheme[c] + 0.003948 so4[c] + 0.000223 thf[c] + 0.000223 thmpp[c] + 0.2537 thr-L[c] + 0.05684 trp-L[c] + 0.1379 tyr-L[c] + 5.5e-005 udcpdp[c] + 0.1441 utp[c] + 0.4232 val-L[c] + 0.003158 zn2[c] -> 59.81 adp[c] + 59.81 h[c] + 59.806 pi[c] + 0.7739 ppi[c]

Flux Balance Analysis

- Overview
- Mathematical Representation of Reactions & Constraints
- Biomass Reaction
- Flux Balance Analysis
- Loopless Flux Balance Analysis
- Parsimonious Flux Balance Analysis

Formulation of Flux Balance Analysis



Orth, J. D., I. Thiele, et al. (2010). "What is flux balance analysis?" Nature biotechnology 28(3): 245-248.

COBRApy Flux Balance Analysis

- Called by "solution = model.optimize()"
- This results in the calculation of the 1) reaction fluxes, 2) reduced costs, and 3) shadow prices for the given environmental conditions of the model.
- The reaction fluxes represent the steady-state fluxes of the model.
- Reduced costs are associated with each reaction's flux and signify the amount by which the objective function is decreased if that reaction flux is increased. For instance, if the input flux of glucose shows a reduced cost of -x, it means that increasing that flux by one unit will increase of the objective function by x units.
- The shadow prices define the incremental change in the objective function if a constraining flux is incrementally changed. The sensitivity of an FBA solution is indicated by shadow prices. They indicate how much the addition of a given metabolite will increase or decrease the objective.

solution.fluxes

ACALD	0.000000e+00
ACALDt	0.000000e+00
ACKr	0.000000e+00
ACONTa	6.007250e+00
ACONTb	6.007250e+00
ACt2r	0.000000e+00
ADK1	0.000000e+00
AKGDH	5.064376e+00
AKGt2r	0.000000e+00
ALCD2x	0.000000e+00
ATPM	8.390000e+00
ATPS4r	4.551401e+01
Biomass_Ecoli_core	8.739215e-01
CO2t	-2.280983e+01
CS	6.007250e+00
CYTBD	4.359899e+01
D_LACt2	0.000000e+00
ENO	1.471614e+01
ETOHt2r	0.000000e+00

solution.reduced_costs

ACALD	6.938894e-18
ACALDt	0.000000e+00
ACKr	1.040834e-17
ACONTa	0.000000e+00
ACONTb	1.387779e-17
ACt2r	0.000000e+00
ADK1	-0.000000e+00
AKGDH	1.318390e-16
AKGt2r	1.387779e-17
ALCD2x	0.000000e+00
ATPM	-1.018497e-02
ATPS4r	0.000000e+00
Biomass_Ecoli_core	1.942890e-16
CO2t	-1.387779e-17
CS	4.163336e-17
CYTBD	-8.673617e-19
D_LACt2	1.301043e-18
ENO	0.000000e+00
ETOHt2r	0.000000e+00

solution.shadow_prices

13dpg_c	-4.710549e-02
2pg_c	-4.201301e-02
3pg_c	-4.201301e-02
6pgc_c	-9.166475e-02
6pgl_c	-9.039162e-02
ac_c	-2.418931e-02
ac_e	-2.291619e-02
acald_c	-3.437428e-02
acald_e	-3.437428e-02
accoa_c	2.673555e-02
acon_C_c	-7.129480e-02
actp_c	-2.928179e-02
adp_c	5.092486e-03
akg_c	-6.238295e-02
akg_e	-6.110983e-02
amp_c	1.018497e-02
atp_c	0.000000e+00
cit_c	-7.129480e-02
co2_c	1.387779e-17

FBA_using_COBRApy.ipynb

COBRApy Simulations based on Flux Balance Analysis

(FBA_using_COBRApy.ipynb)

Flux Balance Analysis using COBRApy

Simulations using flux balance analysis can be solved using COBRApy model.optimize() command. This module will maximize or minimize (maximizing is the default) flux based on the objective reactions.

To start with, import the cobra.test package and load the *E.coli* core model which is labeled as "textbook."

