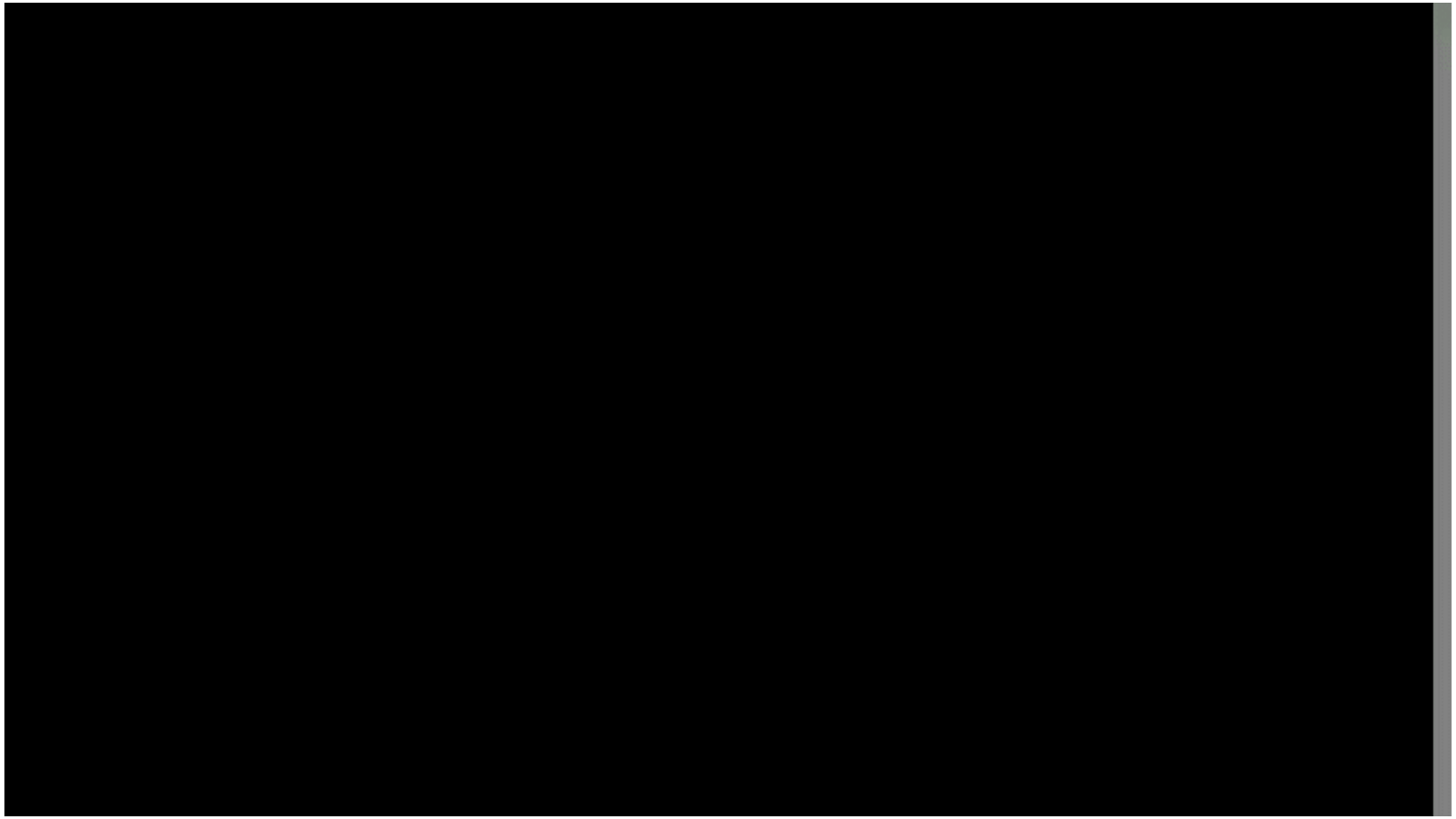
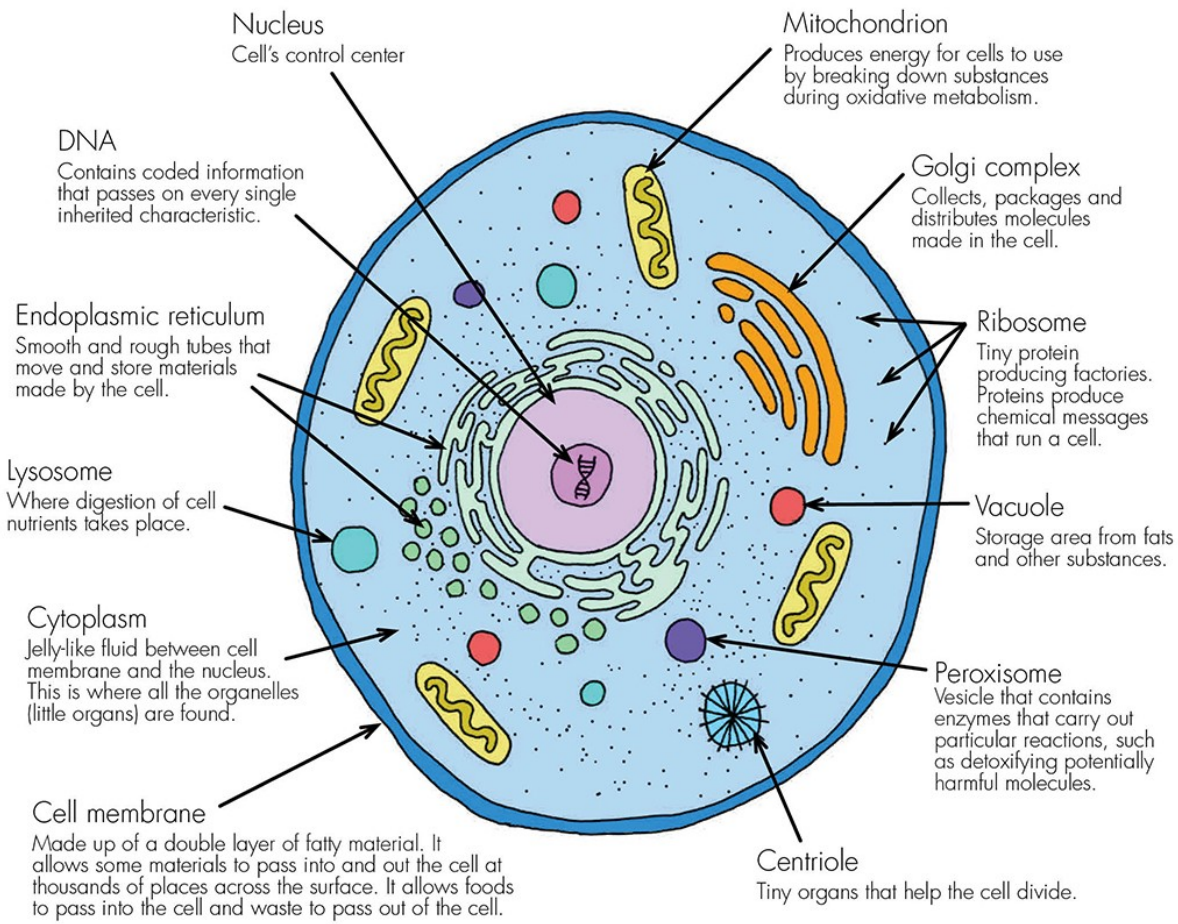
# Introduction to Systems Biology

