Metabolic Models: From DNA to Physiology (and Back)

Marcio Argollo de Menezes

Abstract Metabolic reconstructions constitute translations from genomic data to biochemical processes and serve as valuable tools to assess, along with mathematical models, the viability of organisms on different environments or the overproduction of industrially valuable metabolites following controlled manipulation of specific reaction rates. In the following, we review FBA, a constraint-based mathematical method which successfully predicts genome-wide metabolic fluxes, most notably the rate of accumulation of biomass precursors with stoichiometry determined by the cellular biomass composition. The practical implementation of the method on a synthetic metabolic model is offered as computer codes written for GNU-Octave, an open-source language with powerful numerical tools.

Systems biology is an emerging research field which integrates information from very distinct, well-established areas to deal with the (rather puzzling) question, "What is life and what underlies its agency?" [30], so that the innumerous molecular structures and procedures encoded therein can be exploited for diverse purposes [10, 22, 29], from drug design and crop yield optimization [12, 27] to tissue remodelling and winemaking [15, 28].

This chapter is a practical introduction to metabolic modelling, where one tests the flux capabilities of a given map of metabolic reactions with a mathematical representation of the map and computational techniques that solve numerical problems associated with tests of hypotheses, inference of behaviors, and generation of predictions [14, 22]. In the first section, I introduce the basic ideas and fundamental biological discoveries behind metabolic modelling, from map design to mathematical modelling. Next, I will describe the mathematical formulation and practical implementation of flux balance analysis (FBA) [25], one of the most successful computational techniques for organism-wide prediction of metabolic reaction fluxes and the effects of their modulation. We will find the maximal growth rate and determine the essentiality of genes/reactions for non-zero biomass

M. A. de Menezes (✉)
Instituto de Física, Universidade Federal Fluminense (UFF), Niterói, Brazil
e-mail: marcio@mail.if.uff.br

F. A. B. da Silva et al. (eds.), *Theoretical and Applied Aspects of Systems Biology*, Computational Biology 27, https://doi.org/10.1007/978-3-319-74974-7_4

production [9, 12, 33, 36, 40] on an artificial metabolic model, with reactions derived from genes of a model cell and a pseudoreaction describing accumulation of biomass precursors. Numerical solutions will be implemented on *GNU Octave*, an open-source scripting language with many libraries for numerical calculus [6].

1 Metabolic Models

Advances on sequencing techniques and bioinformatics algorithms made it possible to reconstruct, from genomic sequence, the entire set of biochemical reactions and transport processes available to an organism [2, 5, 16, 24, 35].[1]

With this information one can build topological maps akin to metabolic pathways, where generation of a selected set of products, given a particular set of available substrates, is written in terms of reactions whose integrated interconversion of metabolites link the desired products to elements in the set of available substrates.

One can investigate, for instance, growth capacity in different environments, one of the central hypotheses behind the idea of life as self-replicating, autocatalytic sets [11, 34]: a cell must accumulate, from a basic food source, the set of metabolites which constitute its physical structure in amounts defined by the cellular composition, which is further (self) organized into a new (identical) cell [19]. This process can be incorporated in metabolic models in the form of a pseudoreaction (usually called biomass production reaction), where biomass precursor metabolites are the substrates with stoichiometric indices defined by their relative amounts in cellular composition [8].

Given enough precision in the reconstruction process [39], the physiological strategy for biomass generation, that is, the pathway design of metabolic reaction fluxes experimentally observed in living organisms [20, 38] should be one of the possible designs from the reconstructed map [8]. It is common sense to suppose that, under competition, evolution drives organisms toward maximization of fitness [31], which in simple prokaryotic organisms translates almost entirely to growth rate. This hypothesis is prone to be answered by mathematical and computational modelling [14, 26], which are very helpful tools, accelerating discovery and generating reproducible basic knowledge from biologically inspired hypotheses. Constraint-based analysis [3] is a mathematical methodology in which flux of metabolic reactions are predicted by systematic narrowing of the search space by the addition of biologically inspired constraints, the most notable being flux balance in freely dividing cells: when cells evolve with constant duplication time (constant growth rate), as expected in nutrient-rich media, their molecular composition remains unchanged after duplication. This steady state, easily reproducible in chemostats [21, 41], constrains metabolic reaction fluxes to values leading to balanced synthesis and consumption rates of every intracellular metabolite [17–19].