In [1]:
```
import cobra.test
model = cobra.test.create_test_model("textbook")
```

Running a FBA simulation

Place the results of the simulation in an object called "solution" and then print the result of the simulation which is the numerical value of the objective function which in the case corresponds to the growth rate

In [2]:
```
solution = model.optimize()
print(solution)
```

```
<Solution 0.874 at 0x7f9fd0667820>
```

The **model.optimize()** function will return a Solution object. A solution object has several attributes:

- objective_value: the objective value
- status: the status from the linear programming solver
- fluxes: a pandas series with flux indexed by reaction identifier. The flux for a reaction variable is the difference of the primal values for the forward and reverse reaction variables.
- shadow_prices: a pandas series with shadow price indexed by the metabolite identifier.

FBA Lecture Examples

(FBA_Lecture_Examples.ipynb)

FBA Lecture Examples

In [1]:
```
import cobra.test
#import cobra.test
import pandas as pd
from cobra.util.solver import linear_reaction_coefficients
import escher
from escher import Builder
```

Aerobic Simulation - *E.coli* Core Model

In [2]:
```
import cobra.test
# Load the model
model = cobra.test.create_test_model("textbook")
# Set the inputs
model.reactions.EX_o2_e.lower_bound = -1000
model.reactions.EX_glc__D_e.lower_bound = -20
# Optimize
solution = model.optimize()
model.summary()
```

Out[2]:

Objective

1.0 Biomass_Ecoli_core = 1.7905689707194798

Flux Balance Analysis

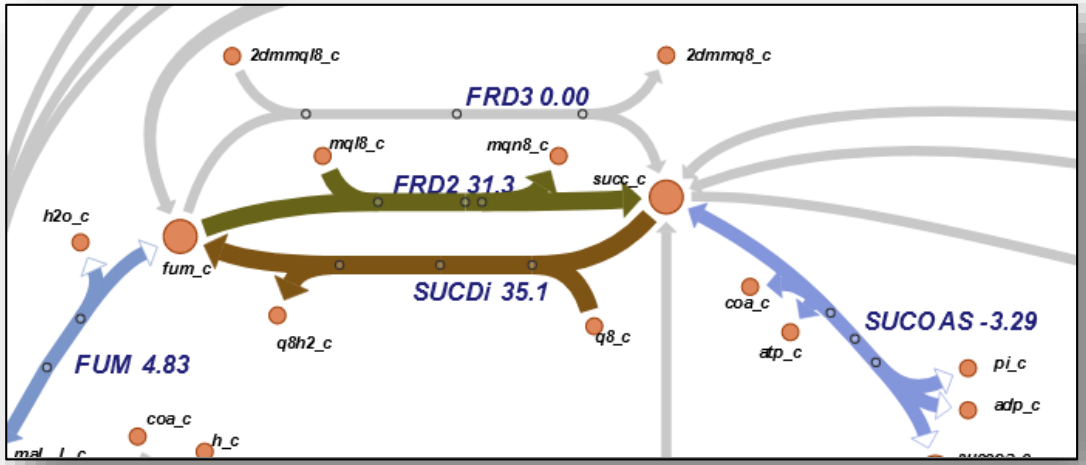
- Overview
- Mathematical Representation of Reactions & Constraints
- Biomass Reaction
- Flux Balance Analysis
- Loopless Flux Balance Analysis
- Parsimonious Flux Balance Analysis

Loopless Flux Balance Analysis

- There are times when the FBA solution will include loops. This is seen by high flux values within the loop. To the right the flux value of the loop reactions is ~10 times larger than the flux entering or leaving the loop.
- Classical loopless approaches are computationally expensive to solve due to the added mixed-integer constraints.
- A much faster, and pragmatic approach is instead to post-process flux distributions to simply set fluxes to zero wherever they can be zero without changing the fluxes of any exchange reactions in the model. CycleFreeFlux is an algorithm that can be used to achieve this and in cobrapy it is implemented in the

  cobra.flux_analysis.loopless_solution

  function. This loopless_solution will identify the closest flux distribution to the original one.



https://cobrapy.readthedocs.io/en/latest/loopless.html

Loopless Flux Balance Analysis

Loopless_FBA.ipynb

Show the difference between clissical FBA and Loopless FBA.