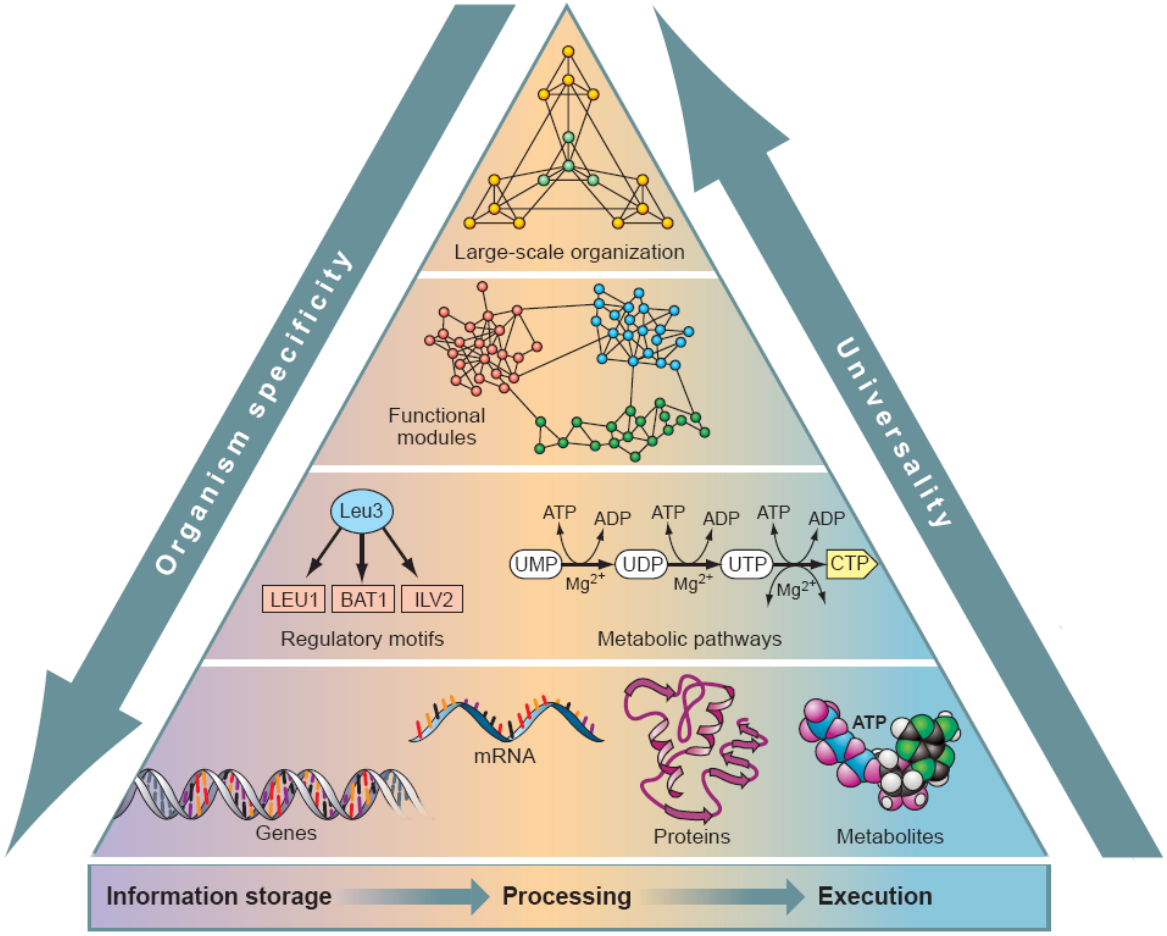
James Gomes
KSBS, IITD

# Inner Life of a Cell



# Cells Exhibit a High Degree of Complexity



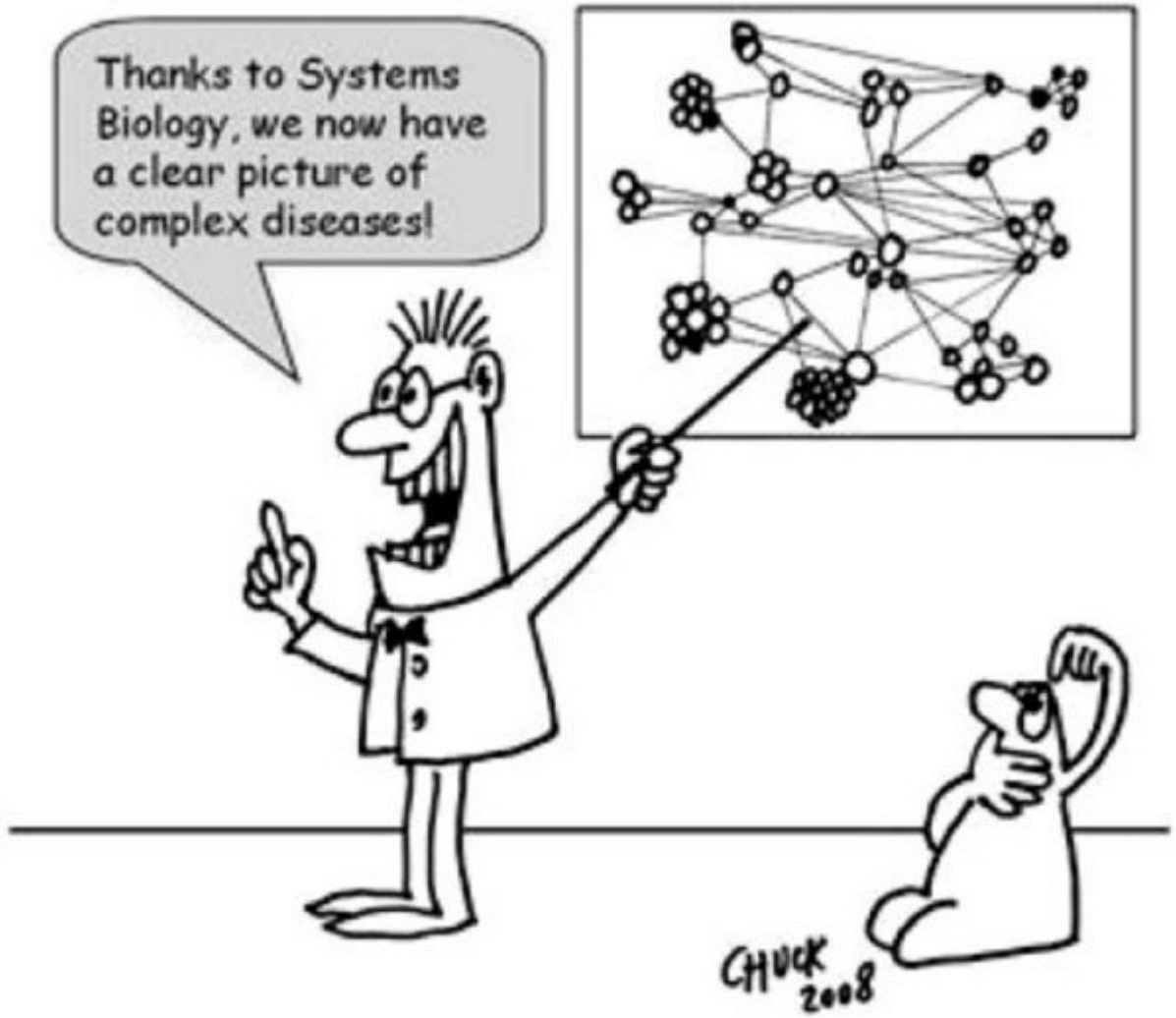




**What are the different types of cells?**

**Is "Life" an emergent property?**

# What is Systems Biology?

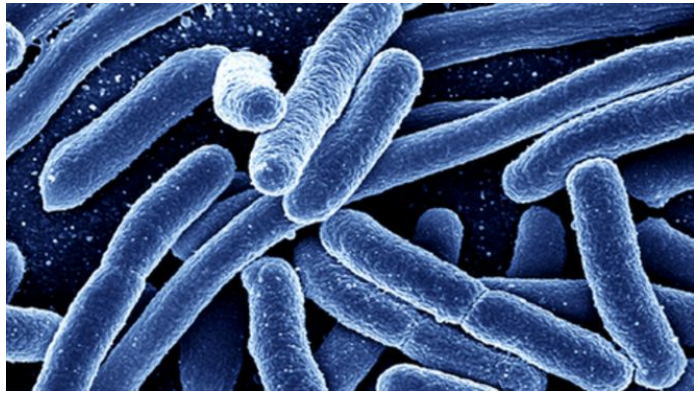
**Study of the interaction of the components of a biological systems and how these interactions give rise to the functions of the system**



# Why do we need to study Systems Biology?

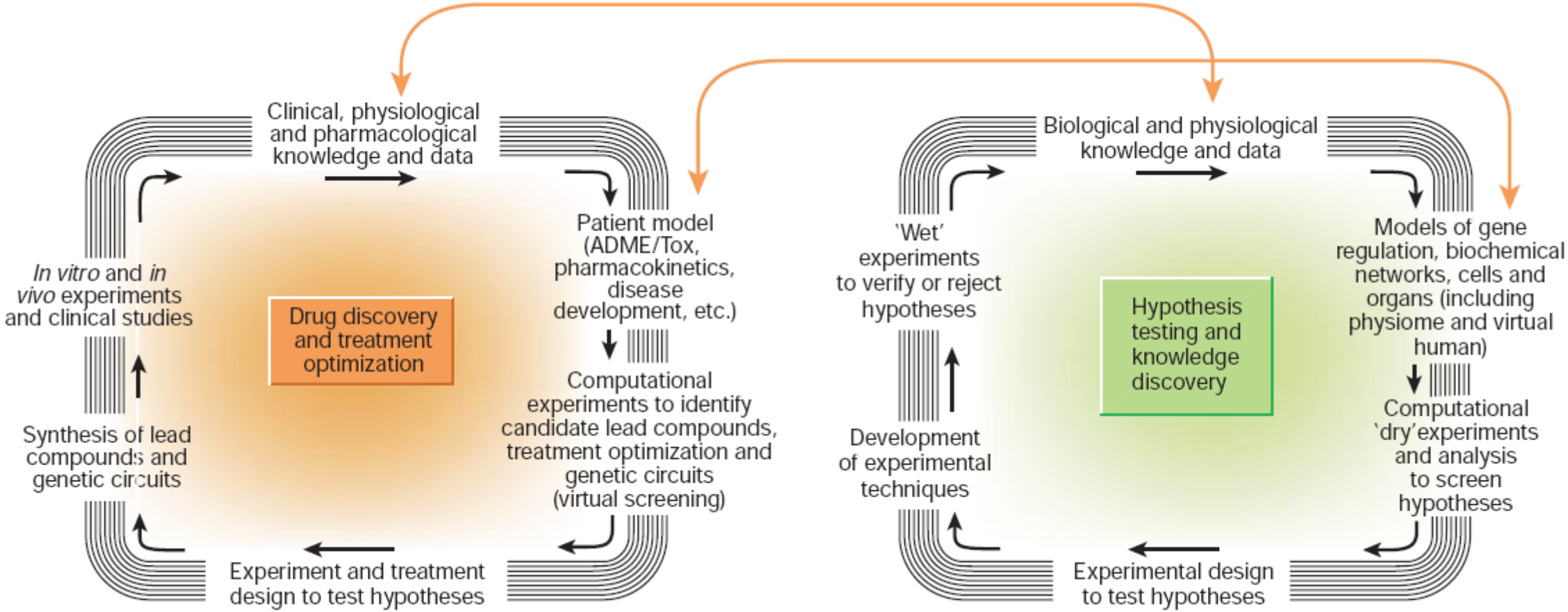
**To discover and understand the unifying principles of the Biological World**