[1] Check http://systemsbiology.ucsd.edu/InSilicoOrganisms/OtherOrganisms or https://www.ebi.ac.uk/biomodels-main/ for an updated list.

Based on the premise that prokaryotes such as *Escherichia coli* have maximized their growth performance along evolution, flux balance analysis (FBA) [25] is used to predict the expected (physiological) metabolic phenotype of bacteria evolving in rich media (with constant growth rate) as the one, among all flux sets satisfying the stationarity constraint in the metabolic reconstruction, with the largest rate of biomass production (as noted by J. Monod [17], physical limits on uptake rates constrain growth rates to finite values). With simple mathematical formulation [26], FBA successfully connects cell's physiology to the capabilities of itś underlying metabolic network given flux constraints imposed by environmental nutrient composition and cellular state [7, 13, 23, 32].

2 FBA: Predicting Metabolic Phenotypes

Let's discuss the basic ideas behind FBA with its practical implementation on an artificial metabolic model containing $M = 7$ metabolites and $N = 12$ reactions (Fig. 1). Eleven reactions are gene-related (five occurring inside the cell and six transporting metabolites through cell boundary), and one is a pseudoreaction ($R12$, not gene-related) describing the accumulation of biomass precursor metabolites. Reversible reactions are split into the actual direct and reverse processes, and reaction fluxes are all positive-definite.

When reactions occur simultaneously in the intracellular medium, metabolite concentrations change in time as result of the difference between their rates of synthesis and consumption in all participating reactions. Writing the flux rate of reaction j as f_j and the stoichiometry of metabolite i in reaction j as S_{ij} (zero if not present in the reaction, negative if substrate and positive otherwise), the concentration of metabolite i evolves in time as

$$\frac{d\,[m_i]}{dt} = \sum_{j=1}^{N} S_{ij}\, f_j . \tag{1}$$

where $N = 12$ is the number of reactions and $M = 7$ the number of metabolites in the metabolic model. The stoichiometric matrix carries all the information contained

(R1,g1) m1_e → m1
(R2,g1) m1 → m1_e
(R3,g2) m2_e → m2
(R4,g2) m2 → m2_e
(R5,g3) m3_e → m3
(R6,g4) m4 → m4_e
(R7,g5) 2 m2+ m3 → m1 + m4
(R8,g6) m1+3 m3 → 2 m5
(R9,g7) m5 + m2 → m6
(R10,g7) m6 → m5 + m2
(R11,g8) m5 + 2 m6 → m7 + m1
(R12, ø) 0.3 m5 +0.54 m6 +0.16 m7 → Cell biomass

Fig. 1 Metabolic model from hypothetical cell, viewed as a list of reactions. Metabolites marked in red on the reactions list do not occur inside the cell

a)

$$\frac{d[m1]}{dt} = (1)f_1 + (-1)f_2 + (1)f_7 + (-1)f_8 + (1)f_{11}$$
$$\frac{d[m2]}{dt} = (1)f_3 + (-1)f_4 + (-2)f_7 + (-1)f_9 + (1)f_{10}$$
$$\frac{d[m3]}{dt} = (1)f_5 + (-1)f_7 + (-3)f_8$$
$$\frac{d[m4]}{dt} = (-1)f_6 + (1)f_7$$
$$\frac{d[m5]}{dt} = (2)f_8 + (-1)f_9 + (-1)f_{11} + (-0.3)f_{12}$$
$$\frac{d[m6]}{dt} = (1)f_9 + (-1)f_{10} + (-2)f_{11} + (-0.54)f_{12}$$
$$\frac{d[m7]}{dt} = (1)f_{11} + (-0.16)f_{12}$$

b)

$$\frac{d[mi]}{dt} = \sum_{j=1}^{N} S_{ij} f_j$$

$$S = \begin{pmatrix} 1 & -1 & 0 & 0 & 0 & 0 & 1 & -1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & -1 & 0 & 0 & -2 & 0 & -1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & -1 & -3 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & -1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 2 & -1 & 1 & -1 & -0.3 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & -1 & -2 & -0.54 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & -0.16 \end{pmatrix}$$

