Grundlagen der Systembiologie und der Modellierung epigenetischer Prozesse

Sonja J. Prohaska

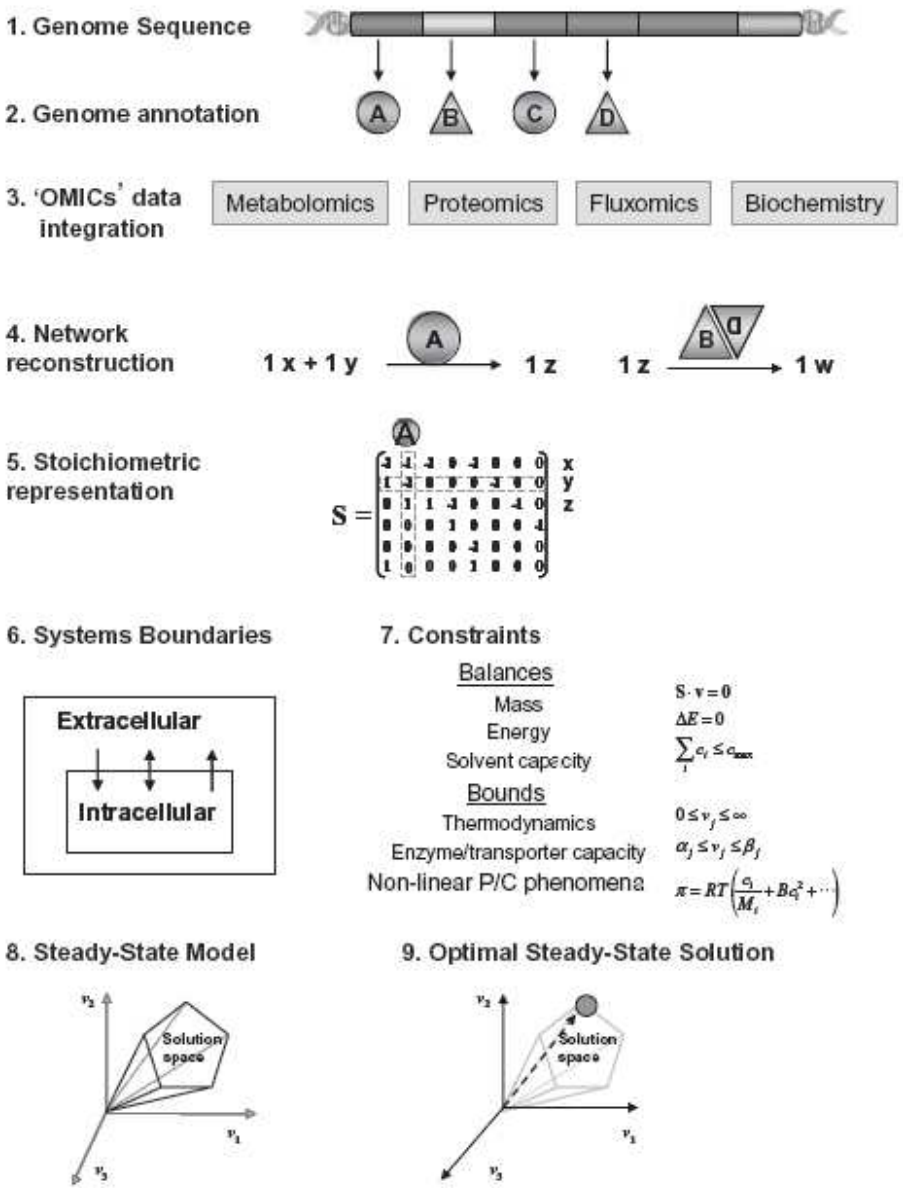
Bioinformatics Group
Institute of Computer Science
University of Leipzig

October 25, 2010

Genome-scale *in silico* Model

- functional *-omics* = annotating *-omics* data
- integrating *-omics* data of different kinds = systems biology
- Represent biological systems by networks.
- "*-omics*" data provide information about network components and their interactions.

