计算系统生物学

王 勇

中国科学院数学与系统科学研究院

课程信息

- http://www.aporc.org/doc/wiki/Course001

- Contains all course-related materials (lecture slides & further readings), regularly updated

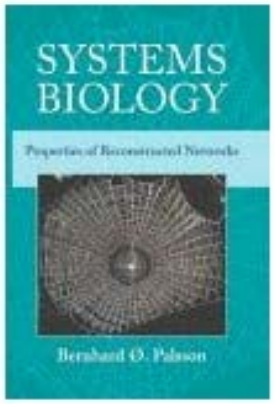
课程信息

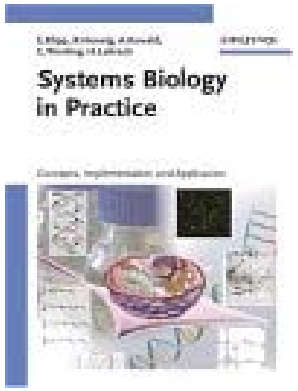
- 答疑&讨论: 课后或思源楼 1003
- 对课程的意见和建议

邮件： ywang@amss.ac.cn

推荐书目

- An introduction to Systems Biology: Design Principles of Biological Circuits
  by Uri Alon
  June 2006, Chapman&Hall/CRC, Taylor and Francis Group
- Systems Biology : Properties of Reconstructed Networks
  by Bernard Palsson
  January 2006, published by Cambridge Univ. Press
- Systems Biology in Practice: Concepts, Implementation And Application
  Klipp, E et al.
  John Wiley & Sons Inc. 2005
- Systems Biology: A Textbook Edda Klipp, et al. 2009





http://www.systems-biology.org

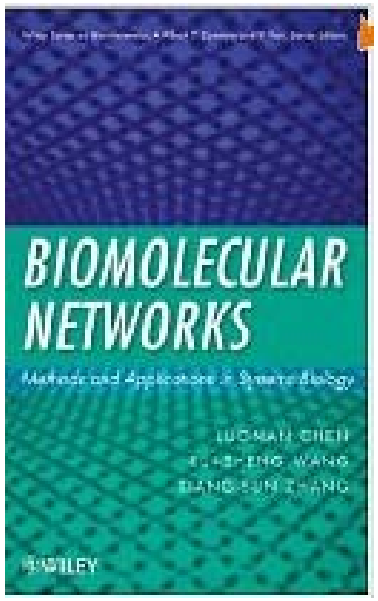
State-of-Arts

- Alerts from Science, Nature, Cell, PNAS

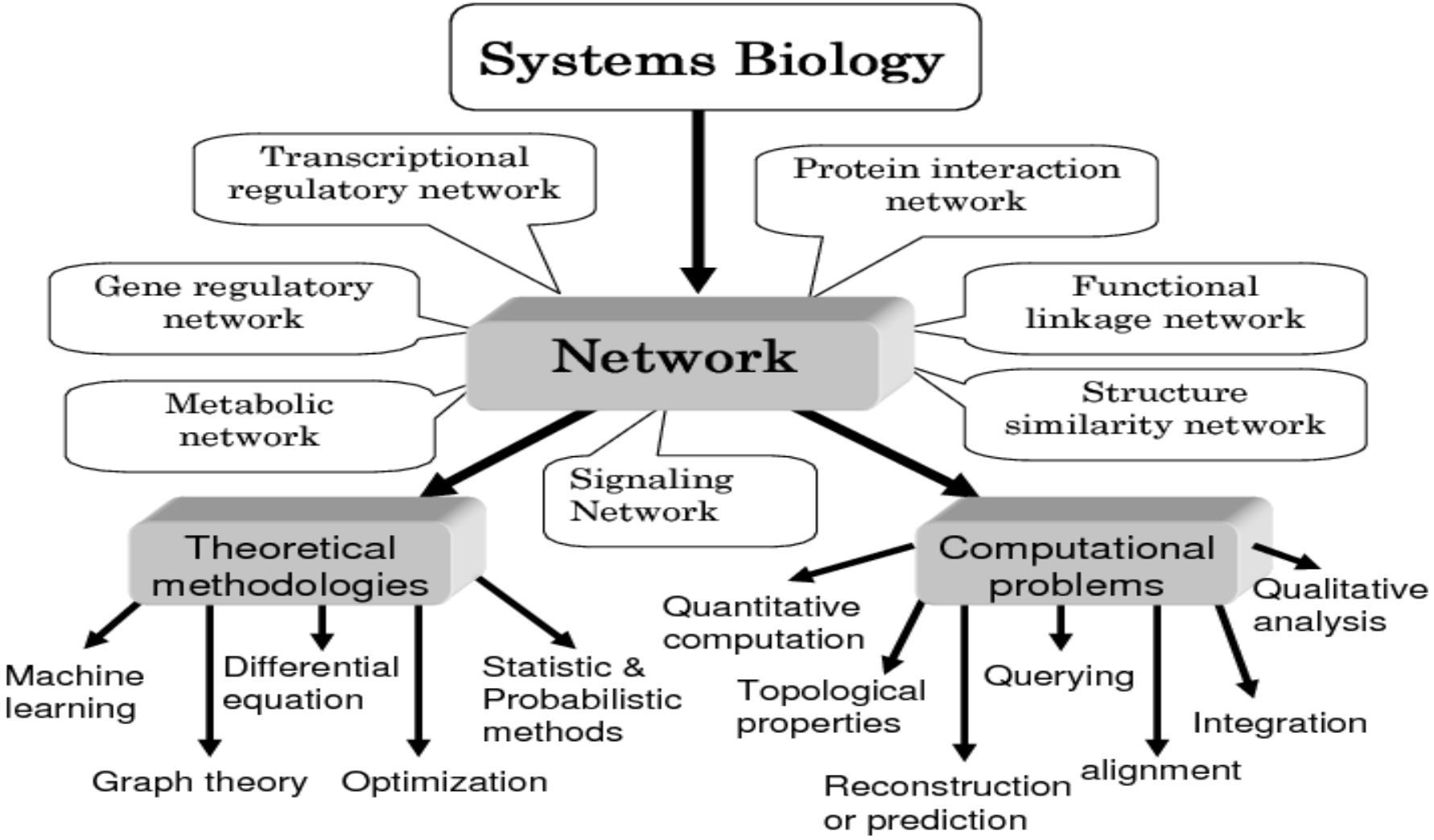
- Nature Molecular Systems Biology
- BMC Systems Biology
- IET Systems Biology
- Other related journals

- Google, Wiki

We focus on biomolecular networks



Luonan Chen, Rui-Sheng Wang, Xiang-Sun Zhang.
Biomolecular Networks: Methods and Applications in Systems Biology.
John Wiley & Sons, Hoboken, New Jersey. July, 2009.



Network Systems Biology

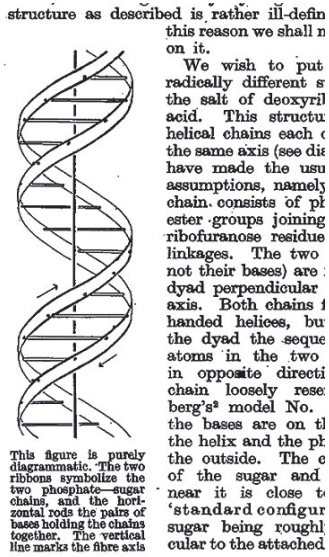
大纲

1. 基因调控网络重建
2. 转录调控网络重建
3. 转录因子合作网络预测
4. 蛋白质相互作用网络预测
5. 生物分子网络分析
6. 生物分子网络比对
7. 生物分子网络motif，模块分析
8. 生物活性通路与网络标记物识别

系统生物学

系统生物学

《Nature》1953



structure as described is rather ill-defin
this reason we shall n
on it.

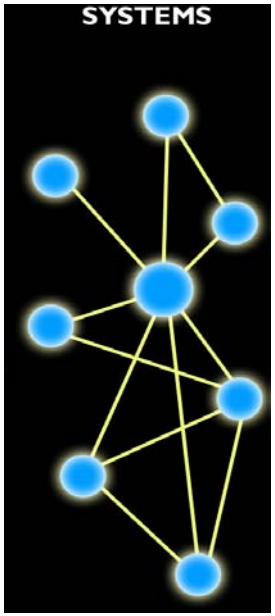
We wish to put
radically different s
the salt of deoxyril
acid. This structu
helical chains each
the same axis (see dia
have made the usu
assumptions, namely
chain consists of ph
ester groups joining
ribofuranose residue
linkages. The two
not their bases) are
dyad perpendicular
axis. Both chains
handed helices, bu
the dyad the seque
atoms in the two
in opposite directi
chain loosely rese
berg's[2] model No.
the bases are on th
the helix and the ph
the outside. The c
of the sugar and
near it is close t
'standard configur
sugar being roughl
cular to the attached

This figure is purely diagrammatic. The two ribbons symbolize the two phosphate—sugar chains, and the horizontal rods the pairs of bases holding the chains together. The vertical line marks the fibre axis