Fig. 2 Dynamics of metabolite concentrations, written in function of reaction fluxes

on the list of reactions and is depicted in Fig. 2, along with the differential equations describing time evolution of concentrations in our metabolic model.

Genome-wide metabolic reconstructions usually have hundreds to thousands of metabolites and reactions, and the solution of this set of coupled differential equations becomes unpractical. Nevertheless, in organisms growing in the stationary phase intracellular metabolite concentrations do not change with time, reducing Eq. 1 to

$$\sum_{j=1}^{N} S_{ij} f_j = 0 \quad \forall\, i \tag{2}$$

that can be written in a compact notation as

$$\mathbf{S}\vec{f} = \vec{0} \tag{3}$$

Equation 2 describes a set of coupled linear equations on fluxes, much simpler to solve than the set of coupled differential equations defined in (1). As there are more reactions than metabolites, multiple solutions exist for the problem [26], reflecting the multitude of strategies inscribed in metabolic networks.[2] This degeneracy can be lifted by the introduction of more constraints to solution. As stated in the introduction, prokaryotic cells growing in rich media should evolve toward maximization of growth rate. One can formulate this problem mathematically as

$$\begin{aligned} &\text{MAX}\{f_{\text{bio}}\} \\ &\text{Given} \sum_{j=1}^{N} S_{ij} f_j = 0 \quad \forall\, i \end{aligned} \tag{4}$$

[2]Many organisms can, for instance, generate ATP either by respiration, fermentation, or both processes simultaneously [37].

which is a very popular problem in mathematics called linear programming [4] for which one finds public libraries implementing its solution with diverse algorithms. We choose to expose our examples in GNU Octave [6], an open-source scripting language with simple syntax and many libraries that solve a vast range of mathematical problems. It comes with an environment, where one can type commands that are interpreted on-the-fly.

2.1 Growth Prediction

To define the stoichiometric matrix of our metabolic model, just type, in the octave environment,

```
octave:1> A=[1 -1 0  0 0  0  1 -1  0  0  1  0;
             0  0 1 -1 0  0 -2  0 -1  1  0  0;
             0  0 0 0  1  0 -1 -3  0  0  0  0;
             0  0 0 0  0 -1  1  0  0  0  0  0;
             0  0 0 0  0  0  0  2 -1  1 -1 -0.3;
             0  0 0 0  0  0  0  0  1 -1 -2 -0.54;
             0  0 0 0  0  0  0  0  0  0  1 -0.16];
```

To find the maximum growth rate of our model cell, we use the linear programming library *glpk*. From the help function

```
octave:1> help glpk
 -- Function File: [XOPT, FMIN, ERRNUM, EXTRA] = glpk (C, A, B,
    LB, UB,
          CTYPE, VARTYPE, SENSE, PARAM)
     Solve a linear program using the GNU GLPK library.

     Given three arguments, 'glpk' solves the following
     standard LP:
          min C'*x
      subject to
          A*x  = b
            x >= 0

     Input arguments:
     C
          A column array containing the objective function
          coefficients.
     A
          A matrix containing the constraints coefficients.
     B
          A column array containing the right-hand side value
          for each constraint in the constraint matrix.
     LB
          An array containing the lower bound on each of the
          variables. If LB is not supplied, the default lower
          bound for the variables is zero.
```

```
UB
     An array containing the upper bound on each of the
     variables. If UB is not supplied, the default upper
     bound is assumed to be infinite.
```