Clown Fish



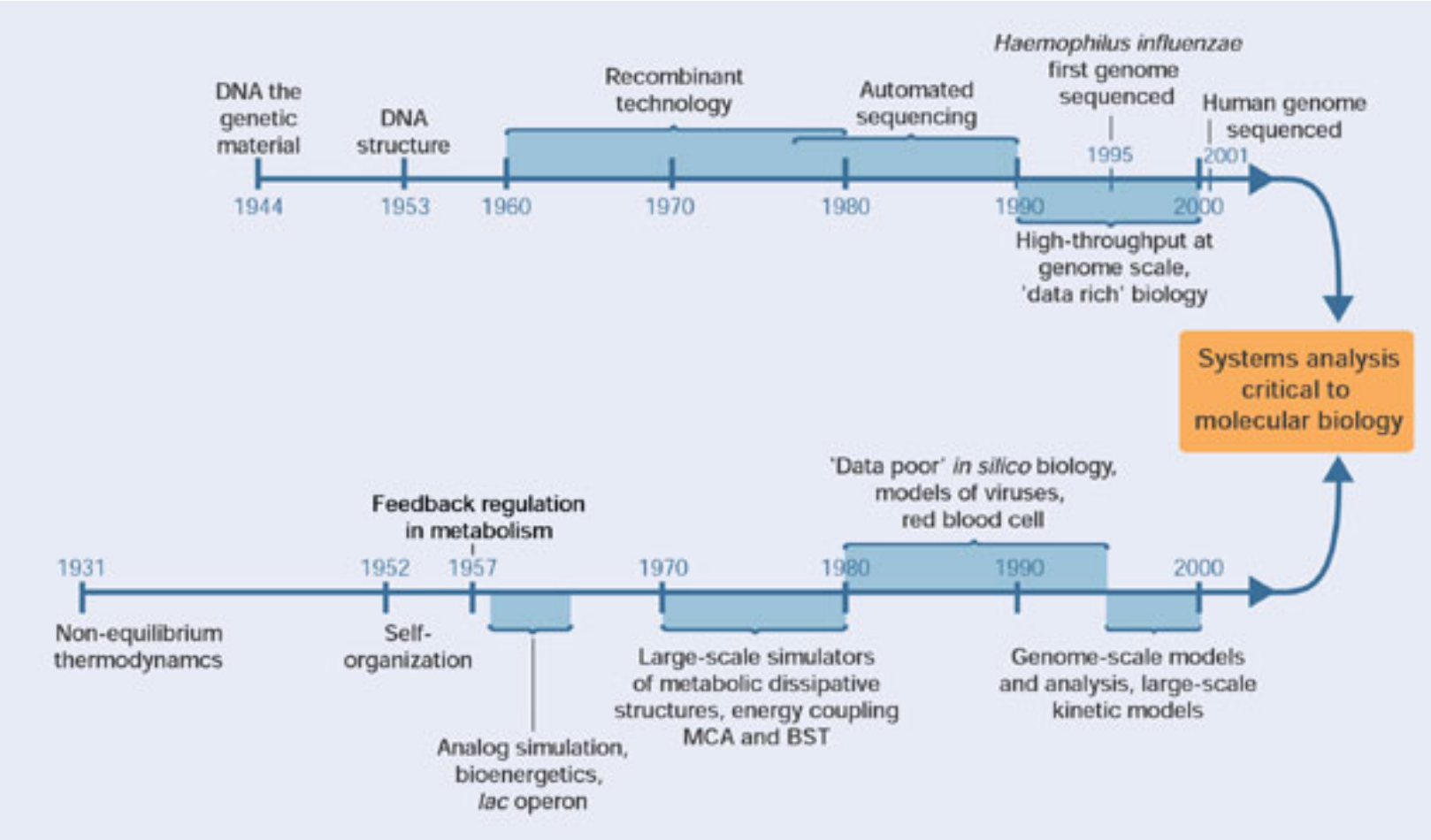
E. coli

# What does Systems Biology consist of?




H. KITANO, NATURE, VOL 420, 2002

# Developments Leading to the Field of Systems Biology

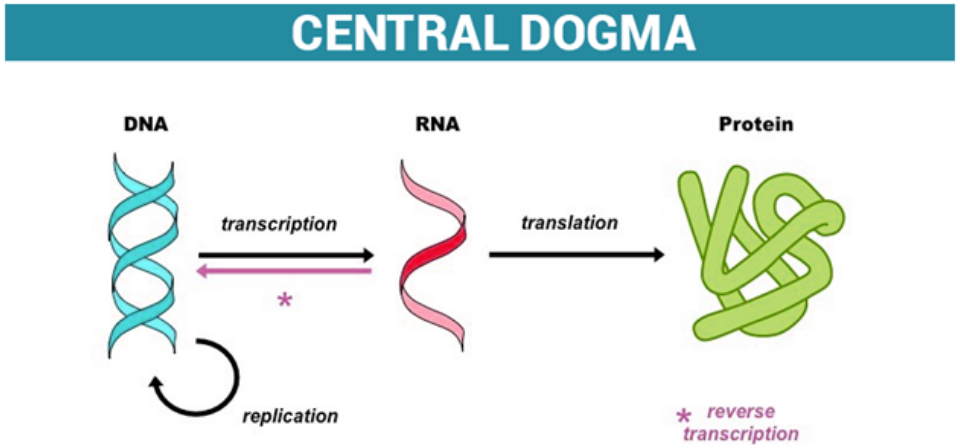



Hans V Westerhoff & Bernhard O Palsson, *Nature Biotechnology* Vol 22 (10) 2004

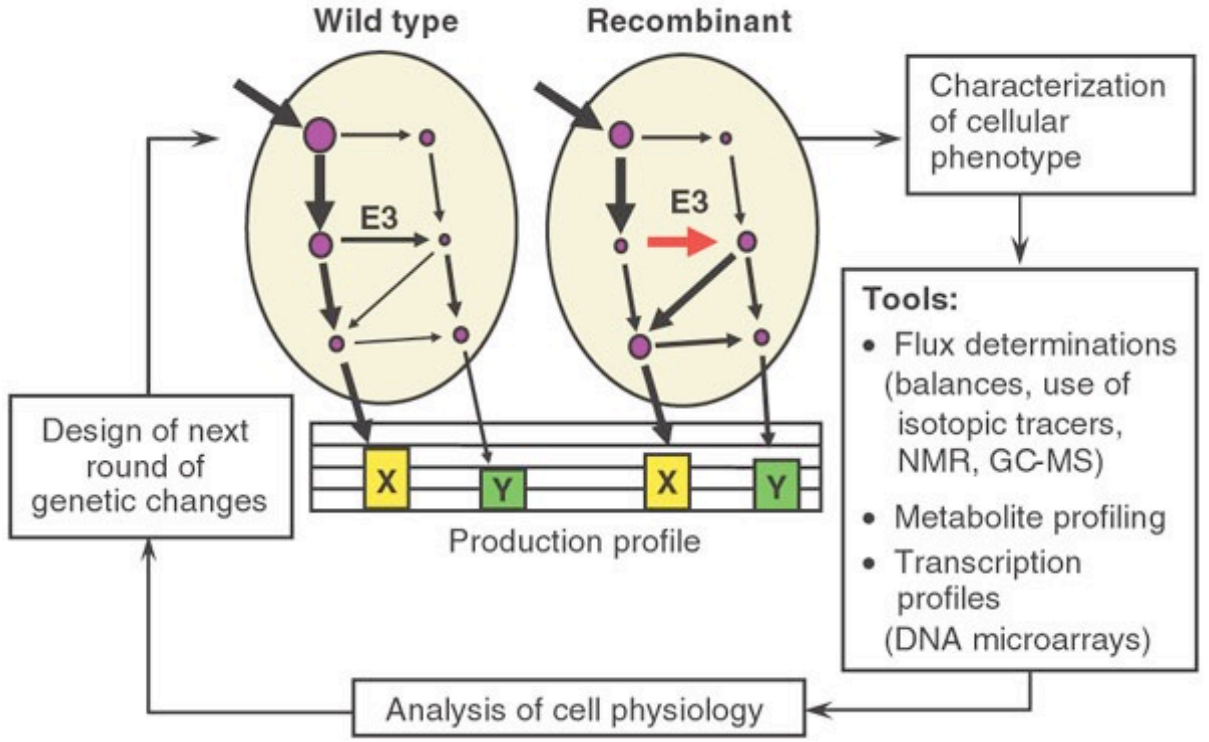
# Intersection of Mathematics, Engineering Science and Biology began in the 1980's