Genome-scale *in silico* Model



The Cell as a Chemical Reaction Network

3. get the nodes of the network and integrate
   - (functional) genomics $\rightarrow$ all functional elements (mainly protein genes) that could be found and annotated in the genome
   - metabolomics $\rightarrow$ all metabolites present in a cell (substrates, cofactors, byproducts, etc. of chemical reactions)
   - proteomics $\rightarrow$ all structural proteins and enzymes (catalysts of chemical reactions) present in a cell
   - fluxomics $\rightarrow$ flows and reaction rates of all chemical reactions in a cell

The Cell as a Chemical Reaction Network

4. get the edges of the network by representing the chemical reactions
   - allow chemical reactions forming or breaking covalent bonds
   - allow chemical reactions that cause association or dissociation of molecules
5. get the stoichiometry of the chemical reactions right
   - balence atom composition (and mass)
   - invariant between organisms, independent of changes to conditions
- get the thermodynamics of the chemical reactions right
  - balence energy, derive relative rates of reactions
  - dependent on changes to (physiological) conditions
  - sequence alteration in binding surfaces can alter the thermodynamics of molecule association in different species
- get direction and absolute rate of reactions
  - determined by enzymes and their activity

The Cell as a Chemical Reaction Network

Stady-State Networks

- biological systems exist in a stady state (rather than in equilibrium)

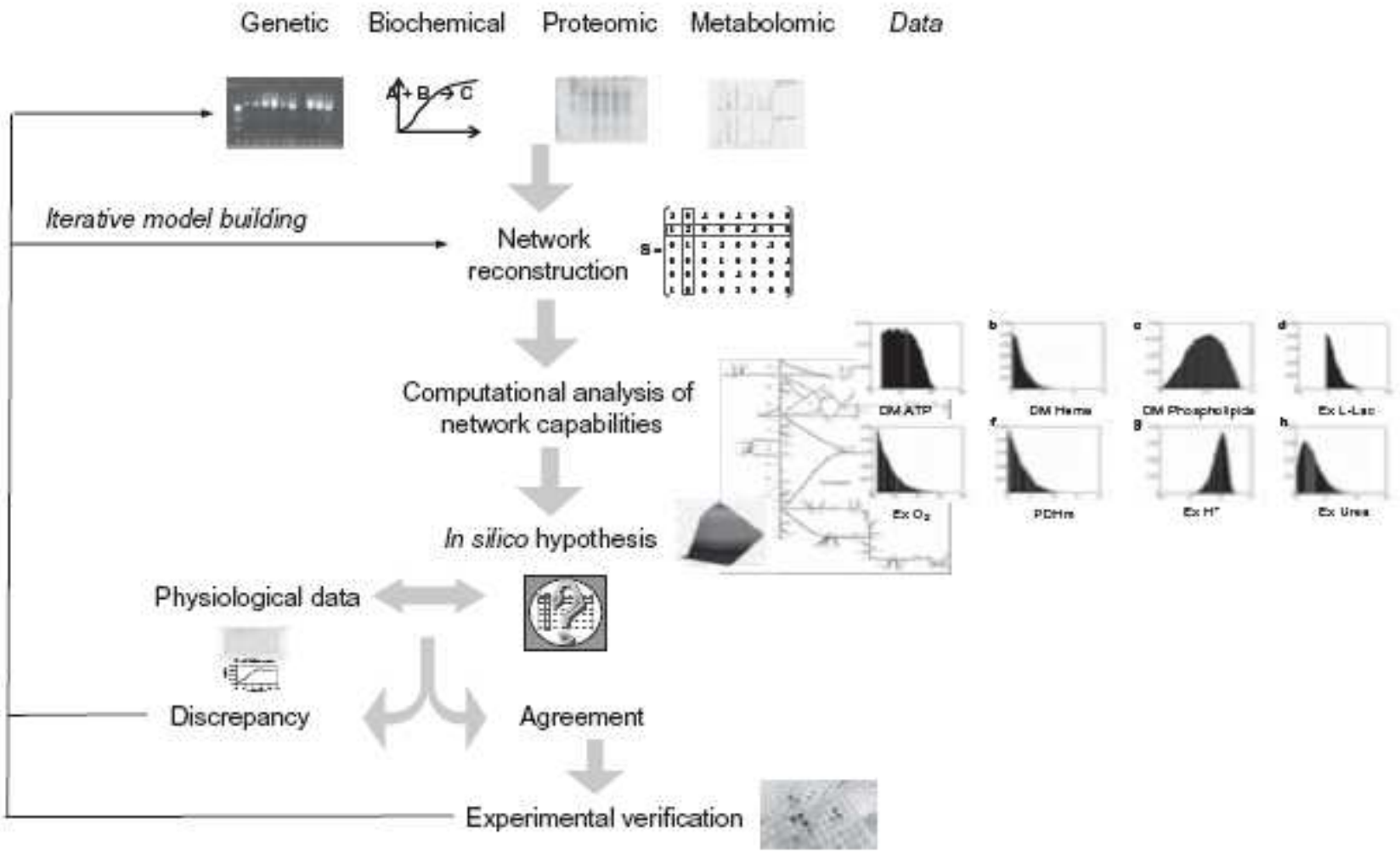
6. boundaries for (Sub-)systems need to be defined