Since we want to maximize growth rate, the vector $\vec{C}$ must have a single non-zero element, C [12], which we set to -1 to reflect the maximization of flux in the biomass production reaction $R12$.

```
octave:2> c = zeros(12,1);
octave:3> c(12)=-1;
```

The lower bound of all reactions is zero, and upper bounds, given no additional information on maximum reaction rates, are set to an arbitrary value (1 in our case).

```
octave:4> lb = zeros(12,1);
octave:5> ub = ones(12,1);
```

As no intracellular metabolite accumulates in time, all components of $\vec{B}$ are set to zero.

```
octave:6> B = zeros(7,1);
```

After setting all input parameters, *glpk* is evoked:

```
octave:7> [x0, FMIN, ERRNUM] = glpk(c,A,B,lb,ub);
```

If $ERRNUM = 0$, a valid solution is found, and the vector $\vec{X0}$ is returned with the respective reaction fluxes. In our case, $FMIN$ returns the growth rate. Typing the variable name in octave environment, one obtains its value

```
octave:8> ERRNUM
ERRNUM = 0
octave:9> x0
x0 =

   0.25253
   0.00000
   0.43434
   0.00000
   1.00000
  -0.00000
   0.00000
   0.33333
   0.43434
   0.00000
   0.08081
   0.50505
```

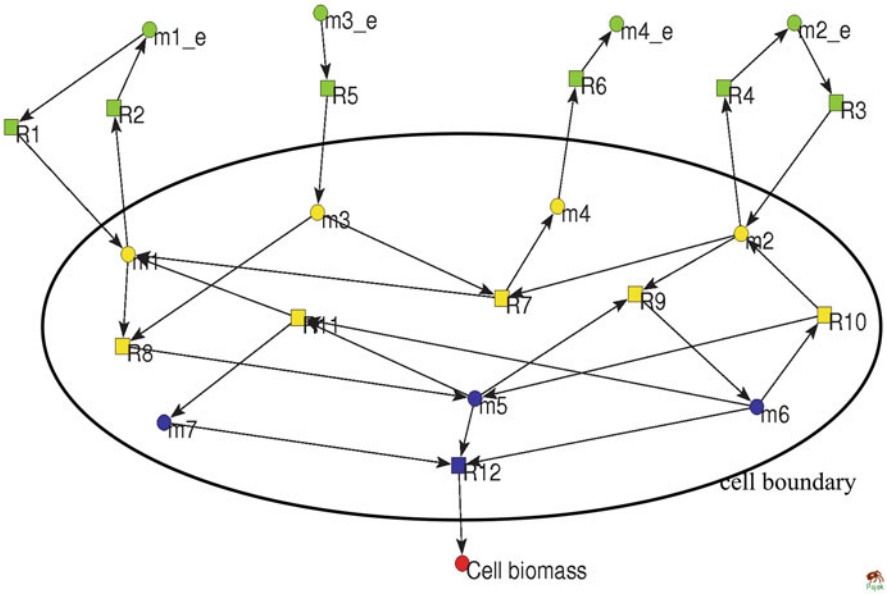

Fig. 3 Metabolic model from hypothetical cell viewed as a network. Metabolites marked in red on the reactions list do not occur inside the cell. In the network representation, metabolites and reactions are nodes (circles and squares, respectively), and directed links connect reaction substrates to their respective reactions and reactions to their products. Green nodes mark external metabolites, and transport reactions and biomass precursors are marked in blue

In order to visualize pathways involved in the strategy of optimal growth, we describe the metabolic model as a network (Fig. 3), with links connecting substrates to reactions and reactions to products. This representation evidences the molecular approach of physiology [19] in which growth is sustained by the uptake and sequential transformations of a small set of metabolites comprising the food source.