- Universal growth model was proposed by M. Savageau in the 1970s
- The single cell model was proposed by Shuler and Domach in the 1980s
- Metabolic Flux Analysis was proposed in the 1990s by Greg Stephanopoulos
- Systems Biology models was developed in the 2000s; one of the founders was Hiroaki Kitano
- Now models for evolution of life itself is being tested; these have be written based on principles of "systems biology"

# Metabolic Flux Analysis Caused a Shift in Paradigm of the Central Dogma




- Will this work?
  - Wild type cells are engineered to overexpress the enzyme E3 with the aim of increasing the yield of Y




Gregory Stephanopoulos, Hal Alper and Joel Moxley, Nature Biotechnology, Vol 22 No 10 Oct 2004

# The Equation that Defines MFA

- MFA begins with reaction network stoichiometries describing how substrates re converted to metabolic products and biomass constituents (macromolecular pools). Consider K metabolites participating in J reactions

$$\frac{d\mathbf{X}_{met}}{dt} = \mathbf{r}_{met} - \mu\mathbf{X}_{met}$$

$$\mathbf{X}_{met} \overset{\Delta}{=} \text{Vector of concentrations of intracellular metabolites}$$

$$\mathbf{r}_{met} \overset{\Delta}{=} \text{rxn vector containing net rates of formation}$$

- **Assumption**: pseudo–steady state
- Reason: 1. High turnover of metabolite pools

  2. Rapid adjustments of metabolite pools even after large perturbations

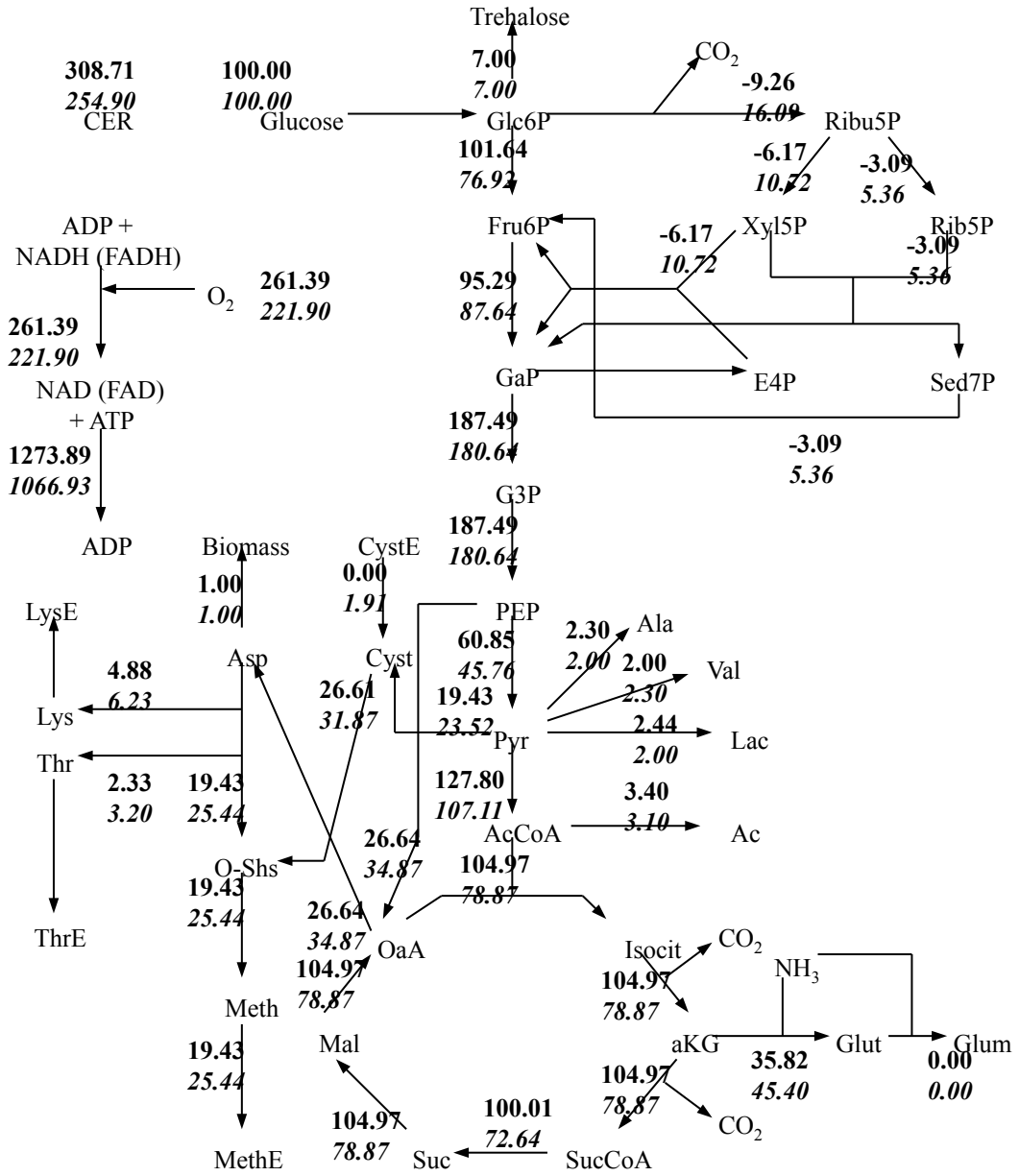
$$\therefore \quad \mathbf{0} = \mathbf{r}_{met} - \mu\mathbf{X}_{met}$$

- $\mu\mathbf{X}_{met}$ describes the dilution effect of metabolites due to growth. Intracellular level of metabolites is usually very low, hence

$$\mu\mathbf{X}_{met} << \mathbf{r}_{met}$$

$$0 = \mathbf{r}_{met} = \mathbf{G}^T\mathbf{v}$$

# One of the results of MFA is a Flux Map that shows the differences in reaction throughputs under different conditions being investigated

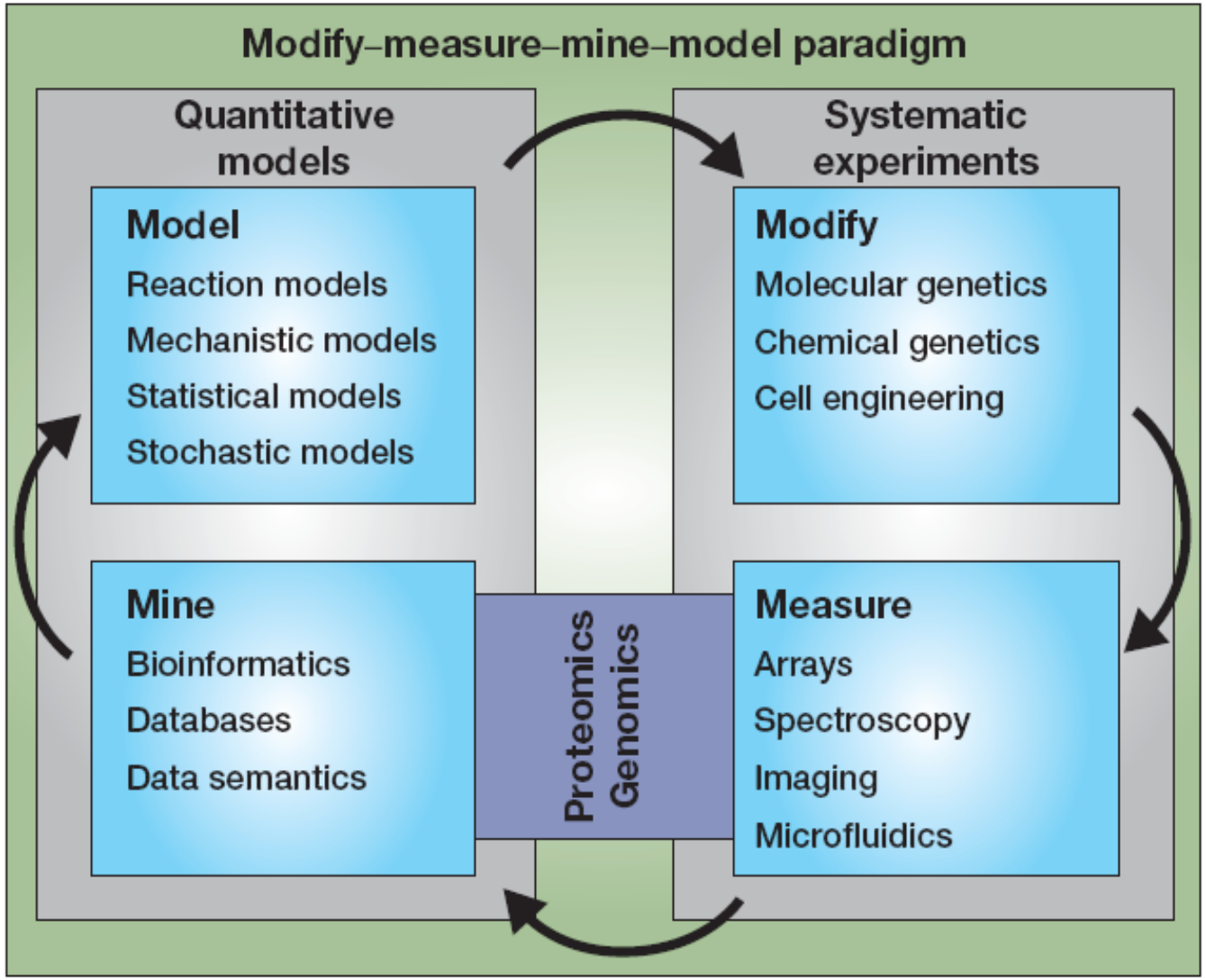



Metabolic flux analysis showing effect of cysteine on of methionine production by mutant *C. lilium*