8. a network is in stady-state if the in-flow is equal to the out-flow (i.e. no accumulation or depletion of molecules occurs)

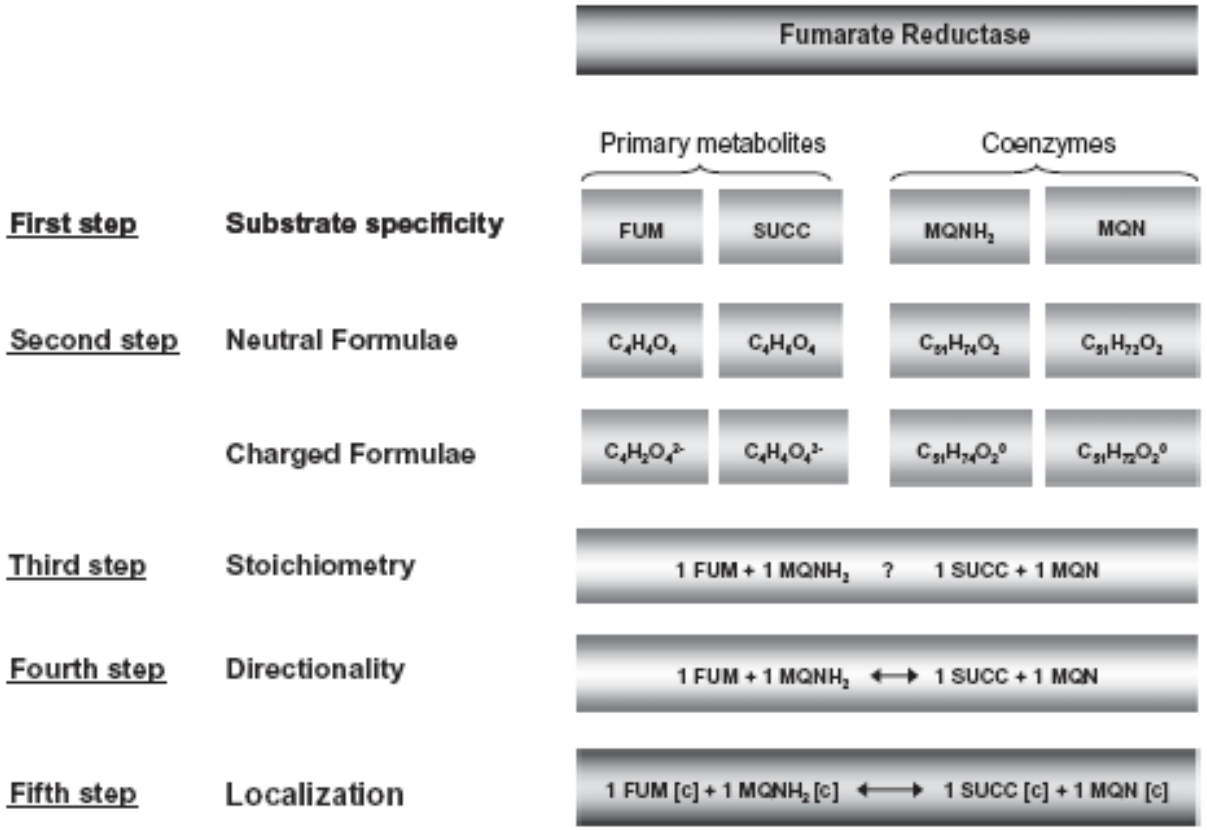
The Interative Process of Network Reconstruction

1. identify relevant metabolic genes from genome annotation
2. translate gene functions into balanced chemical reactions
3. network assembly from individual reactions
4. problem of incomplete data: fill in missing reactions to satisfy stady-state assumption
5. test the model *in silico* and compare results with physiological data
6. use gene essentiality date to validate reconstruction
7. refine interatively

The Interative Process of Network Reconstruction



Formulating Biochemical Reactions



Prokaryotes:
- extracellular space
- cytoplasm
- periplasm

Eukaryotes:
- extracellular space
- cytoplasm
- periplasm
- nucleus
- mitochondria
- peroxisome
- lysosome
- vacuole
- Golgi apparatus
- endoplasmatic reticulum

Constraint-based Modelling Approach

stoichiometric matrix

$$S_{mv} = \begin{pmatrix} 0 & 0 & 1 \\ 0 & -1 & -1 \\ 0 & -1 & 0 \\ -1 & 1 & 1 \\ 0 & 0 & -1 \\ 1 & 0 & 0 \end{pmatrix}$$