Since metabolite $m1$ is synthesized by an internal reaction, there should be another solution for the above problem given a medium without $m1$. In fact, if we set the upper bound of its uptake reaction to zero, we find another strategy for biomass generation (with smaller yield).

```
octave:10> ub(1,1)=0;
octave:11> [x0, FMIN, ERRNUM] = glpk(c,A,r,lb,ub);
octave:12> x0
x0 =

   0.00000
   0.00000
   0.75000
   0.00000
   1.00000
```

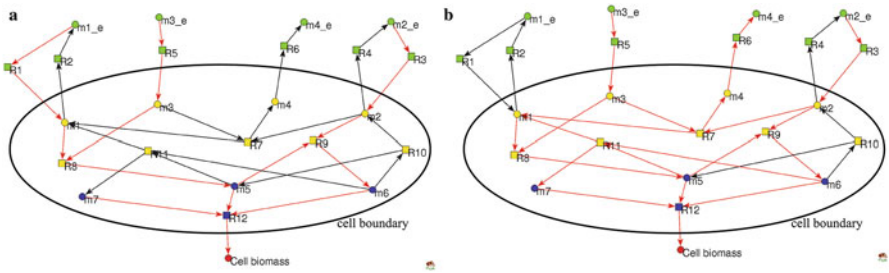

Fig. 4 Different strategies for biomass generation with metabolite $m1$ available as food source (left) and internally synthesized (right). (**a**) Pathway 1: Metabolite m1 available for uptake. (**b**) Pathway 2: Metabolite m1 not available for uptake

```
0.20161
0.20161
0.26613
0.34677
0.00000
0.06452
0.40323
```

The different pathways leading to biomass production are shown in Fig. 4, where active routes (links) are marked in red. In the left network, metabolite $m1$ is provided as a food source, while on the right network, it is not provided as a food source, but synthesized in the internal reaction $R7$ along with the by-product metabolite $m4$. This situation is analogous to the different pathways leading to ATP production in some organisms where the availability of oxygen in the environment determines whether respiration or fermentation takes place, with excretion of by-products evidencing the latter strategy [1].

2.2 *Gene Essentiality*

As stated previously, metabolic reactions are activated by the promotion of genes [1], and one can predict the essentiality of either reactions or genes for an organisms life by testing its capacity of producing biomass [33] in the metabolic reconstruction with upper bound for the selected reactions set to zero. Octave commands can be sequentially given to the interpreter as a script code. We let as a final exercise the interpretation of the code given below which output the critical genes which, when shut down, precludes biomass formation

```
marcio@sumbawa:~/cursos/fiocruz/redes metabolicas 2017/
book_chapter/codes$ ./fba1.m
######## Gene deletion studies ##############################
Gene g2 [R3][R4] is critical
```

```
Gene g3 [R5] is critical
Gene g6 [R8] is critical
Gene g7 [R9][R10] is critical
Gene g8 [R11] is critical
marcio@sumbawa:~/cursos/fiocruz/redes metabolicas 2017/
book_chapter/codes$
```