DNA双螺旋结构发现，
开启分子生物学研究

《Science》2001

人类基因组计划完成



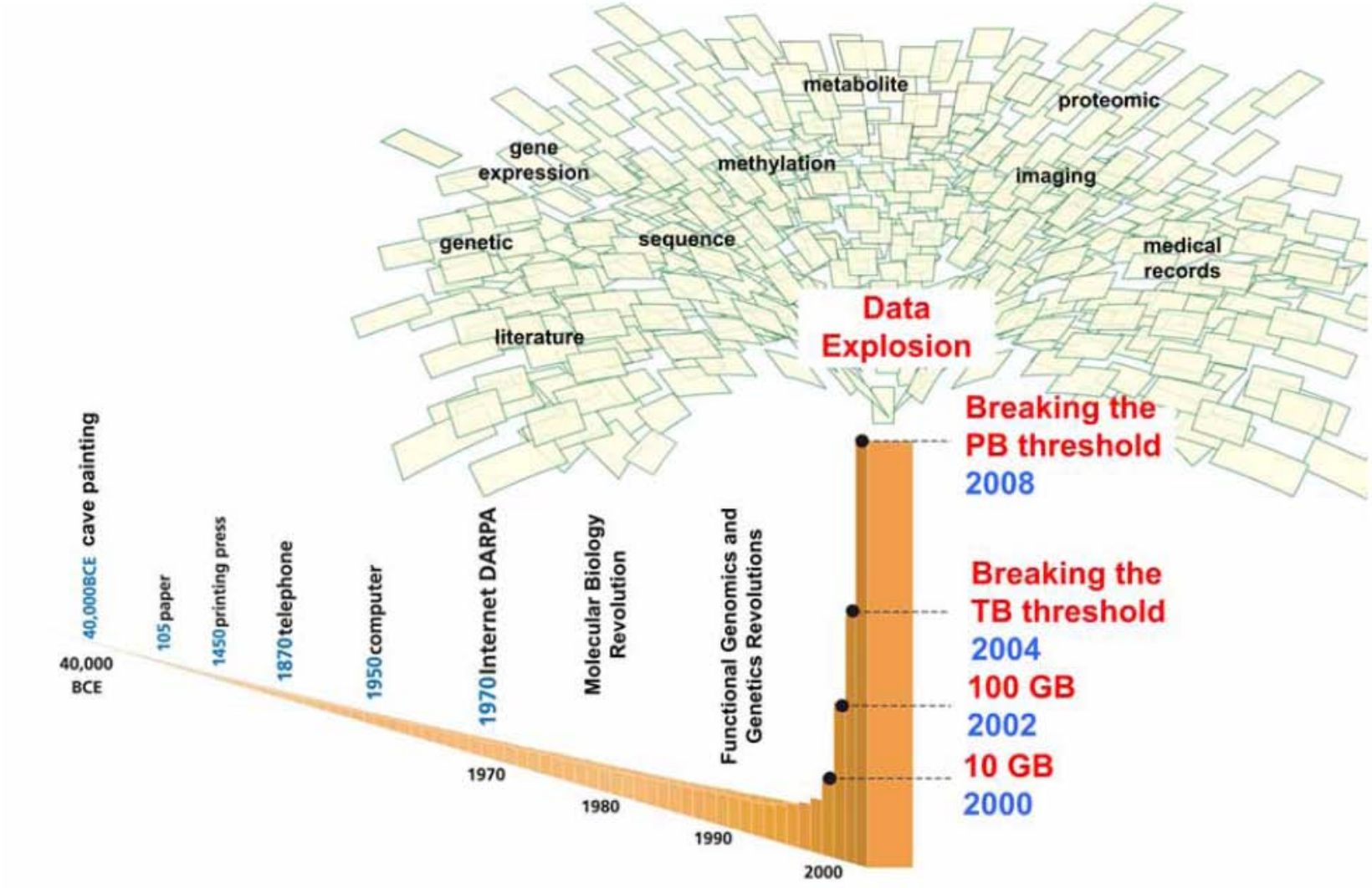
系统生物学

1950

1990

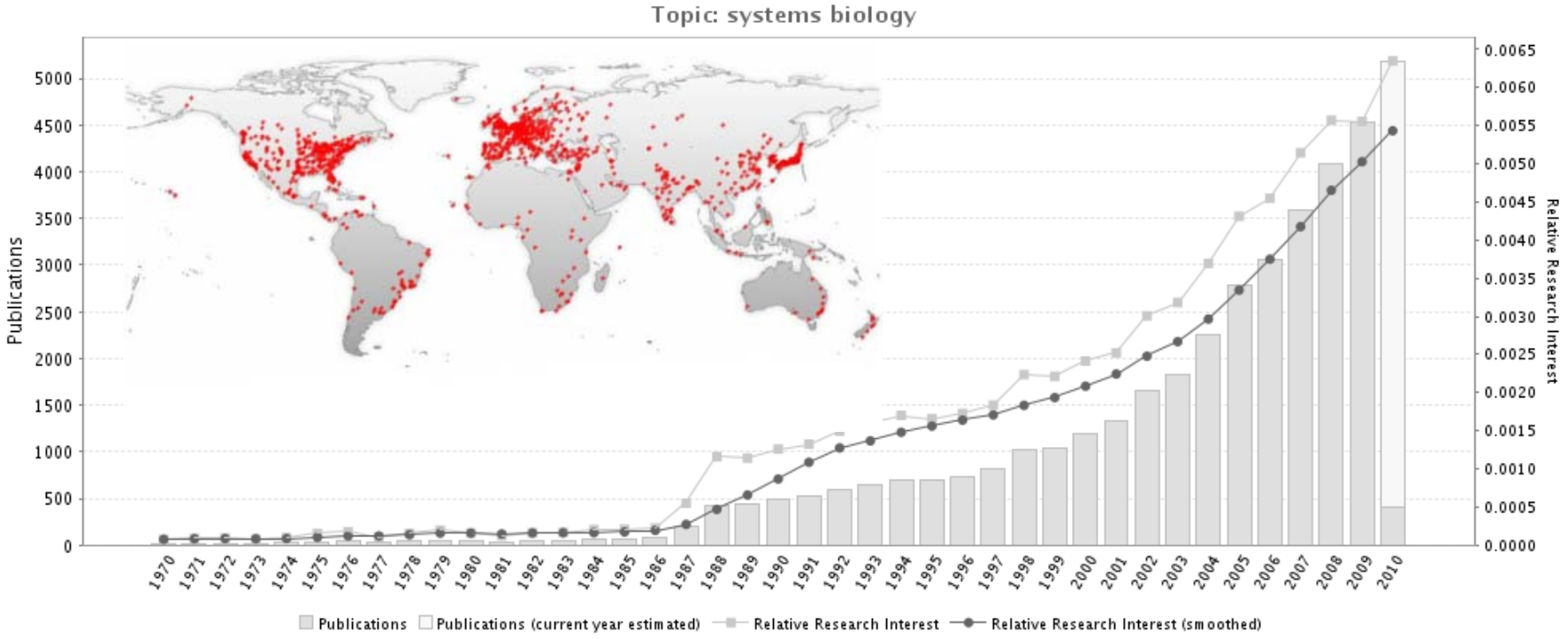
2002

生物信息学



http://www.sagebase.org

GoPubMed 是款新的垂直搜索引擎，内容涉及生命科学的方方面面。输入相应的关键词，即可获得带有高亮显示的搜索结果条目。左侧栏还有分类显示目录，标注不同的检索方式。应该说界面相当友好，且很精致，有需要的可以保存供日后使用。



什么是系统生物学？

- 内涵？
- 外延？
- 问题？

Systems Biology

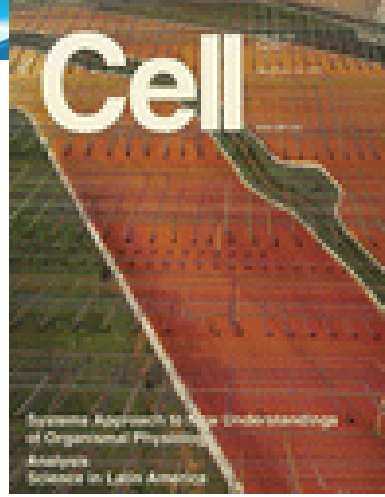
- Networking the whole biological system, rather than studying its isolated parts.

- Integrating large amounts of data in the context of biological network (Sequence, structure, function, gene expression, protein expression, protein interaction, protein-DNA interaction, and literature data).

Procedure---Systems Biology

- System Perturbation
- Generating of comprehensive global data
- Identification of key molecules
- Network modelling
- Generation of hypotheses
- Validation of hypotheses

- Journal club



A Predictive Model for Transcriptional Control of Physiology in a Free Living Cell

Richard Bonneau et. al. Cell, Vol 131, 1354-1365, 28 Dec. 2007

Institute for Systems Biology, Seattle, WA 98103, USA
Center for Genomics & Systems Biology, New York University, New York, NY 10003, USA

On the cover: Brightly colored blooms of halophlic (喜盐的) organisms in the salt flats of the South San Francisco Bay (photograph by Michael Melford, courtesy Getty Images) serve as a vibrant backdrop(背景) for a segment of a predictive environmental and gene regulatory circuit determined for one of this ecosystem's principal inhabitants, the archaeon(古代生物) Halobacterium salinarium NRC-1(一种嗜盐的古生菌，一般只生存在盐水池塘或是盐湖中). This organism possesses a number of fascinating adaptations for life in hypersaline (高盐) environments including the production of membrane pigments (细胞膜上产生色素) that mediate light-driven energy production and flotation devices called gas vesicles for vertical mobility in search of favorable oxic regimes (氧载体). While the availability of unique adaptations is important, the integrated regulation of these and many other core physiological processes (生理学过程) is vital for survival in this dynamic environment.

In this issue, Bonneau et al. report a systems level regulatory circuit for the transcriptional control of 80% of all genes in this organism. This regulatory model accurately predicts the transcriptional changes that occur when Halobacterium is challenged with new environmental and genetic perturbations. Significantly, this study supports the claim that fundamental properties of biological systems and their environments should enable the rapid construction of highly accurate, predictive models of global gene regulation for both traditional model systems and for many more currently uncharacterized organisms.

Faculty of 1000 Biology

- ***"This paper represents an exceptionally important milestone in the field..."***
  Evaluated by Faculty of 1000 Biology member **Charles Auffray** (Centre National de la Recherche Scientifique (CNRS) - UMR 7091, France)

“Faculty of 1000 Biology”创办于2002年1月，根据全球资深科学家的意见，提供对近期发表的生物科学论文的快速评论，目的是帮助广大科研人员遴选和发现有价值的研究工作。

Other comments

- **Research Highlight by *Nature Reviews Microbiology* 6, 92 (February 2008)**

- **Bio-IT World's Systems Biology newsletter.**(In the closing days to 2007, a really nice piece of systems biology work was published in the journal *Cell*)

Why this paper

- Cell publish computational biology work
- From Institute for Systems Biology, Seattle
- The ISB founder, also the founder of systems biology Lee Hood is one of the co-authors.
- To taste the flavor of systems biology (network+perturbation+data integration)

What they studied

- A largely uncharacterized organisms
- Easy to be cultured
- The environment significantly influences the dynamic expression

Methodology

- Experiments:

1. Microarray data: Total 413 experiments (Time-course and steady state, 8 environment effects perturbation, combinatorial perturbation. 33 gene deletion and GTF overexpression)