# Why was there a need for creating a new field of Systems Biology?

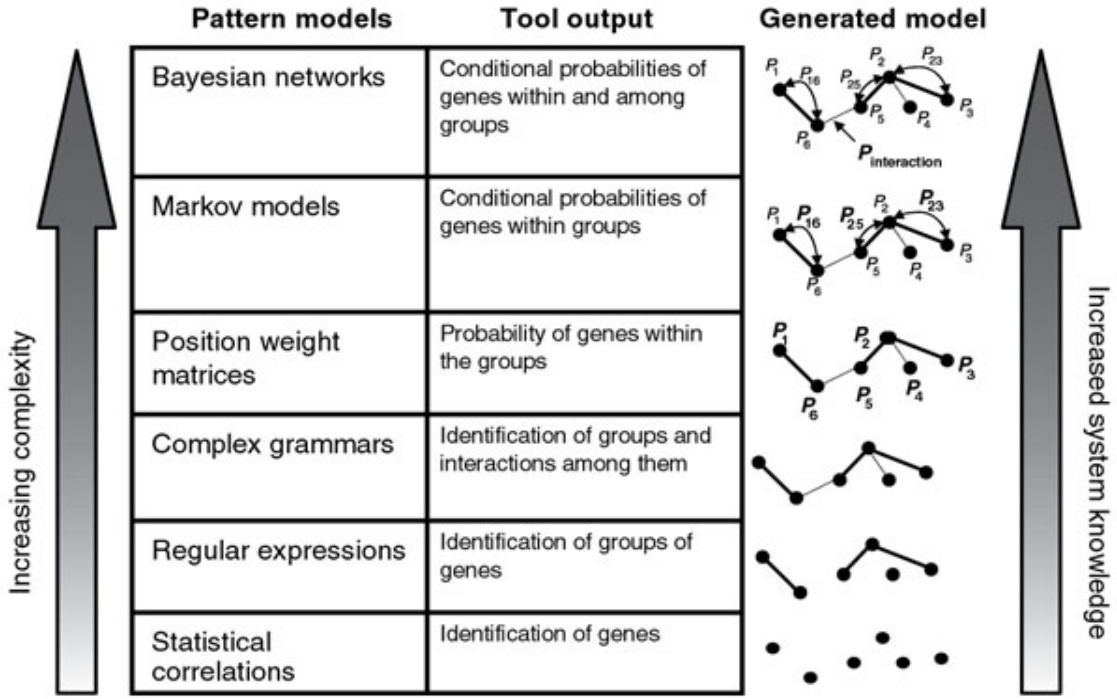
- MFA assumes that all the variables (concentrations of intracellular metabolites, proteins, etc. ) are in PSEUDO-STEADY STATE
- MFA does not include the regulation of gene expression

# Systems Biology - A shift from Pseudo-Steady State




NATURE CELL BIOLOGY, 2008

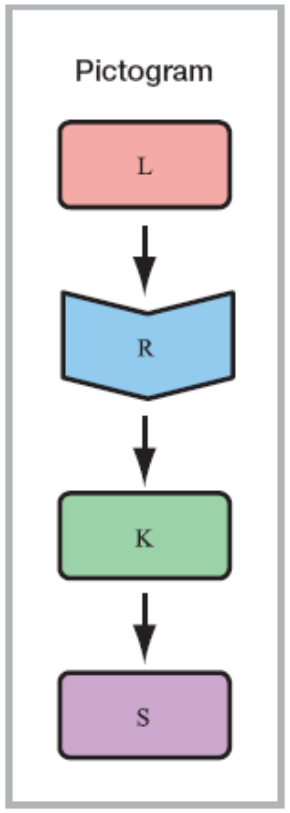
# Systems Analysis Integrates Physiological and Transcriptional Data




**Gregory Stephanopoulos, Hal Alper and Joel Moxley, Nature Biotechnology, Vol 22 No 10 Oct 2004**

- Various pattern models and data analysis techniques can be used for linking data sets
  - Statistical correlations can be used to link microarray data with phenotype
  - Models of increased complexity require more data but are also able to provide deeper insight

# Simple Example of Constructing a Model in Systems Biology




Reaction list

$$R+L \underset{k_{r1}}{\overset{k_{f1}}{\rightleftharpoons}} LR^*$$

$$LR^*+K \underset{k_{r2}}{\overset{k_{f2}}{\rightleftharpoons}} LR^*K$$

$$LR^*K \xrightarrow{k_{f3}} LR^*+K^*$$

$$K^*+S \underset{k_{r4}}{\overset{k_{f4}}{\rightleftharpoons}} K^*S$$

$$K^*S \xrightarrow{k_{f5}} K^*+S^*$$

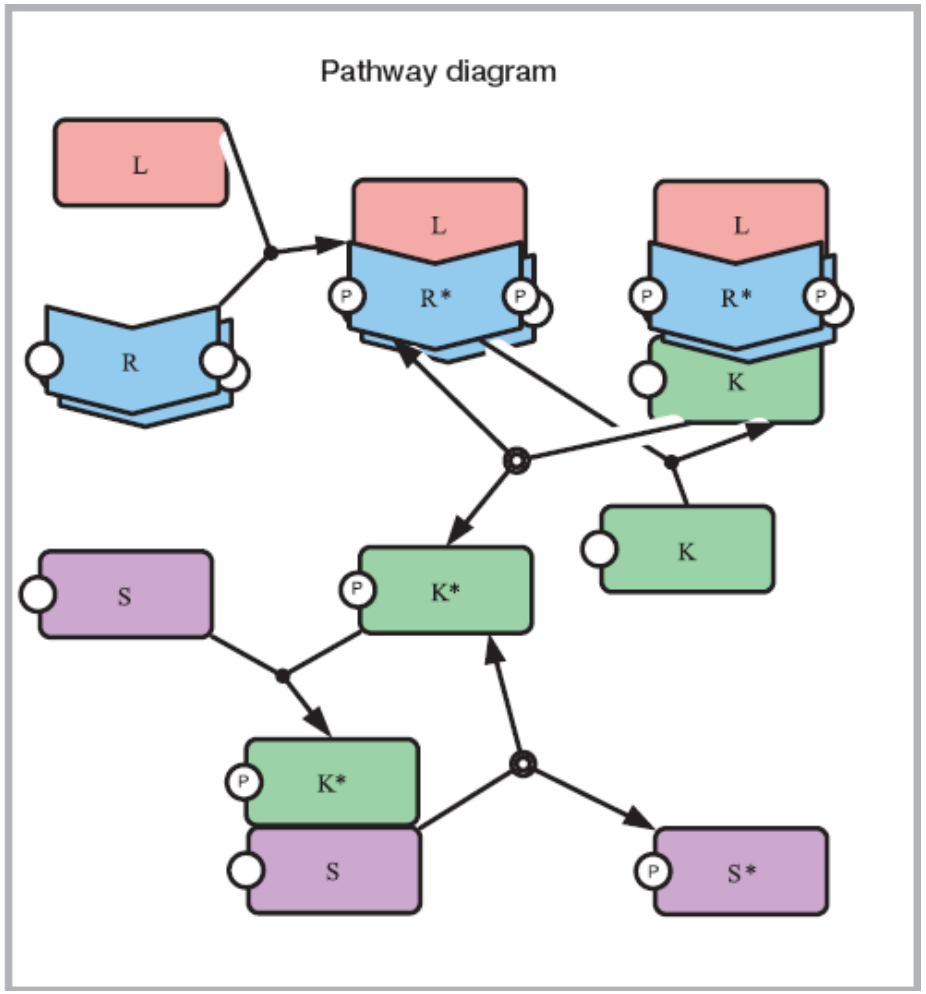
Approximations

if $[S]_0 >> [K^*]_0$: $\dfrac{d[K^*S]}{dt} \approx 0$

$$\frac{d[K^*]}{dt} = k_{f3}[LR^*K] + k_{f5}[K^*S] - \frac{k_{f5}[K^*]_0[S]}{\left(\frac{k_{r4}+k_{f5}}{k_{f4}}\right)+[S]}$$

$$\frac{d[S^*]}{dt} = \frac{k_{f5}[K^*]_0[S]}{\left(\frac{k_{r4}+k_{f5}}{k_{f4}}\right)+[S]}$$