In [1]:
```
import cobra.test
import pandas as pd
pd.set_option('display.max_rows', 500)

from cobrapy_bigg_client import client
model = client.download_model('iJO1366', save=False) # Loading the model to the simulation
#model.reactions.EX_o2_e.lower_bound = -1000
#model.reactions.EX_glc__D_e.lower_bound = -10
```

```
Set parameter Username
Academic license - for non-commercial use only - expires 2022-10-10
```

Compute the regular FBA and the Loopless FBA

In [2]:
```
#from cobra import Reaction, Metabolite, Model
from cobra.flux_analysis.loopless import add_loopless, loopless_solution
model.objective = 'BIOMASS_Ec_iJO1366_core_53p95M'
%time fba_solution = model.optimize()
%time loopless_solution = loopless_solution(model)
```

```
Wall time: 183 ms
Wall time: 443 ms
```

Note how much more time it takes to remove the loops!

Let's create an Escher map showing the regular FBA analysis.

Flux Balance Analysis

- Overview
- Mathematical Representation of Reactions & Constraints
- Biomass Reaction
- Flux Balance Analysis
- Loopless Flux Balance Analysis
- Parsimonious Flux Balance Analysis

Parsimonious FBA

- Parsimonious FBA (pFBA) finds a flux distribution which gives the
  1. optimal growth rate (maximizes biomass function),
  2. but minimizes the total sum of flux (sum of absolute value of all fluxes).
- The underlying assumption is that, under growth pressure, there is a selection for strains that can process the growth substrate the most rapidly and efficiently while using the minimum amount of enzyme.

Lewis, N. E., K. K. Hixson, et al. (2010). "Omic data from evolved E. coli are consistent with computed optimal growth from genome-scale models." Molecular Systems Biology 6: 390.

Parsimonious Flux Balance Analysis

pFBA.ipynb

Parsimonious FBA (pFBA) finds a flux distribution which gives the optimal growth rate (maximize biomass function), but minimizes the total sum of flux.

In [1]:
```
import cobra.test
import pandas as pd
from cobra.flux_analysis.loopless import add_loopless, loopless_solution
from cobrapy_bigg_client import client
pd.set_option('display.max_rows', 500)

model = client.download_model('iJO1366', save=False) # Loading the model from the BIGG database
```

```
Set parameter Username
Academic license - for non-commercial use only - expires 2022-10-10
```

In [2]:
```
model.objective = 'BIOMASS_Ec_iJO1366_core_53p95M'
%time fba_solution = model.optimize()
%time pfba_solution = cobra.flux_analysis.pfba(model)
```

```
Wall time: 181 ms
Wall time: 203 ms
```

Note that pFBA takes more come to calculate.

pFBA finds the optimal biomass growth and then minimizes the other fluxes to keep the same optimal biomass growth. We can see that the biomass function is the same for both FBA and pFBA,

In [3]:
```
print(' FBA Biomass Flux = ',fba_solution.fluxes.BIOMASS_Ec_iJO1366_core_53p95M)
print('pFBA Biomass Flux = ',pfba_solution.fluxes.BIOMASS_Ec_iJO1366_core_53p95M)
```

```
 FBA Biomass Flux =  0.9823718127269769
pFBA Biomass Flux =  0.9823718127269769
```

Flux Balance Analysis

- Overview
- Mathematical Representation of Reactions & Constraints
- Biomass Reaction
- Flux Balance Analysis
- Loopless Flux Balance Analysis
- Parsimonious Flux Balance Analysis

Learning Objectives

Each student should be able to:

- Explain flux balance analysis (FBA).
- Explain mass balanced linear equations.
- Explain the role of the biomass reaction.
- Explain how to create a stoichiometric matrix.
- Explain the purpose of loopless FBA
- Explain the purpose of parsimoneous FBA
- Demonstrate the ability to use Flux Balance Analysis to calculate metabolic fluxes with a COBRA model.