2.3 Octave Code with FBA Analysis

```
#!/usr/bin/octave
printf("######## Metabolic reactions of model cell ########\n");
printf("# (R1, gene g1)     --> m1                          #\n");
printf("# (R2, gene g1)  m1 -->                            #\n");
printf("# (R3, gene g2)     --> m2                          #\n");
printf("# (R4, gene g2)  m2 -->                            #\n");
printf("# (R5, gene g3)    --> m3                          #\n");
printf("# (R6, gene g4)  m4 -->                            #\n");
printf("# (R7, gene g5)  2 m2 + m3 --> m1 + m4             #\n");
printf("# (R8, gene g6)  m1 + 3 m3 --> 2 m5                #\n");
printf("# (R9, gene g7)  m5 + m2 --> m6                    #\n");
printf("# (R10, gene g7)  m6 --> m5 + m2                   #\n");
printf("# (R11, gene g8) m5 + 2 m6 --> m7 + m1             #\n");
printf("# (R12, no gene) 0.3 m5 + 0.54 m6 + 0.16 m7 -->    #\n");
printf("#####################################################\n");
printf("#R12 is a pseudo-reaction describing accumulation #\n");
printf("#of biomass precursors in proportions defined by  #\n");
printf("#cellular composition. Its flux mimics growth rate#\n");
printf("#####################################################\n");

A=[1 -1 0  0 0  0  1 -1  0  0  1  0
   0  0 1 -1 0  0 -2  0 -1  1  0  0
   0  0 0 0  1  0 -1 -3  0  0  0  0
   0  0 0 0  0 -1  1  0  0  0  0  0
   0  0 0 0  0  0  0  2 -1  1 -1 -0.3
   0  0 0 0  0  0  0  0  1 -1 -2 -0.54
   0  0 0 0  0  0  0  0  0  0  1 -0.16];

biomass_reaction = 12;
a_uptake = 1;
b_uptake = 4;
c_uptake = 5;
d_excretion = 6;

[M,N]=size(A);
lb = zeros(N,1);
ub = ones(N,1);
r = zeros(M,1);
c = zeros(N,1);
c(biomass_reaction)=-1; # Maximization of biomass production
flux
```

```
[x0, FMIN, STATUS] = glpk(c,A,r,lb,ub);
if(STATUS==0)
  max_biomass=x0(biomass_reaction);
  printf("Maximum biomass production flux (pathway 1)=%f\n",x0
  (biomass_reaction));
  for i=1:N
    printf("Rxn %d flux=%f\n",i,x0(i));
  endfor
endif

ub(1,1)=0;
[x0, FMIN, STATUS] = glpk(c,A,r,lb,ub);
if(STATUS==0)
  max_biomass=x0(biomass_reaction);
  printf("Maximum biomass production flux (pathway 2)=%f\n",x0
  (biomass_reaction));
  for i=1:N
    printf("Rxn %d flux=%f\n",i,x0(i));
  endfor
endif

printf("######## Gene deletion studies ##############\n");
# (R1, gene g1)     --> A
# (R2, gene g1)  A -->
# (R3, gene g2)     --> B
# (R4, gene g2)  B -->
# (R5, gene g3)    --> C
# (R6, gene g4)  D -->
# (R7, gene g5)  2B + C --> A + D
# (R8, gene g6)  A + 3C --> 2X
# (R9, gene g7)  X + B --> Y
# (R10, gene g7)  Y --> X + B
# (R11, gene g8) X + 2Y --> Z + A
ngenes=8;
nrxns_gene(1)=2;
nrxns_gene(2)=2;
nrxns_gene(3)=1;
nrxns_gene(4)=1;
nrxns_gene(5)=1;
nrxns_gene(6)=1;
nrxns_gene(7)=2;
nrxns_gene(8)=1;
gene_rxn(1,1)=1;
gene_rxn(1,2)=2;
gene_rxn(2,1)=3;
gene_rxn(2,2)=4;
gene_rxn(3,1)=5;
gene_rxn(4,1)=6;
gene_rxn(5,1)=7;
gene_rxn(6,1)=8;
gene_rxn(7,1)=9;
```

```
gene_rxn(7,2)=10;
gene_rxn(8,1)=11;


c = zeros(N,1);
c(biomass_reaction)=-1;
lb = zeros(N,1);
ub = ones(N,1);
for i=1:ngenes
  for j=1:nrxns_gene(i)
    k=gene_rxn(i,j);
    old_ub(k,1)=ub(k,1);
    old_lb(k,1)=lb(k,1);
    ub(k,1)=0;
    lb(k,1)=0;
  endfor
  [x0, FMIN, STATUS] = glpk(c,A,r,lb,ub);
  if(x0(biomass_reaction)<1e-6)
    printf("Gene g%d ",i);
    for j=1:nrxns_gene(i)
      k=gene_rxn(i,j);
      printf("[R%d]",k);
    endfor
    printf(" is critical\n");

    endif
  for j=1:nrxns_gene(i)
    k=gene_rxn(i,j);
    ub(k,1)=old_ub(k,1);
    lb(k,1)=old_lb(k,1);
  endfor
endfor
```