# Simple Example of Constructing a Model in Systems Biology




Differential equations

$$\frac{d[R]}{dt} = -k_{f1}[L][R] + k_{r1}[LR^*]$$

$$\frac{d[LR^*]}{dt} = k_{f1}[L][R] - k_{r1}[LR^*] - k_{f2}[LR^*][K] + k_{r2}[LR^*K] + k_{f3}[LR^*K]$$

$$\frac{d[LR^*K]}{dt} = k_{f2}[LR^*][K] - k_{r2}[LR^*K] - k_{f3}[LR^*K]$$

$$\frac{d[K]}{dt} = -k_{f2}[LR^*][K] + k_{r2}[LR^*K]$$

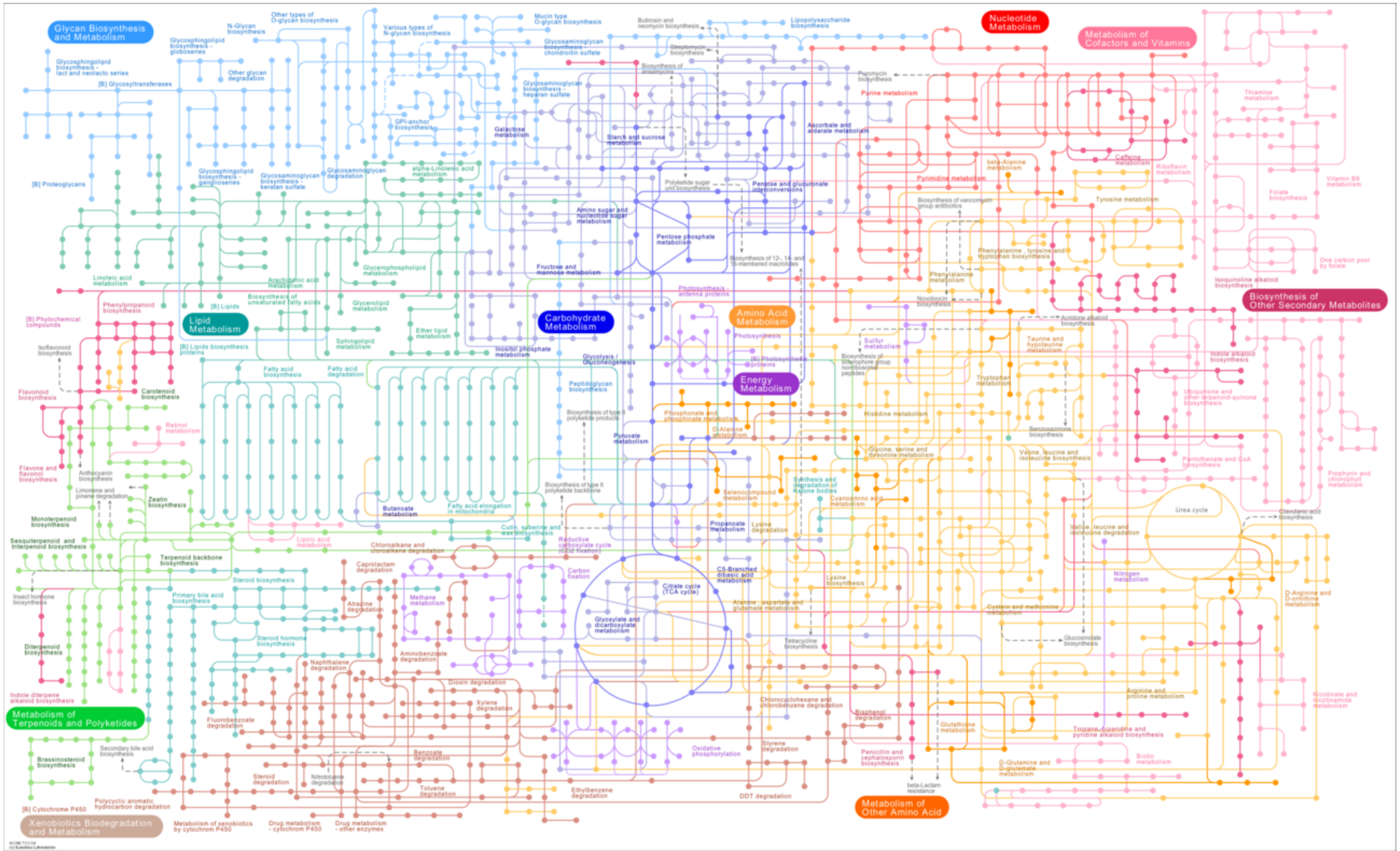
$$\frac{d[K^*]}{dt} = k_{f3}[LR^*K] - k_{f4}[K^*][S] + k_{r4}[K^*S] + k_{f5}[K^*S]$$

$$\frac{d[S]}{dt} = -k_{f4}[K^*][S] + k_{r4}[K^*S]$$

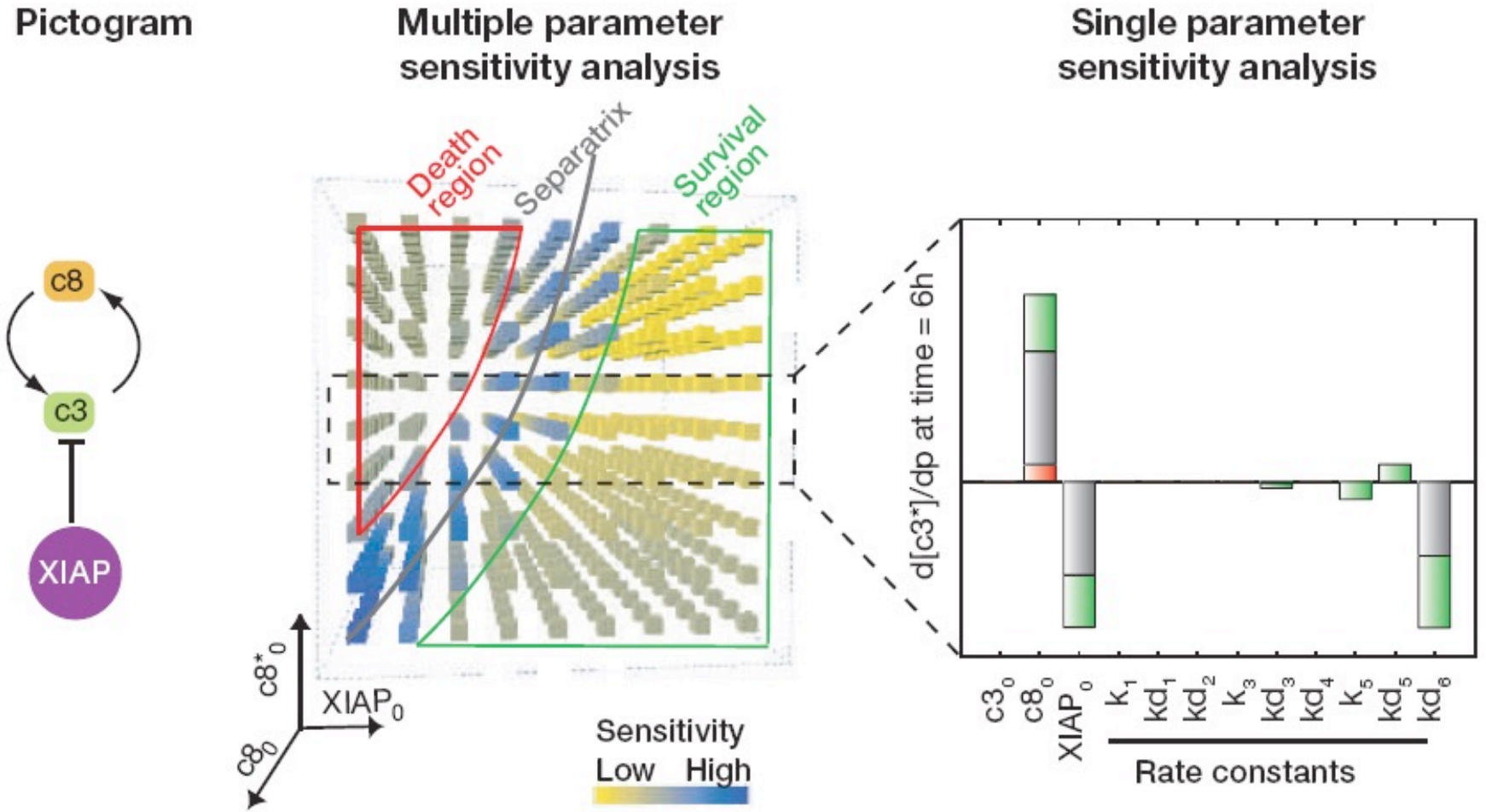
$$\frac{d[K^*S]}{dt} = k_{f4}[K^*][S] - k_{r4}[K^*S] - k_{f5}[K^*S]$$