   266 experiments in training set

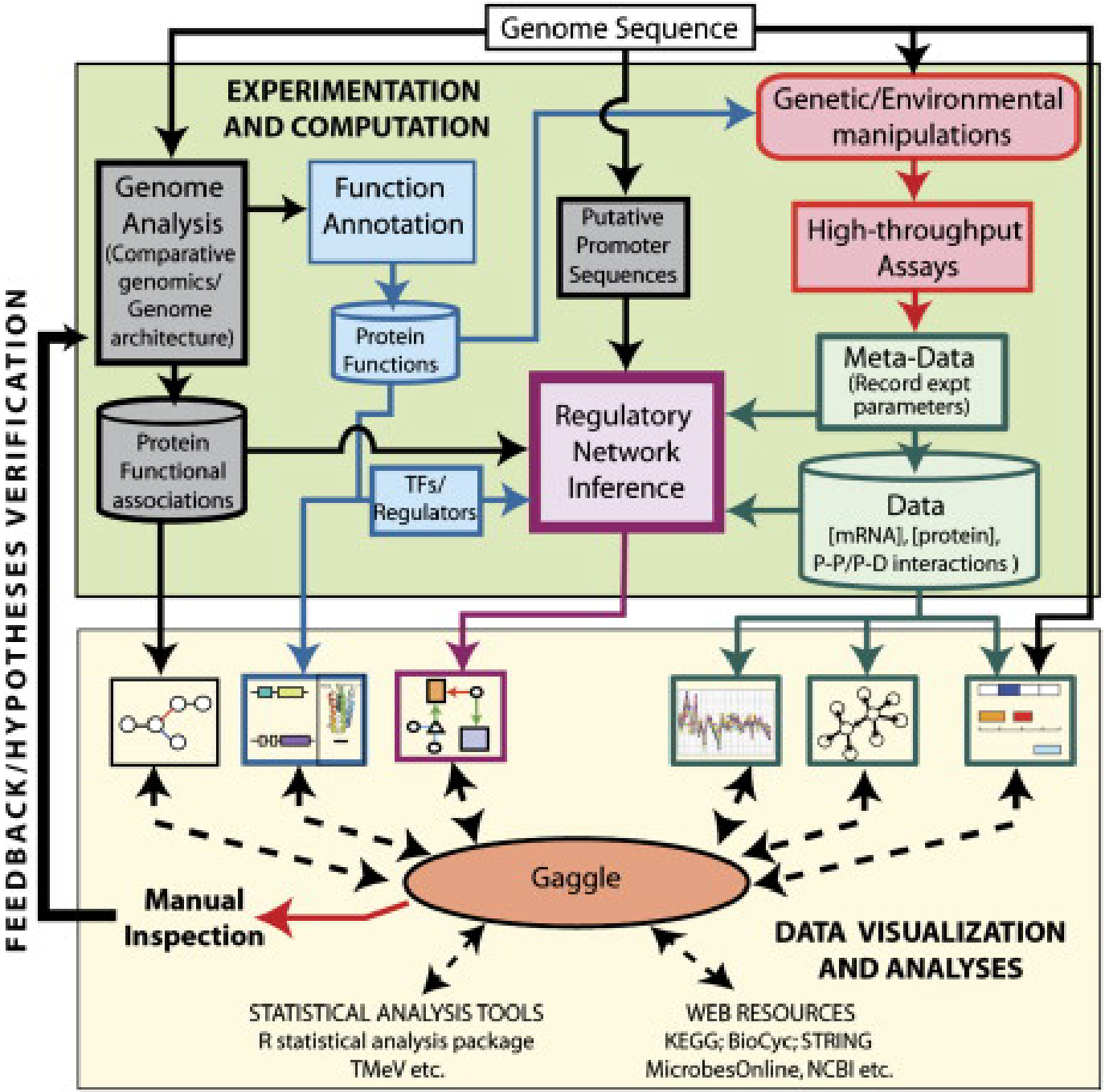
   147 new experiments

2. ChIP-chip data

http://baliga.systemsbiology.net/egrin.php

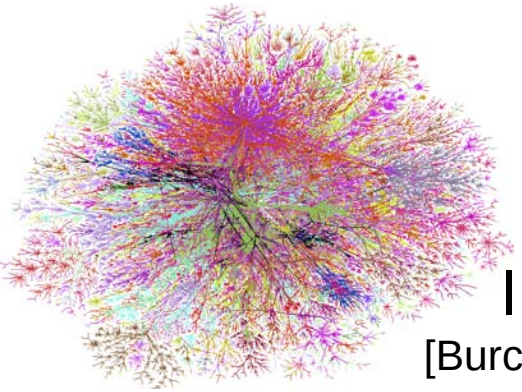
- Computational prediction

1. Protein structure prediction
2. Function annotation algorithm
3. Biclustering algorithm (Data integration and dimensionality reduction)
4. Transcriptional regulatory network inference

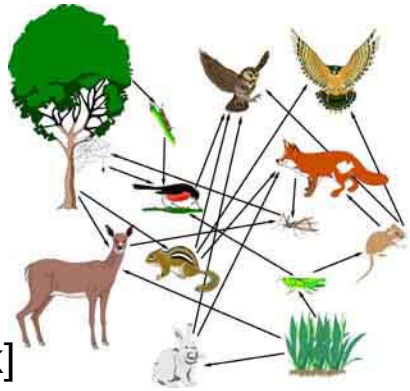

Genome Sequence
EXPERIMENTATION AND COMPUTATION
Genetic/Environmental manipulations
Genome Analysis (Comparative genomics/ Genome architecture)
Function Annotation
Putative Promoter Sequences
High-throughput Assays
Protein Functions
Meta-Data (Record expt parameters)
Protein Functional associations
Regulatory Network Inference
TFs/ Regulators
Data [mRNA], [protein], P-P/P-D interactions)
FEEDBACK/HYPOTHESES VERIFICATION
Gaggle
Manual Inspection
DATA VISUALIZATION AND ANALYSES
STATISTICAL ANALYSIS TOOLS
R statistical analysis package
TMeV etc.
WEB RESOURCES
KEGG; BioCyc; STRING
MicrobesOnline, NCBI etc.

生物分子网络

Networks as a universal language

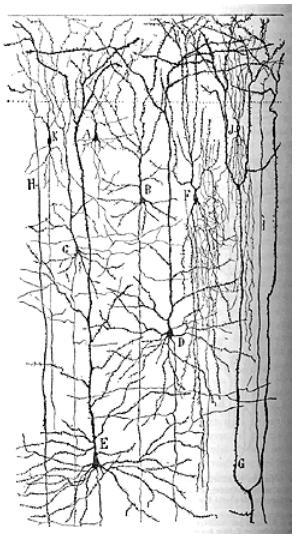


Internet
[Burch & Cheswick]

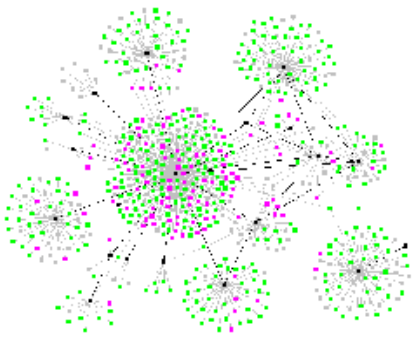


Food Web

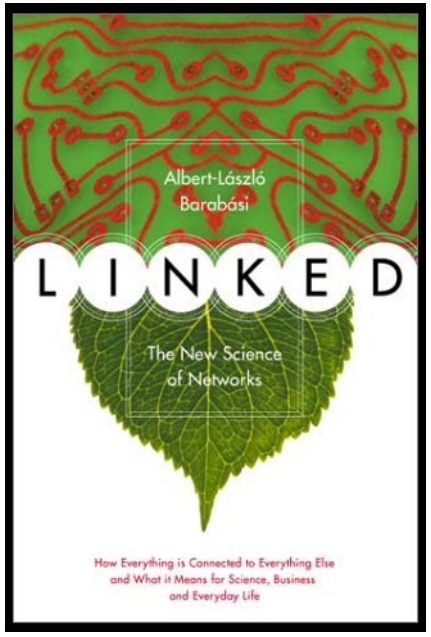
Electronic Circuit



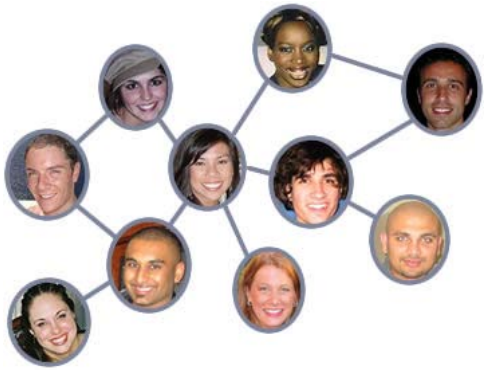
Neural Network
[Cajal]



Disease Spread
[Krebs]



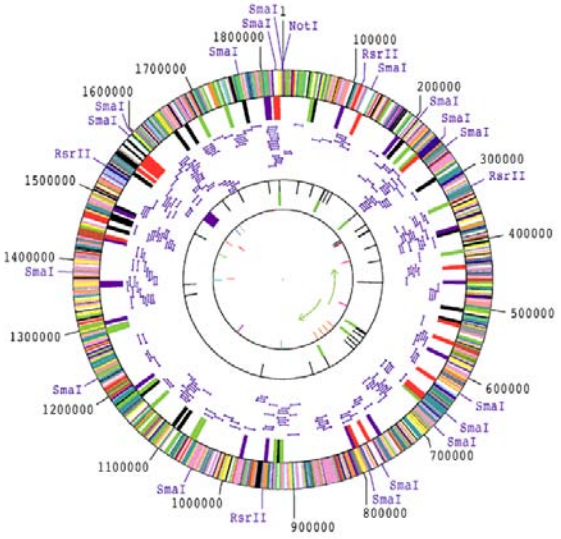
Protein Interactions
[Barabasi]



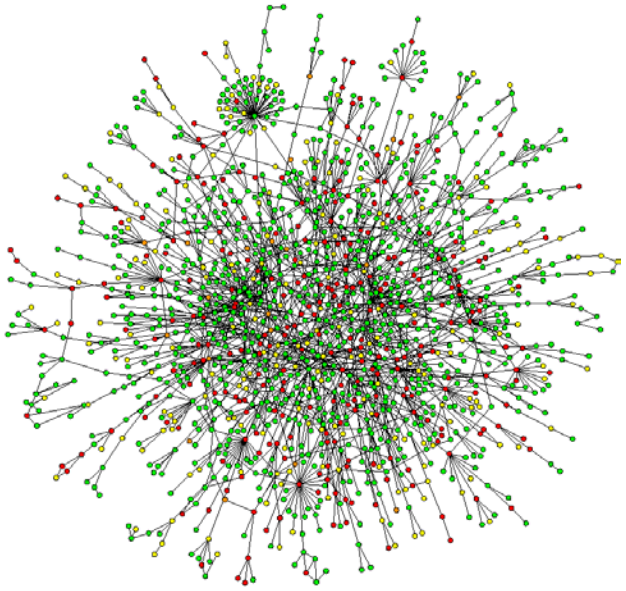
Social Network

生物分子网络

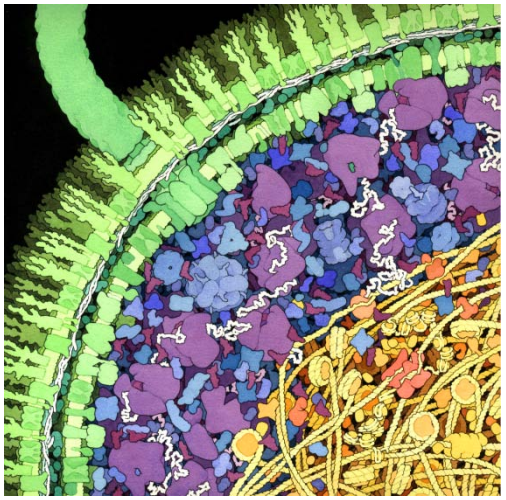
- 图作为基本工具用来强调相互作用并直观表示复杂的生物系统
- 节点代表生物分子，边代表他们之间在生命过程中的某种关系