References

1. Alberts B, Johnson A, Lewis J, Morgan D, Raff M, Roberts K, Walter P. Molecular biology of the cell. 500 Tips. New York: Garland Science; 2014.
2. Bartell JA, Blazier AS, Yen P, Thgersen JC, Jelsbak L, Goldberg JB, Papin JA. Reconstruction of the metabolic network of Pseudomonas aeruginosa to interrogate virulence factor synthesis. Nat Commun. 2017;8:14631 EP. Article.
3. Bordbar A, Monk JM, King ZA, Palsson BO. Constraint-based models predict metabolic and associated cellular functions. Nat Rev Genet. 2014;15:107 EP. Review Article.
4. Cormen TH, Leiserson CE, Rivest RL, Stein C. Introduction to algorithms. MIT electrical engineering and computer science series. Cambridge: MIT Press; 2001.
5. Dias O, Rocha M, Ferreira EC, Rocha I. Reconstructing genome-scale metabolic models with merlin. Nucleic Acids Res. 2015l43(8):3899–910. 25845595[pmid].
6. Eaton JW, Bateman D, Hauberg S. GNU octave version 4.2.2 manual: a high-level interactive language for numerical computations. https://www.gnu.org/software/octave/doc/v4.2.2.

7. Feist AM, Palsson BO. The growing scope of applications of genome-scale metabolic reconstructions using escherichia coli. Nat Biotechnol. 2008;26:659–67.
8. Feist AM, Palsson BO. The biomass objective function. Curr Opin Microbiol. 2010;13(3):344–49. 20430689[pmid].
9. Joyce AR, Palsson B. Predicting gene essentiality using genome-scale in silico models. In: Osterman AL, Gerdes SY, editors. Microbial gene essentiality: protocols and bioinformatics. Totowa: Humana Press; 2008. p. 433–57.
10. Karr JR, Sanghvi JC, Macklin DN, Gutschow MV, Jacobs JM, Bolival B, Assad-Garcia N, Glass JI, Covert MW. A whole-cell computational model predicts phenotype from genotype. Cell. 2012;150:389–401.
11. Kauffman SA. Autocatalytic sets of proteins. J Theor Biol. 1986;119(1):1–24.
12. Kohanski MA, Dwyer DJ, Collins JJ. How antibiotics kill bacteria: from targets to networks. Nat Rev Microbiol. 2010;8(6):423–35. 20440275[pmid].
13. Mahadevan R, Palsson B, Lovley DR. In situ to in silico and back: elucidating the physiology and ecology of Geobacter spp. using genome-scale modelling. Nat Rev Microbiol. 2011;9:222 EP, Erratum.
14. Maranas CD, Zomorrodi AR. Optimization methods in metabolic networks. Hoboken: Wiley; 2016.
15. Mendoza SN, Can PM, Contreras N, Ribbeck M, Agosn E. Genome-scale reconstruction of the metabolic network in oenococcus oeni to assess wine malolactic fermentation. Front Microbiol. 2017;8:534. 28424673[pmid].
16. Monk J, Nogales J, Palsson BO. Optimizing genome-scale network reconstructions. Nat Biotechnol. 2014;32:447 EP.
17. Monod J. The growth of bacterial cultures. Annu Rev Microbiol. 1949;3(1):371–94.
18. Monod J. Recherches sur la croissance des cultures bactériennes. Actualités scientifiques et industrielles. Hermann; 1958.
19. Neidhardt FC. Bacterial growth: constant obsession with dN/dt. J Bacteriol. 1999;181(24):7405–08. 1365[PII].
20. Niedenfhr S, Wiechert W, Katharina NH. How to measure metabolic fluxes: a taxonomic guide for 13c fluxomics. Curr Opin Biotechnol. 2015;34(Supplement C):82–90. Systems biology Nanobiotechnology.
21. Novick A, Szilard L. Description of the chemostat. Science. 1950;112(2920):715–6.
22. O'Brien EJ, Monk JM, Palsson BO. Using genome-scale models to predict biological capabilities. Cell. 2015;161:971–87.
23. Oberhardt MA, Palsson BO, Papin JA. Applications of genome-scale metabolic reconstructions. Mol Syst Biol. 2009;5:1–15.
24. Oberhardt MA, Puchaka J, Martins dos Santos VAP, Papin JA. Reconciliation of genome-scale metabolic reconstructions for comparative systems analysis. PLoS Comput Biol. 2011;7(3):e1001116. 10-PLCB-RA-2544R2[PII].
25. Orth JD, Thiele I, Palsson B. What is flux balance analysis? Nat Biotechnol. 2010;28:245.
26. Palsson BØ. Systems biology: properties of reconstructed networks. Cambridge: Cambridge University Press; 2006.
27. Peyraud R, Dubiella U, Barbacci A, Genin S, Raffaele S, Roby D. Advances on plantpathogen interactions from molecular toward systems biology perspectives. Plant J. 2017;90(4):720–37.
28. Rajagopalan P, Kasif S, Murali TM. Systems biology characterization of engineered tissues. Annu Rev Biomed Eng. 2013;15(1):55–70.
29. Santos FB, Vos WM, Teusink B. Towards metagenome-scale models for industrial applicationsthe case of lactic acid bacteria. Curr Opin Biotechnol. 2013;24:200–6.
30. Schrödinger E. What is life? with mind and matter and autobiographical sketches. Cambridge paperback library. Cambridge: Cambridge University Press; 1992.
31. Shakiba N, Zandstra PW. Engineering cell fitness: lessons for regenerative medicine. Curr Opin Biotechnol. 2017;47(Supplement C):7–15. Tissue, cell and pathway engineering.