$$\frac{d[S^*]}{dt} = k_{f5}[K^*S]$$

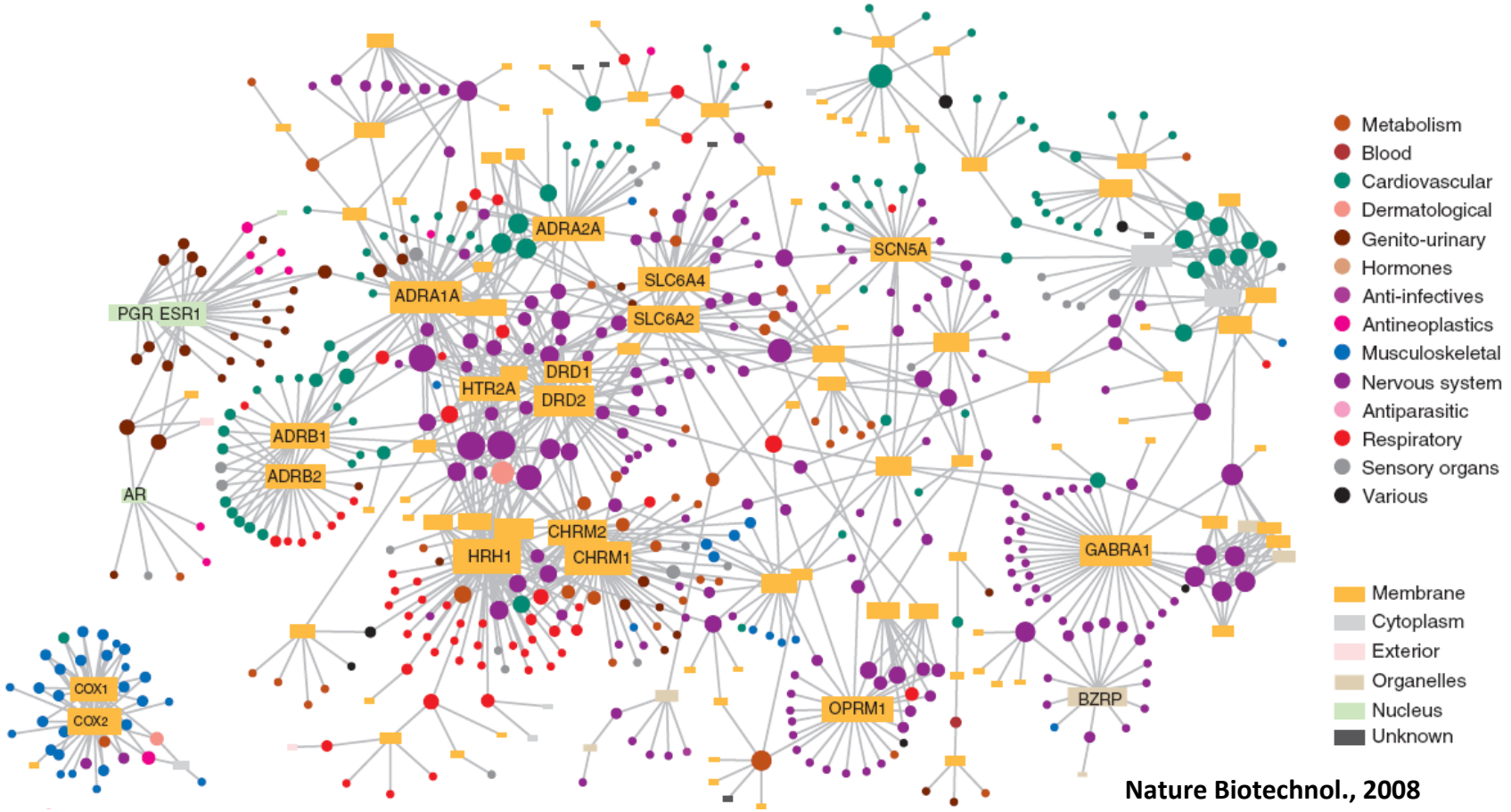
# Kyoto Encyclopedia of Genes and Genomes
# KEGG Atlas



# Sensitivity Analysis and Parameter Estimation in Systems Biology



# Understanding and Depiction of Networks for Designing Drug Targets




**Nature Biotechnol., 2008**

The DT network is generated by using the known associations between FDA-approved drugs and their target proteins. Circles and rectangles correspond to drugs and target proteins. Drug nodes (circles) are colored according to their Anatomical Therapeutic Chemical Classification, and the target proteins (rectangular boxes) are colored according to their cellular component obtained from the Gene Ontology database

# What are the Challenges facing Systems Biology?

- Data quality and standardization
  - Depends heavily on public domain data
  - Data sets are incomplete, not standardized, not properly annotated; very often uncertain
- Network topology
  - Development of new theoretical methods to understand, analyze and visualize the vast data acquired or generated for a problem
- Computation and organization
  - Search and analyze massive volumes of data
- Miniaturized automated microfluidic devices
  - Need to obtain faster, accurate and repeatable high throughput data
  - Bridge theory and experiments through fundamental principles
- Imaging
  - Dynamic spatial and temporal data for discovering new drug targets

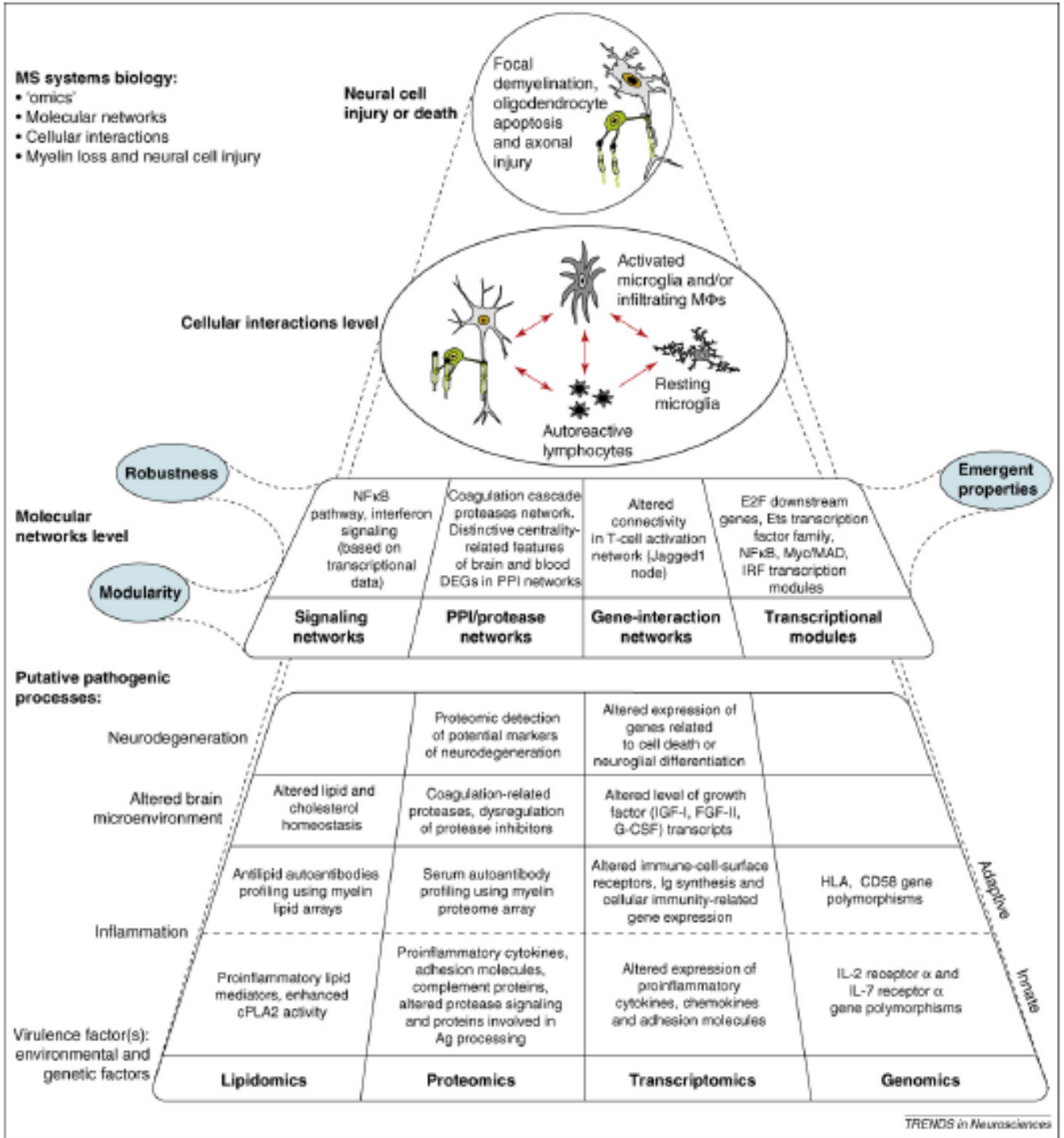
# What are the Elements for Analyses?