1D: Complete Genetic Partslist



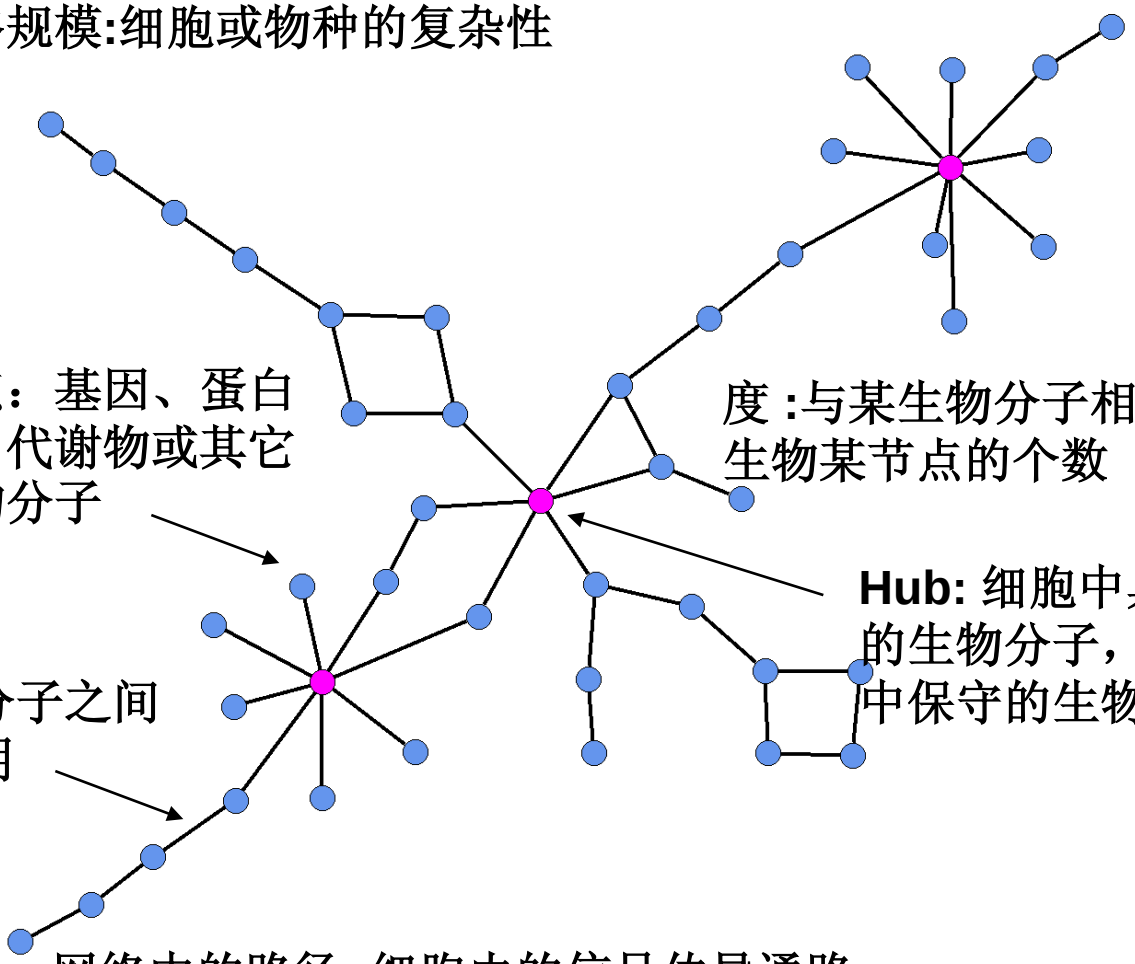
~2D: Bio-molecular Network



3D: Detailed structural understanding of cellular machinery

生物学语言下的网络

网络规模:细胞或物种的复杂性



节点：基因、蛋白质、代谢物或其它生物分子

度 :与某生物分子相互作用的其它生物某节点的个数

Hub: 细胞中具有重要功能的生物分子，或者是进化中保守的生物分子

边：生物分子之间的相互作用

网络中的路径: 细胞内的信号传导通路

At molecular level

Emerging area

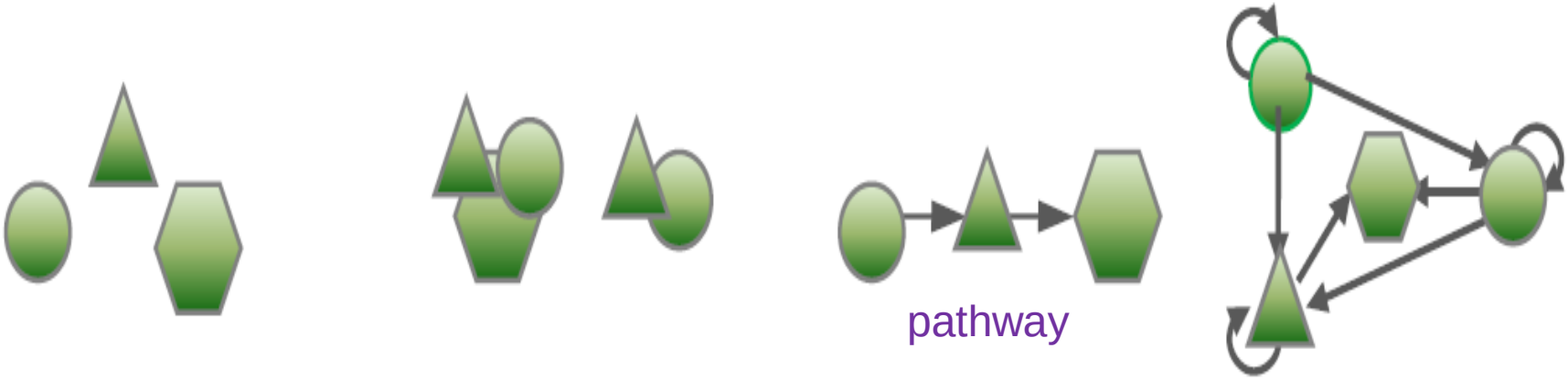
Network Systems Biology?

Instead of analyzing individual components or aspects of an organism,

We study how an organism, viewed as a **dynamical network** of biomolecules and biochemical reactions, eventually gives rise to a complex life.

Main ingredients ?

Molecules, Interactions, Pathways, Networks



Individual molecules ⇨ Pairwise interactions ⇨ Local structures ⇨ Global networks

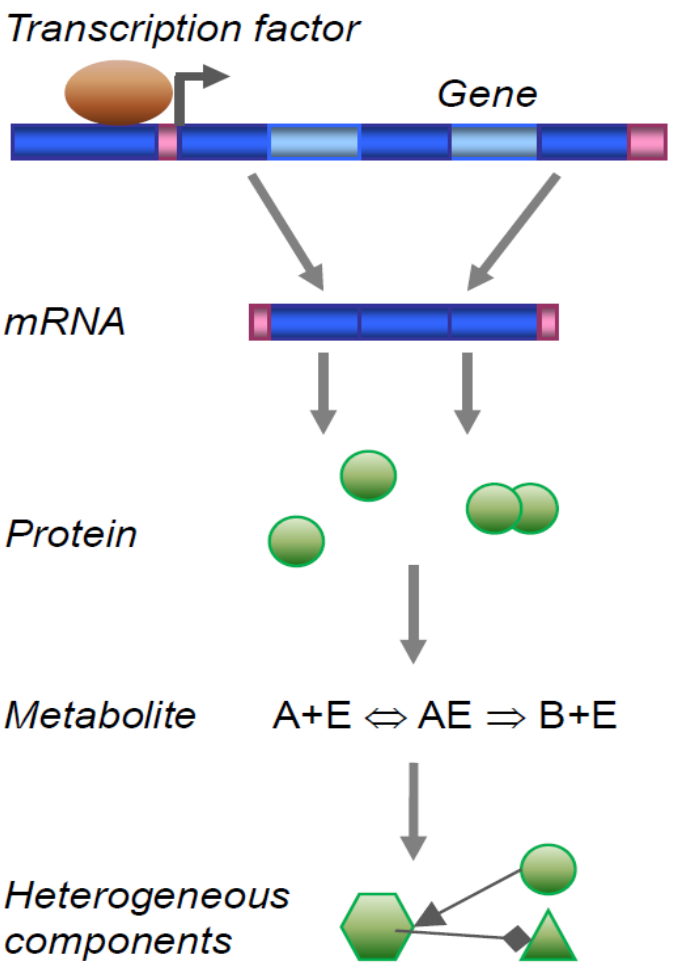
Local → Global

Hierarchical Relations

Biological networks

- Elements: gene, proteins, etc. Node
- Interactions: Regulation, physical interaction, genetic interaction, metabolic reaction, modification, etc. Edges
- Local structures: Motif, community, functional module, complex, etc.
- Global structures: Scale-free, small world, etc.

Biomolecular Networks ?



Transcription regulatory network
(DNA-TF interactions)

Gene regulatory network
(gene-gene interactions)

Protein interaction network
(protein-protein interactions)

Metabolic network
(enzyme-substrate interactions)

Signaling network
(molecule-molecule interactions)

Molecules vs. **Networks**

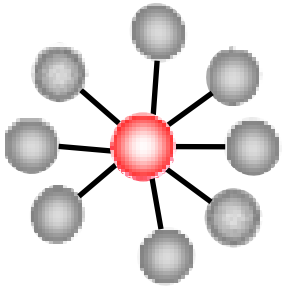
Central Dogma of Biology

Types of Biological Networks

Genetic network: Interactions between genes, gene products, small molecules

Transcriptional regulatory network: Network of control decisions to turn genes on / off, Subset of genetic network

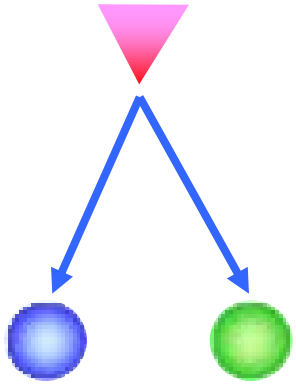
Signal transduction network: Network of the movement of signals from outside the cell to inside.