32. Smith CA, Neidhardt FC, Ingraham JL, Schaechter M. Physiology of the bacterial cell: a molecular approach. Sunderland: Sinauer Associates; 1990. p. 507; 43:95. ISBN: 0878936084; 2010;20:124–5.
33. Snitkin ES, Dudley AM, Janse DM, Wong K, Church GM, Segr D. Model-driven analysis of experimentally determined growth phenotypes for 465 yeast gene deletion mutants under 16 different conditions. Genome Biol. 2008;9(9):R140.
34. Sousa FL, Hordijk W, Steel M, Martin WF. Autocatalytic sets in E. coli metabolism. J Syst Chem. 2015;6(1):4. 9[PII].
35. Thiele I, Palsson B. A protocol for generating a high-quality genome-scale metabolic reconstruction. Nat Protoc. 2010;5:93–121.
36. Tobalina L, Pey J, Rezola A, Planes FJ. Assessment of FBA based gene essentiality analysis in cancer with a fast context-specific network reconstruction method. PLoS One. 2016;11(5):e0154583. PONE-D-15-35442[PII].
37. Vazquez A. Overflow metabolism: from yeast to marathon runners. Saint Louis: Elsevier Science; 2017.
38. Vinaixa M, Rodrguez MA, Aivio S, Capellades J, Gmez J, Canyellas N, vis Stracker TH, Yanes O. Positional enrichment by proton analysis (pepa): a one-dimensional 1h-nmr approach for 13c stable isotope tracer studies in metabolomi cs. Angew Chem Int Ed. 2017;56(13):3531–5.
39. Walsh JR, Schaeffer ML, Zhang P, Rhee SY, Dickerson JA, Sen TZ. The quality of metabolic pathway resources depends on initial enzymatic function assignments: a case for maize. BMC Syst Biol. 2016;10:129. 369[PII].
40. Xavier JC, Patil KR, Rocha I. Integration of biomass formulations of genome-scale metabolic models with experimental data reveals universally essential cofactors in prokaryotes. Metab Eng. 2017;39:200–8.
41. Ziv N, Brandt NJ, Gresham D. The use of chemostats in microbial systems biology. J Vis Exp. 2013;14(80):50168. 50168[PII].