Reflective Questions

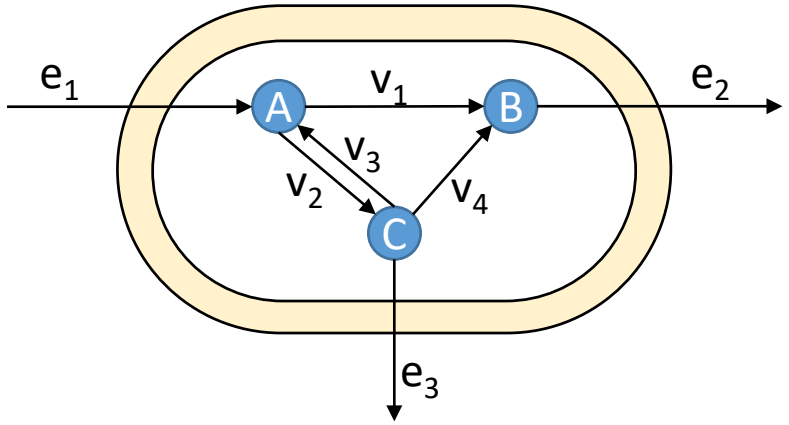
What is flux balance analysis?
What is a stoichiometric matrix?
What are the units of flux used in COBRA models?
What are some possible objective functions?
What is the purpose of the biomass reaction?
Why is the product of the stoichiometric matrix and the flux set to zero?
What is the purpose of linear programming?
What is the purpose on an objective function?
What are biomass precursors?
What is the purpose of growth associated maintenance (GAM)?
What is the purpose of non-growth associated maintenance energy (NGAM)?
What is the purpose of the biomass reaction?
What are reduced costs?
What are shadow prices?
What is the purpose of loopless flux balance analysis?
What is the purpose of parsimonious flux balance analysis?
What are the limits of FBA?

References

1. Lewis, N. E., H. Nagarajan, et al. (2012). "Constraining the metabolic genotype-phenotype relationship using a phylogeny of *in silico* methods." Nature reviews. Microbiology 10(4): 291-305.
2. Orth, J. D., I. Thiele, et al. (2010). "What is flux balance analysis?" Nature biotechnology 28(3): 245-248.
3. Becker, S. A., A. M. Feist, et al. (2007). "Quantitative prediction of cellular metabolism with constraint-based models: the COBRA Toolbox." Nature protocols 2(3): 727-738.
4. Reed, J. L., I. Famili, et al. (2006). "Towards multidimensional genome annotation." Nature reviews. Genetics 7(2): 130-141.
5. REI601M, Introduction to Systems Biology, Dr. Innes Thiele,2012, https://systemsbiology.hi.is/wiki/REI601M
6. Thiele, I. and B. O. Palsson(2010). "A protocol for generating a high-quality genome-scale metabolic reconstruction." Nature protocols 5(1): 93-121.
7. M. Schaechter, et al, Microbe, ASM Press, 2006, p. 116
8. Lewis, N. E., K. K. Hixson, et al. (2010). "Omic data from evolved *E. coli* are consistent with computed optimal growth from genome-scale models." Molecular Systems Biology 6: 390.

Dynamic Mass Balance

A simple network



Linear Transformation

$$\frac{d\mathbf{x}}{dt} = \mathbf{S} \cdot \mathbf{v}$$

$$\begin{bmatrix} \frac{dA}{dt} \\ \frac{dB}{dt} \\ \frac{dC}{dt} \end{bmatrix} = \begin{array}{c} \begin{matrix} v_1 & v_2 & v_3 & v_4 & e_1 & e_2 & e_3 \end{matrix} \\ \begin{bmatrix} -1 & -1 & 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & -1 & 0 \\ 0 & 1 & -1 & -1 & 0 & 0 & -1 \end{bmatrix} \end{array} \begin{bmatrix} v_1 \\ v_2 \\ v_3 \\ v_4 \\ e_1 \\ e_2 \\ e_3 \end{bmatrix}$$

Stoichiometric Matrix

Linear Differential Equations

$$\frac{dA}{dt} = -v_1 - v_2 + v_3 + e_1$$

$$\frac{dB}{dt} = v_1 + v_4 - e_2$$

$$\frac{dA}{dt} = v_2 - v_3 - v_4 - e_3$$

Dynamic Mass Balance (Steady State)