Undirected

Protein-protein interaction network: Physical or genetic network

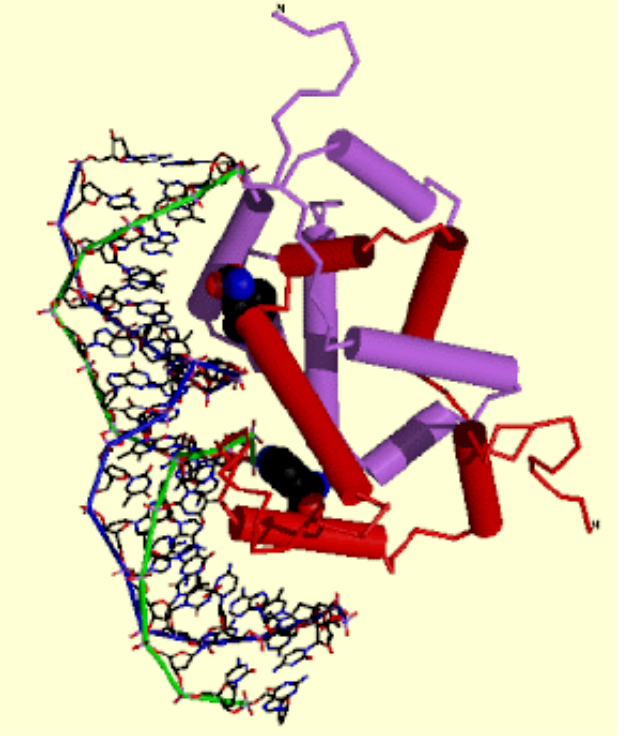
Metabolic network: Network of interactions between proteins, Synthesize / break down molecules (enzymes, cofactors)

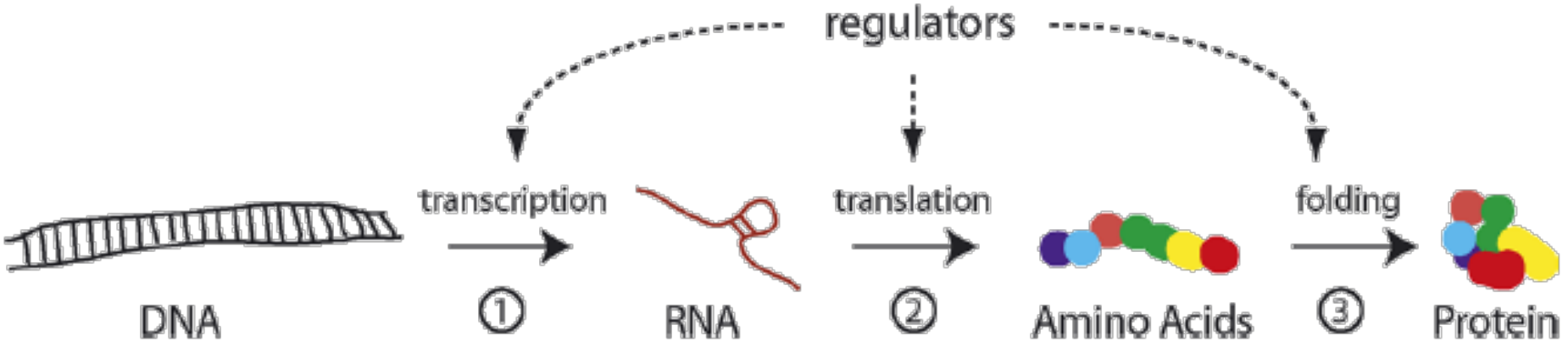


Directed

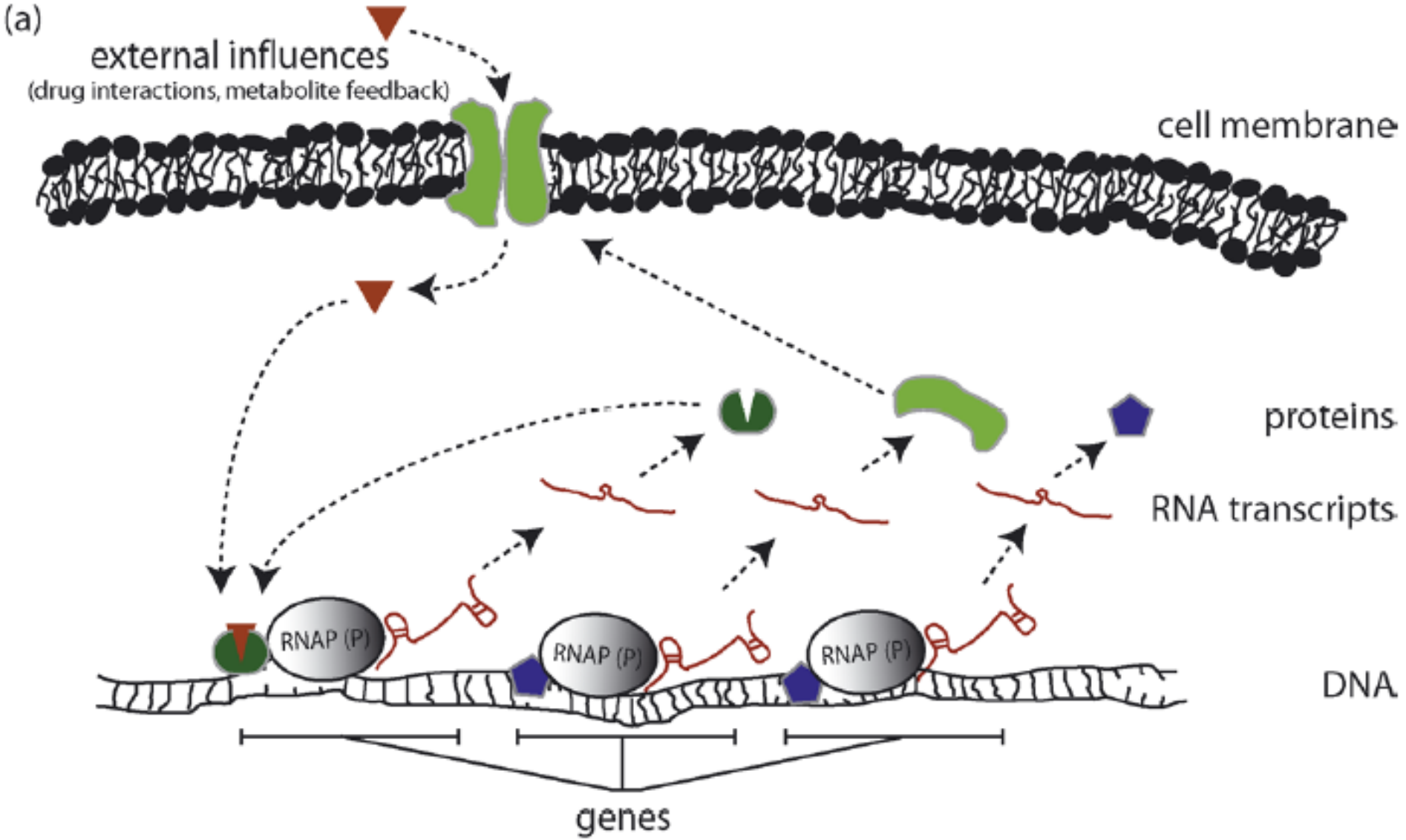
Others: Gene co-expression network, Functional linkage network, protein structure network, protein folding network, neural network, Domain interaction network