- System Structure Identification
  - Gene regulatory network
  - Metabolic reaction networks
- System Control
  - Cellular level control
  - External modification to repair defective control mechanism (treatment of diseases)
- System Design
  - Grow organs from the patients own cells
  - Using biological material for robotics and computation
- System Behaviour Analysis
  - Sensitivity to perturbations
  - Speed and characteristic of response

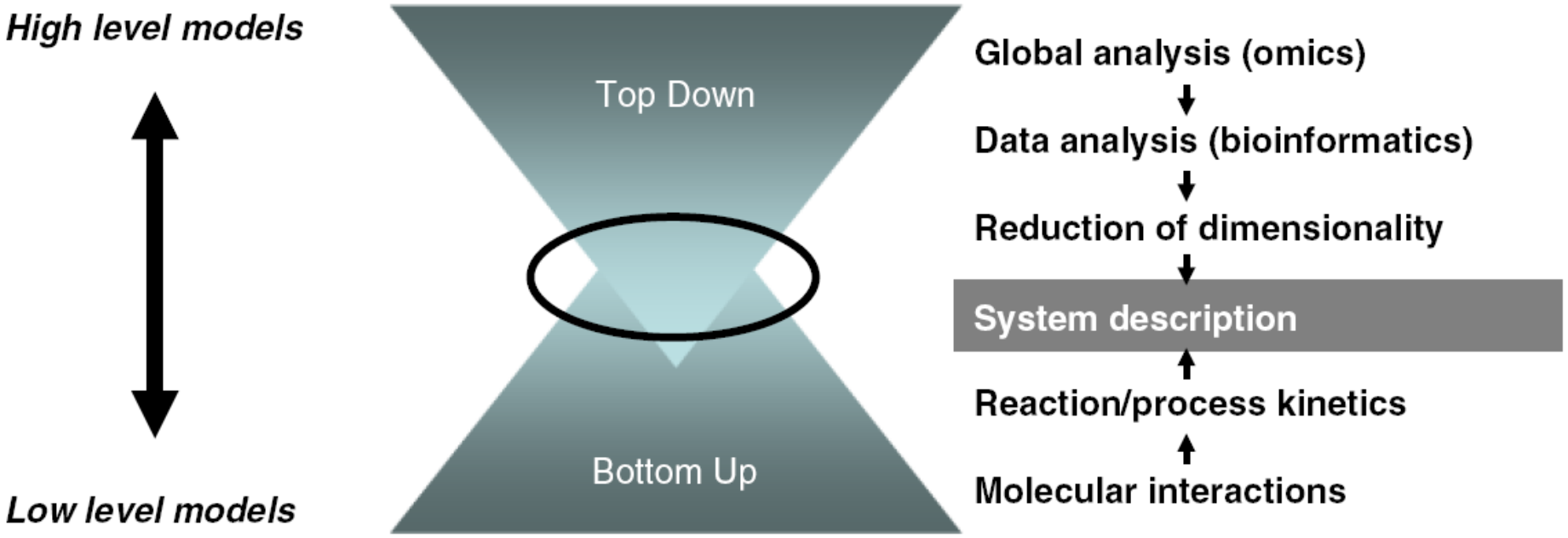
# Development of viable simulators

- Properties if the simulator
  - Functional for biological processes
  - Accurate, efficient and fast
  - Friendly and logical GUI
- No single simulator can answer all the problems
  - Simulators of events – e.g. rhythmic behavior
  - Simulators for gene expression, metabolism and signaling – e.g gene expression
  - Stochastic process, noise and uncertainty
  - phenotype simulation

# Systems biology of multiple sclerosis

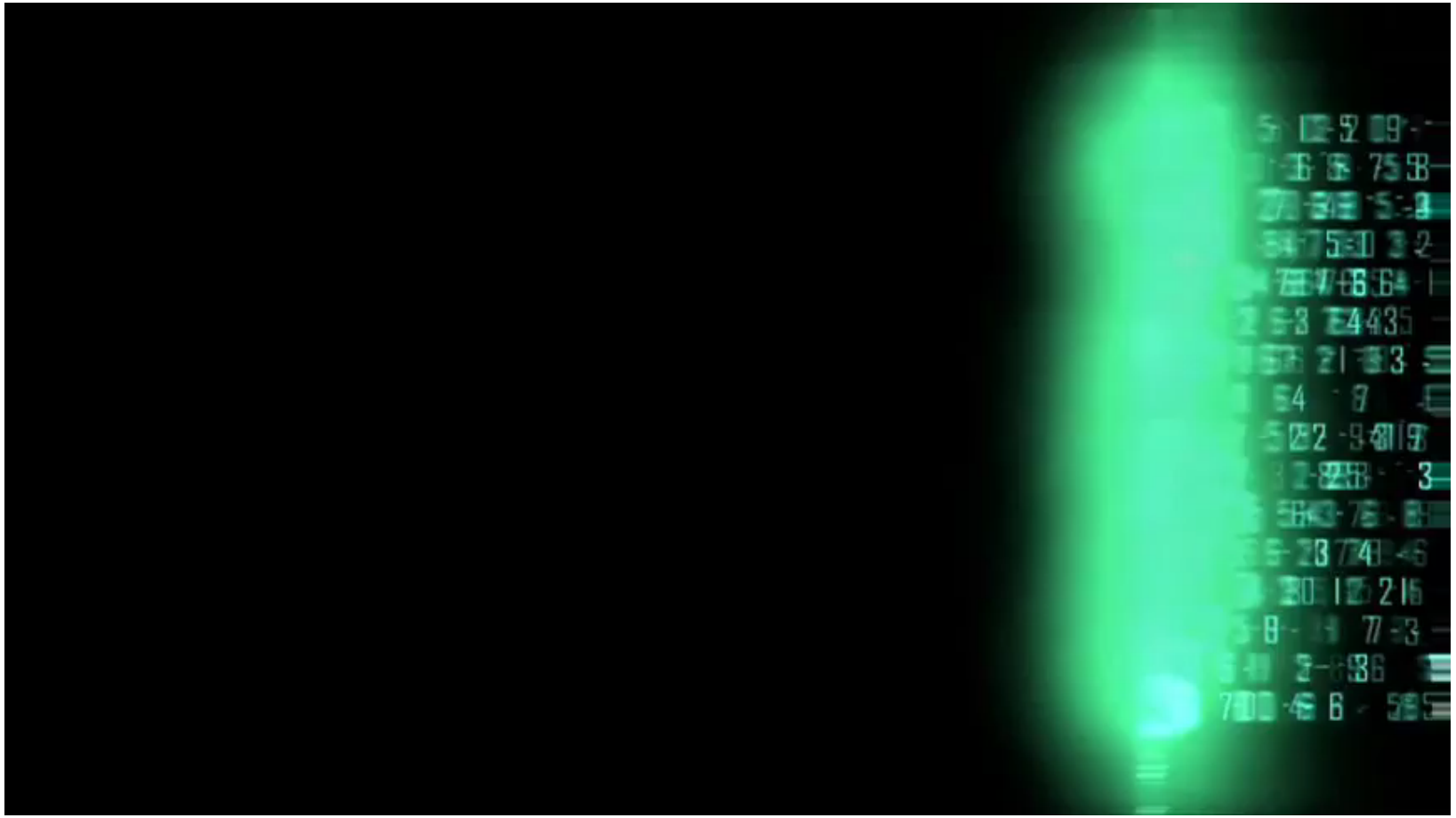


# Top-down Bottom-up Systems Biology




- In the top-down approach, high-throughput data are applied for identification of structures, connectivity, and possible information on the quantitative interaction between different components
- In the bottom-up approach, the system is reconstructed based on biological knowledge, e.g. on molecular interactions

# Systems Biology Simulator



# SYNTHETIC BIOLOGY logically follows Systems Biology

- 1980: used by Barbara Hobom to describe bacteria that had been genetically engineered using recombinant DNA technology
- 2000: term ‘synthetic biology’ was again Eric **Kool** and other speakers at the annual ACS Meeting to describe the synthesis of unnatural organic molecules that function in living systems

# Enigma of Synthetic Biology

- **Make discoveries and overturn paradigms**
- Success or failure as an engineering discipline depending on where independent approximations become useful in the continuum between the atomic and macroscopic worlds
- Assemble existing biological parts into machines, and create artificial systems that reproduce the emergent properties of living systems

# Interchangeability Leading Synthetic Genetic Systems

- The Watson-Crick genetic code
  - Is this the only solution for “life” across the universe?
  - Optimality of the genetic code
- Synthesis of “nucleobases” that can support an artificial genetic coding system
  - A synthetic genetic alphabet with up to 12 independently replicatable nucleobase pairs can be supported by an extended set of Watson–Crick rules
  - Protein engineering converts natural polymerases into polymerases that accept components of an expanded genetic alphabet in a polymerase chain reaction

# Challenges in Synthetic Biology

- Proteins
  - Proteins do not possess the repeating charge present in nucleotides
- Vision
  - synthetic biologists would first alter the behaviour of proteins by replacing amino acid
  - the behaviour of a protein is not a simple combination of independent contributions from the constituent amino acids
- More seriously
  - even the simplest of molecular interaction are poorly understood.
  - the chemical theory cannot retrodict the freezing point of water, the solubility of simple salts in water, or the packing of crystals of simple organic molecules

# The Developmental Stages of the Foetus in Reptiles is Similar