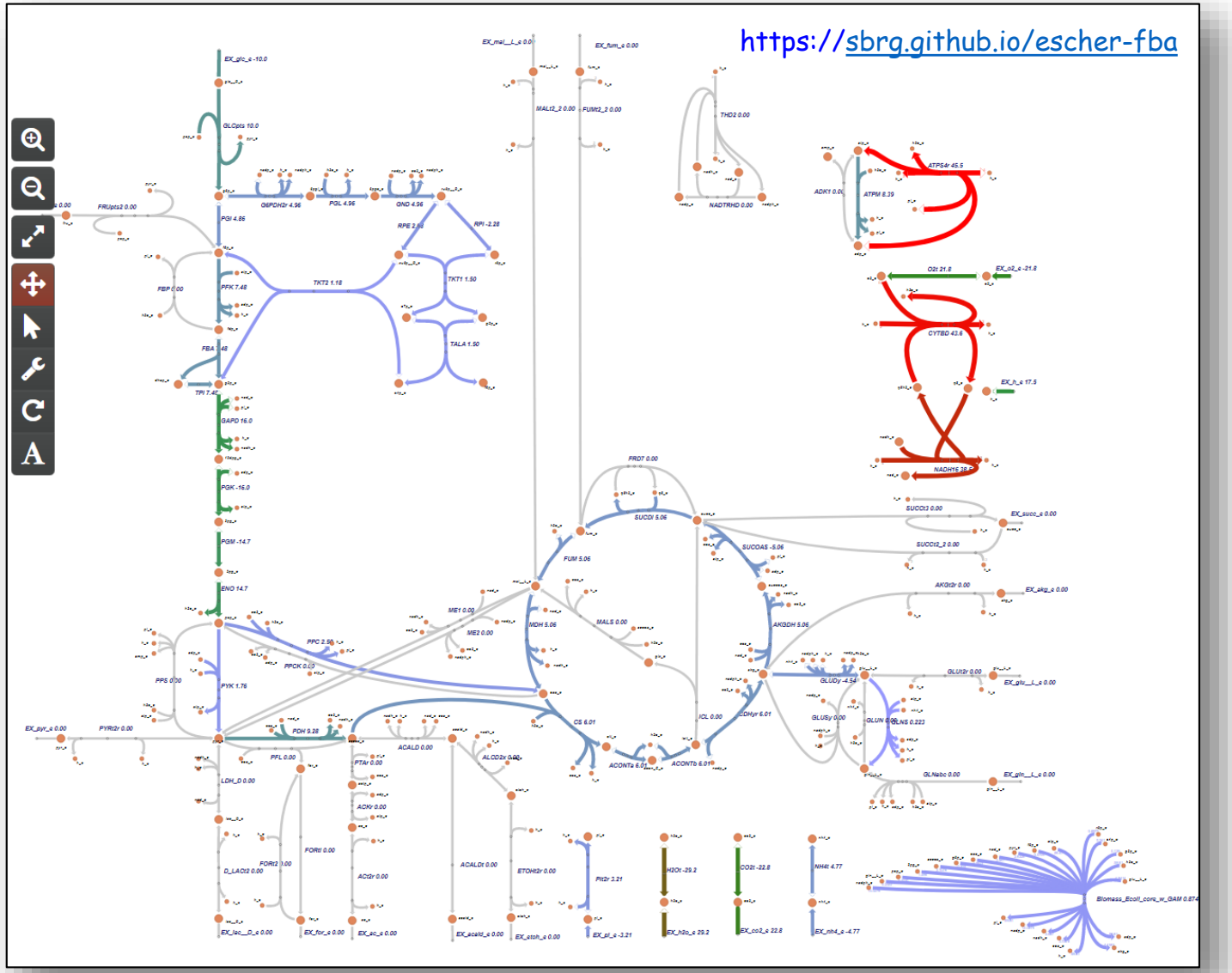
$$0 = \mathbf{S} \cdot \mathbf{v}$$

$$\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} -1 & -1 & 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & -1 & 0 \\ 0 & 1 & -1 & -1 & 0 & 0 & -1 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ v_3 \\ v_4 \\ e_1 \\ e_2 \\ e_3 \end{bmatrix}$$

Note: More unknown variables than equations, thus no unique solutions! Need constraints!