- while m are the metabolites, v are the fluxes/reactions
- a stoichiometric matrix S transforms the flux vector $v = (v_1, v_2, ..., v_n)$ into a vector of time derivates of the concentration vector $x = (x_1, x_2, ..., x_n)$
- $\frac{dx}{dt} = Sv$, stady state balance $Sv = 0$
- $\frac{dx_i}{dt} = \sum_k S_{ik} v_k$ is the sum of all fluxes producing or consuming x_i

Constraint-based Modelling Approach

stoichiometric matrix

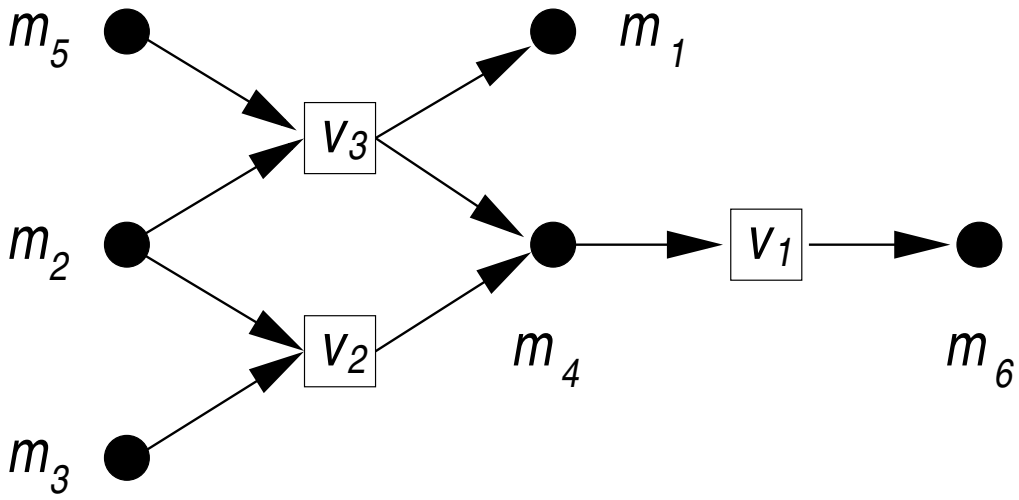
$$S_{mv} = \begin{pmatrix} 0 & 0 & 1 \\ 0 & -1 & -1 \\ 0 & -1 & 0 \\ -1 & 1 & 1 \\ 0 & 0 & -1 \\ 1 & 0 & 0 \end{pmatrix}$$

$m_1 = c_1; m_2 = c_2; m_3 = c_3; m_4 = c_2 c_3; m_5 = c_1 c_3; m_6 = c_3 c_2$

- reversible convertion: $c_2 c_3 \xrightarrow{v_1} c_3 c_2$
- bi-molecular association: $c_s + c_3 \xrightarrow{v_2} c_2 c_3$
- cofactor-coupled reaction: $c_2 + c_1 c_3 \xrightarrow{v_3} c_2 c_3 + c_1$ with c_1 as co-factor

Constraint-based Modelling Approach

network representation



Constraint-based Modelling Approach

- physiochemical constraints (inviolable)
  - mass, energy and momentum conserved
  - slow diffusion of macromolecules in viscous medium
  - reaction rates determined by local consentrations
  - reactions proceed in the direction of negative free-energy change
- spatial contraints
  - transport, structures
  - e.g. length, packaging and accessibility constrain arrangement of DNA
- environmental contraints
  - e.g. nutrient availability, temperature and osmolarity
  - important to determine phenotypic properties and fitness
- regulatory (self-imposed) constraints
  - allow the cell to eliminate suboptimal phenotypic states
  - e.g. transcriptional, translational, enzyme activity regulation

Given the Network...

- sample the network and study network properties
  - population of the flux space
  - interdipendencies and complexity
  - robustness to disturbance
  - flexibility to adopt to changing environments
- given an objective function linear optimization or linear programming can be used to calculate one optimal reaction network state (e.g. optimal growth)
- in large, more interconected networks alternative optima can be examined with mixed-integer LP algorithms
- optimize overproduction of a product: simultaneously optimize growth and secrection of the target product by (multiple) gene deletion.

References


[Choi, 2007] Sangdun Choi. *Introduction to Systems Biology*.


[Kaneko, 2006] Kunihiko Kaneko. *Life: An Introduction to Complex Systems Biology*.