Regulation

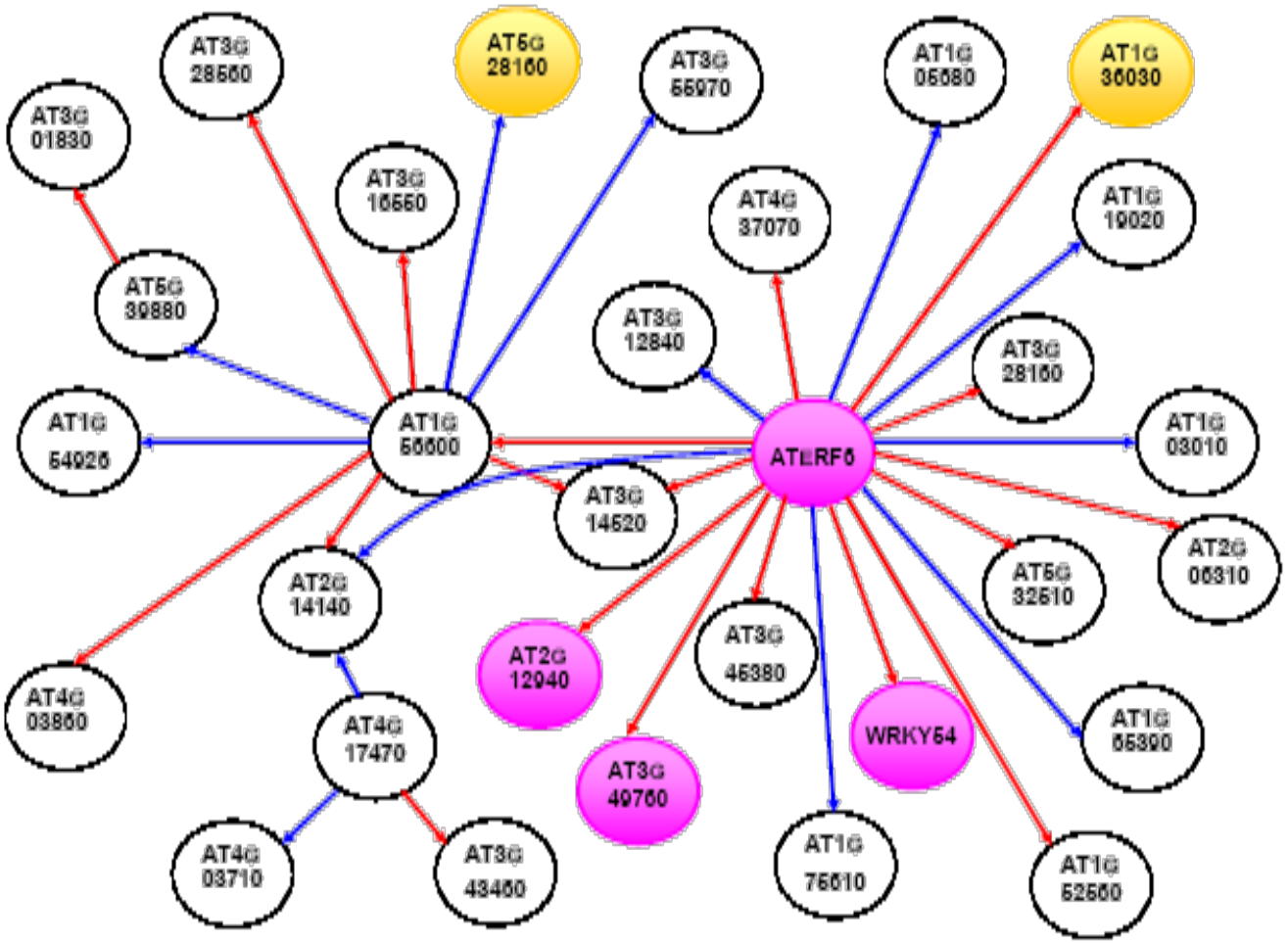



regulators
transcription
translation
folding
DNA
①
RNA
②
Amino Acids
③
Protein

Direct and indirect regulation



Genetic network



Node: gene

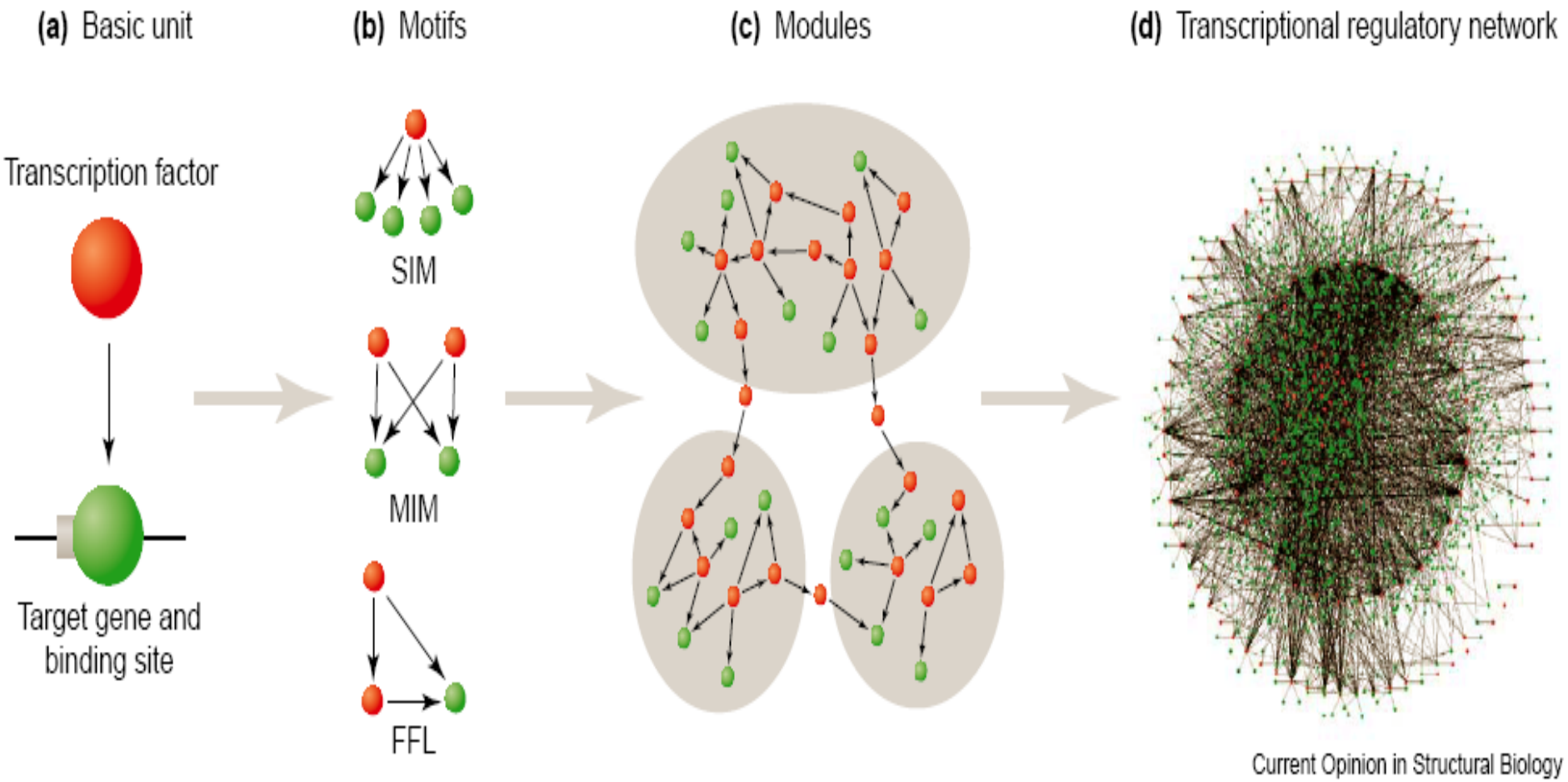
Edge: causality regulatory relationships

Directed

Edge weighted

Genetic network consists of set of genes, proteins, small molecules, and their mutual regulatory interactions。