Escher-FBA

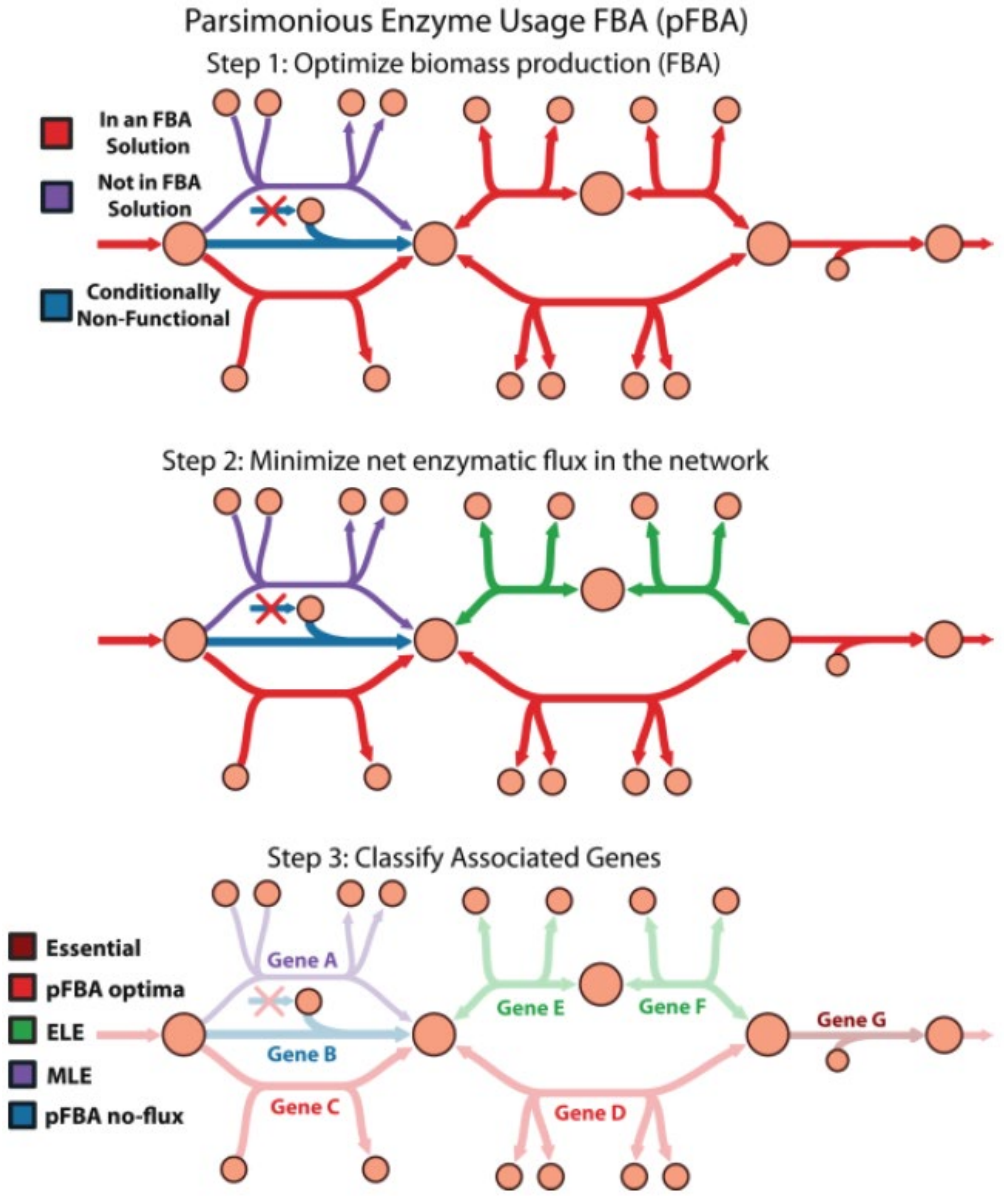
- A web application for interactive FBA simulations within a pathway visualization.
- Allows users to set flux bounds, knock out reactions, change objective functions, upload metabolic models, and generate high-quality figures without downloading software or writing code.
- Designed to replicate FBA simulations to generate scientific hypotheses.



EG Rowe, BO Palsson, ZA King (2018) Escher-FBA: a web application for interactive flux balance analysis. *BMC Sys Bio* 12:84 https://doi.org/10.1186/s12918-018-0607-5

Parsimonious Enzyme Usage

- Gene A, classified as MLE, represents an enzyme that uses a suboptimal co-factor to catalyze a reaction, thereby reducing the growth rate if used.
- Gene B, classified as pFBA no-flux, cannot carry a flux in this example since it is unable to take up or produce a necessary precursor metabolite.
- Genes E and F in this example require two different enzymes to catalyze the same transformation which Gene D can do alone; therefore they are classified as ELE.
- Gene G is essential, since its removal will stop the flux through all pathways.
- Genes C and D represent the most efficient (topologically and metabolically) pathway and therefore are part of the pFBA optima.



Lewis, N. E., K. K. Hixson, et al. (2010). "Omic data from evolved E. coli are consistent with computed optimal growth from genome-scale models." Molecular Systems Biology 6: 390.