Transcriptional regulatory network

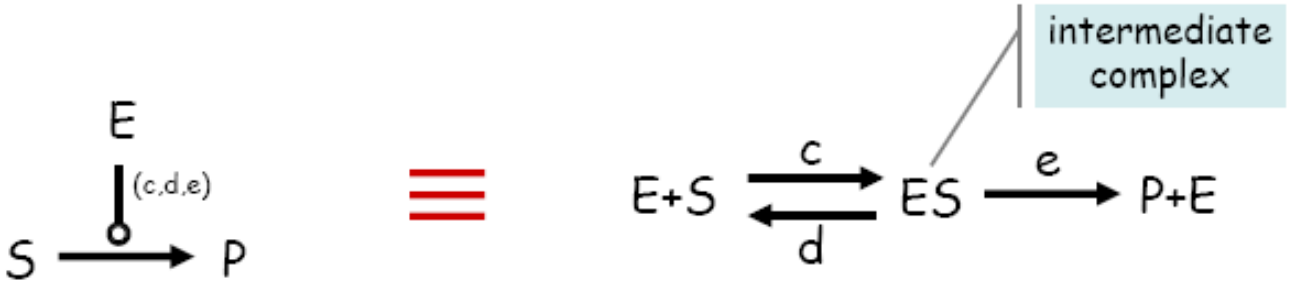


Subset of genetic network

Node: TF and genes, Edge: regulation relationships
Directed, Edge weighted

Enzymatic Reactions

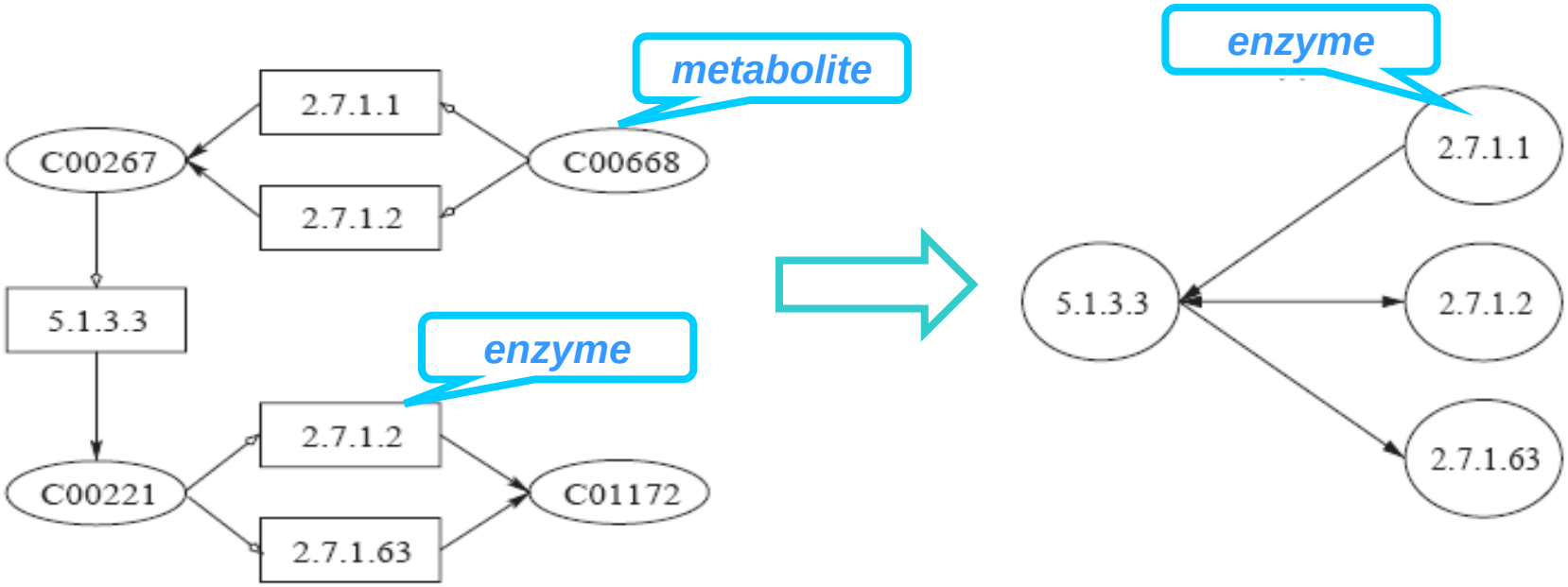
Reaction View

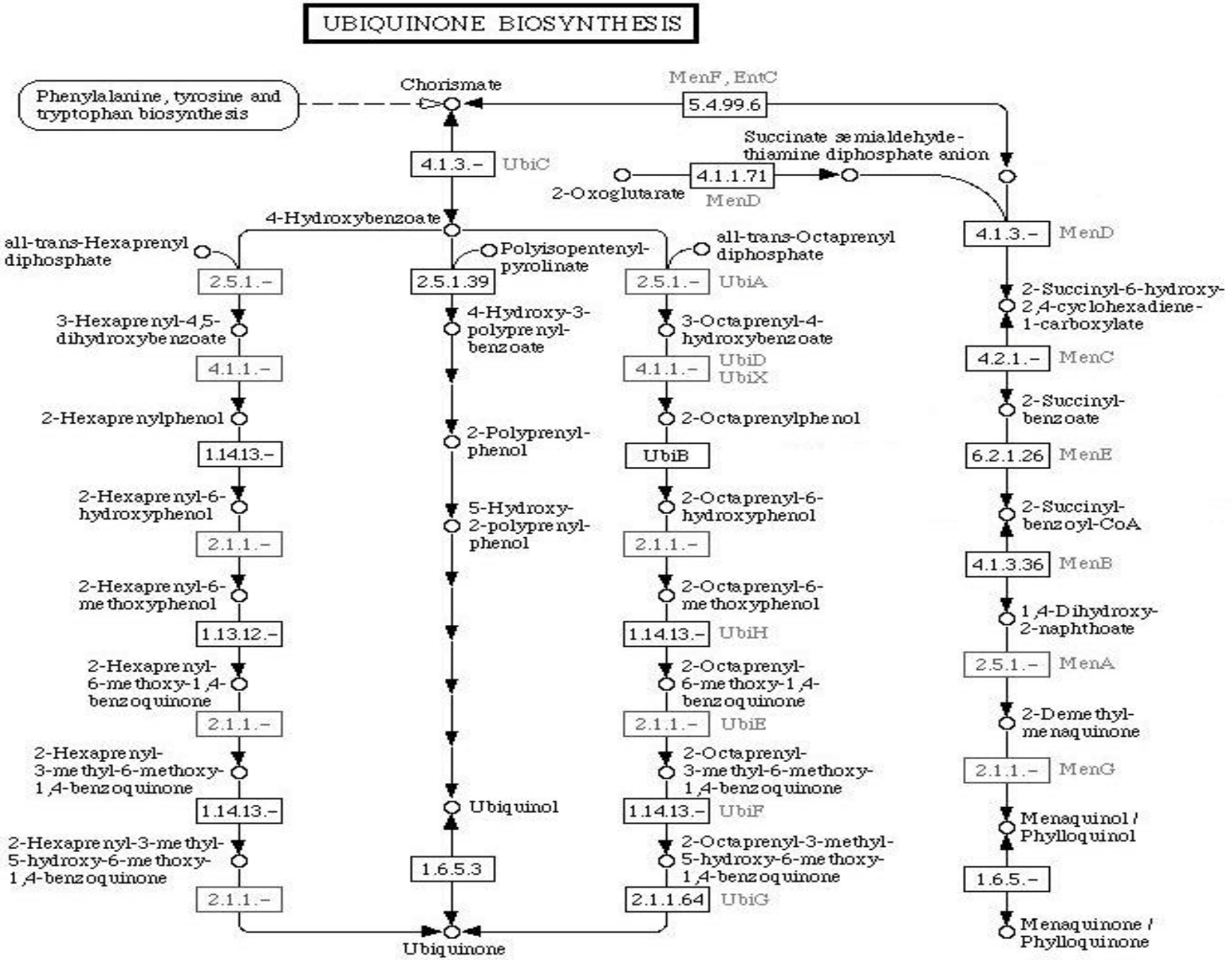


- **Definition of Metabolic network G*(V,E)* :**
  - **For every enzyme z_i in Z - a node v_i exists**
  - **(v_i,v_j) in E iff z_j consumes the product of z_i**

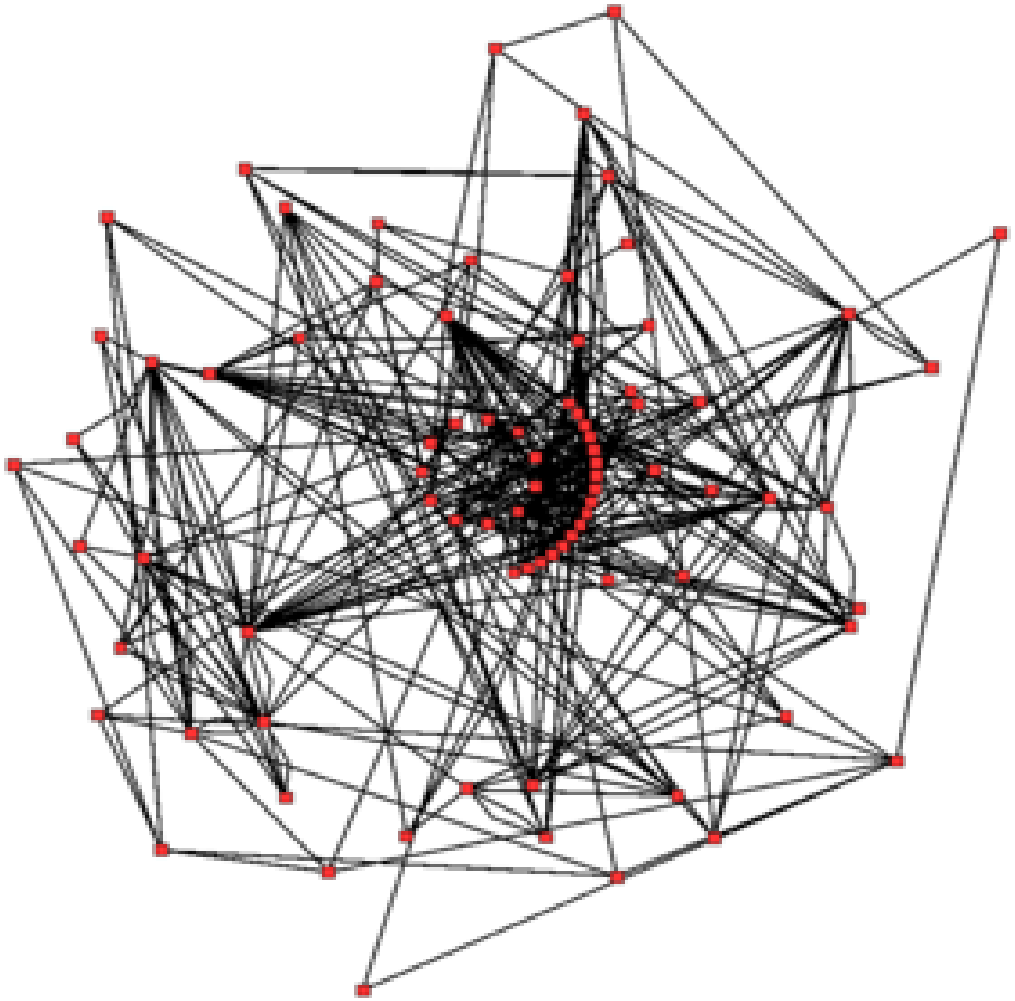
  酶A的一个产物是酶B 的一个底物

- **Example:**




UBIQUINONE BIOSYNTHESIS
Phenylalanine, tyrosine and tryptophan biosynthesis
Chorismate
MenF, EntC
5.4.99.6
4.1.3.- UbiC
Succinate semialdehyde-thiamine diphosphate anion
2-Oxoglutarate
4.1.1.71 MenD
4-Hydroxybenzoate
4.1.3.- MenD
all-trans-Hexaprenyl diphosphate
Polyisopentenyl-pyrolinate
all-trans-Octaprenyl diphosphate
2.5.1.-
2.5.1.39
2.5.1.- UbiA
2-Succinyl-6-hydroxy-2,4-cyclohexadiene-1-carboxylate
3-Hexaprenyl-4,5-dihydroxybenzoate
4-Hydroxy-3-polyprenyl-benzoate
3-Octaprenyl-4-hydroxybenzoate
4.2.1.- MenC
4.1.1.-
4.1.1.- UbiD UbiX
2-Succinyl-benzoate
2-Hexaprenylphenol
2-Polyprenyl-phenol
2-Octaprenylphenol
6.2.1.26 MenE
1.14.13.-
UbiB
2-Hexaprenyl-6-hydroxyphenol
5-Hydroxy-2-polyprenyl-phenol
2-Octaprenyl-6-hydroxyphenol
2-Succinyl-benzoyl-CoA
2.1.1.-
2.1.1.-
4.1.3.36 MenB
2-Hexaprenyl-6-methoxyphenol
2-Octaprenyl-6-methoxyphenol
1,4-Dihydroxy-2-naphthoate
1.13.12.-
1.14.13.- UbiH
2.5.1.- MenA
2-Hexaprenyl-6-methoxy-1,4-benzoquinone
2-Octaprenyl-6-methoxy-1,4-benzoquinone
2-Demethyl-menaquinone
2.1.1.-
2.1.1.- UbiE
2.1.1.- MenG
2-Hexaprenyl-3-methyl-6-methoxy-1,4-benzoquinone
2-Octaprenyl-3-methyl-6-methoxy-1,4-benzoquinone
Menaquinol / Phylloquinol
1.14.13.-
Ubiquinol
1.14.13.- UbiF
1.6.5.-
2-Hexaprenyl-3-methyl-5-hydroxy-6-methoxy-1,4-benzoquinone
2-Octaprenyl-3-methyl-5-hydroxy-6-methoxy-1,4-benzoquinone
1.6.5.3
2.1.1.-
2.1.1.64 UbiG
Menaquinone / Phylloquinone
Ubiquinone

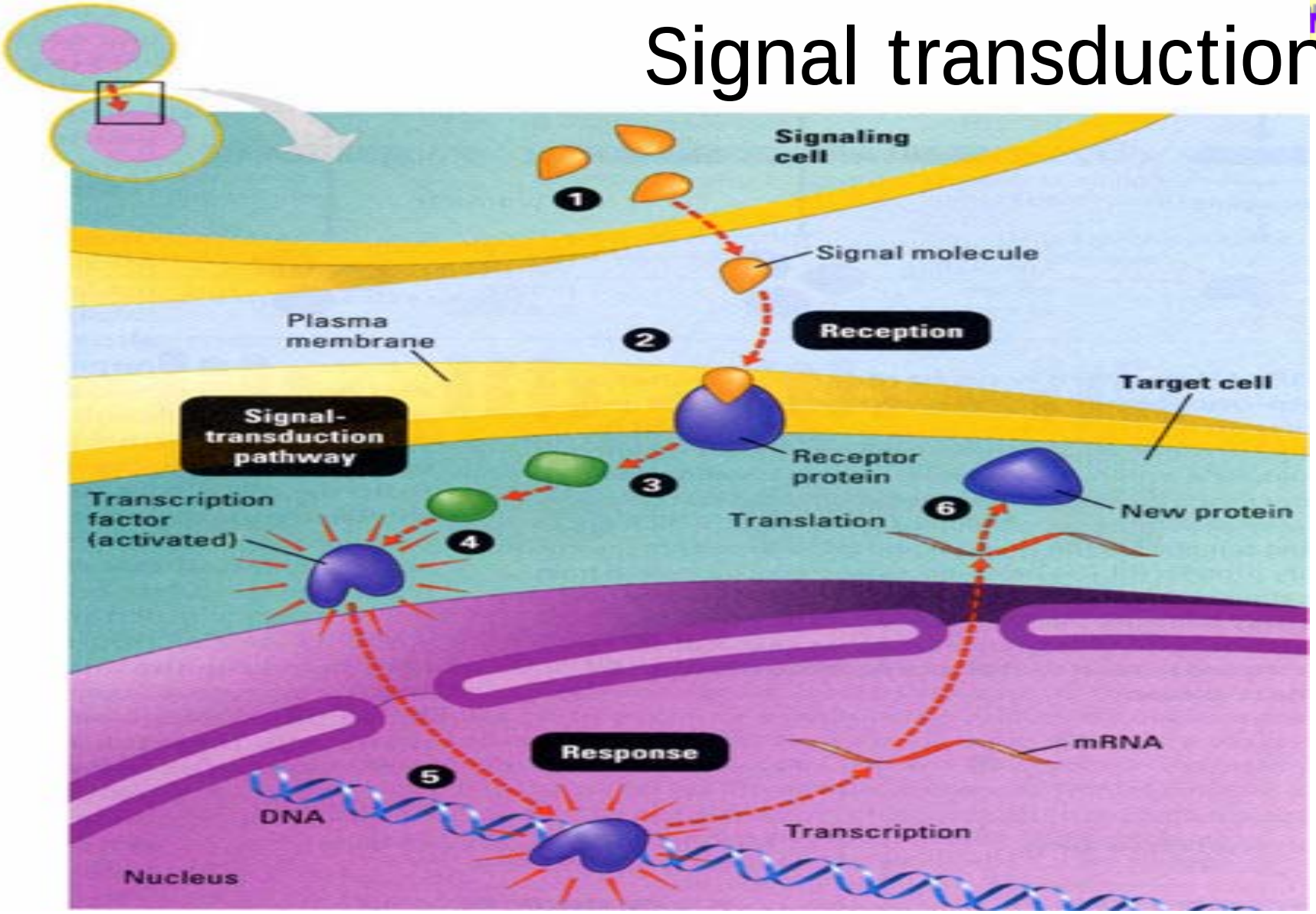
Metabolic network



Representation 1:
Node: metabolites
Edge: enzymatic steps
Directed
Unweighted

Representation 2:
Node: Enzymes
Edge: Enzymes working on adjacent steps
Undirected
Unweighted

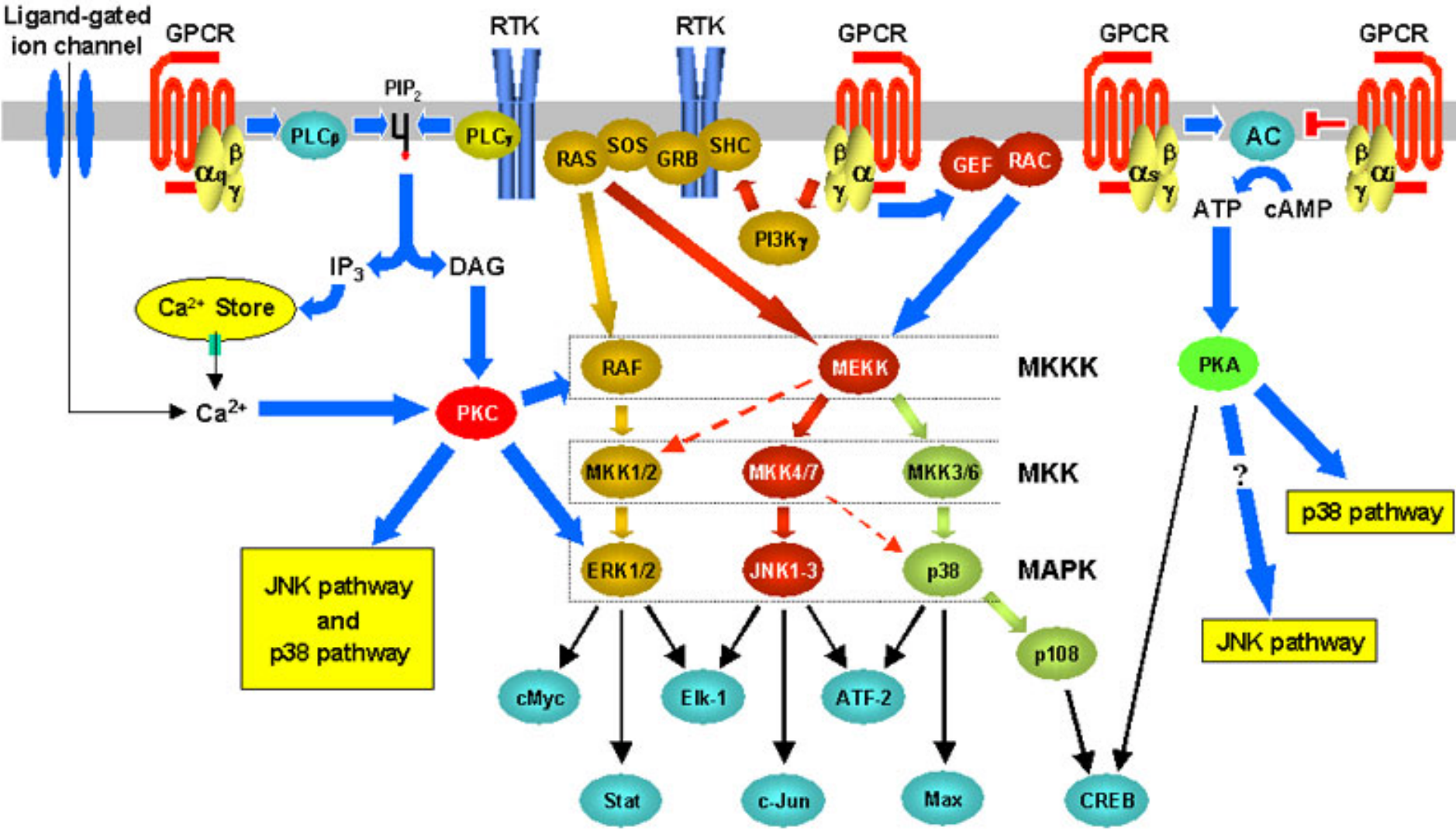
Signal transduction

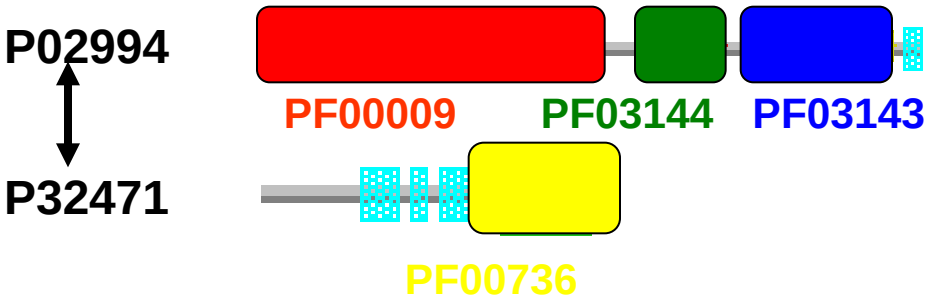


Plants have developed sophisticated signal transduction mechanisms to be able to respond to changing environmental conditions.

Signal transduction network

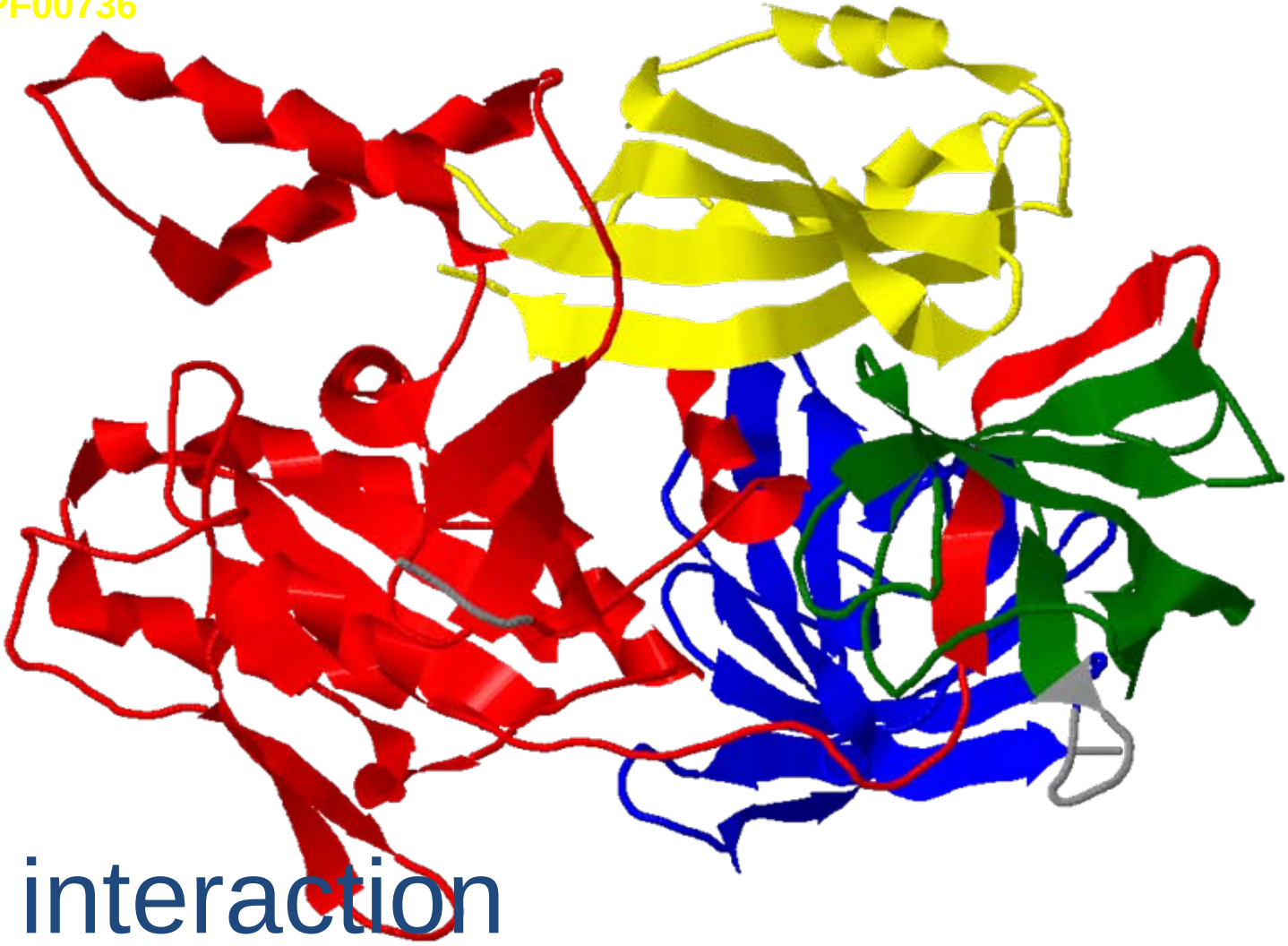
Node: proteins, signal molecules, Edge: interaction relationships, Undirected, Unweighted





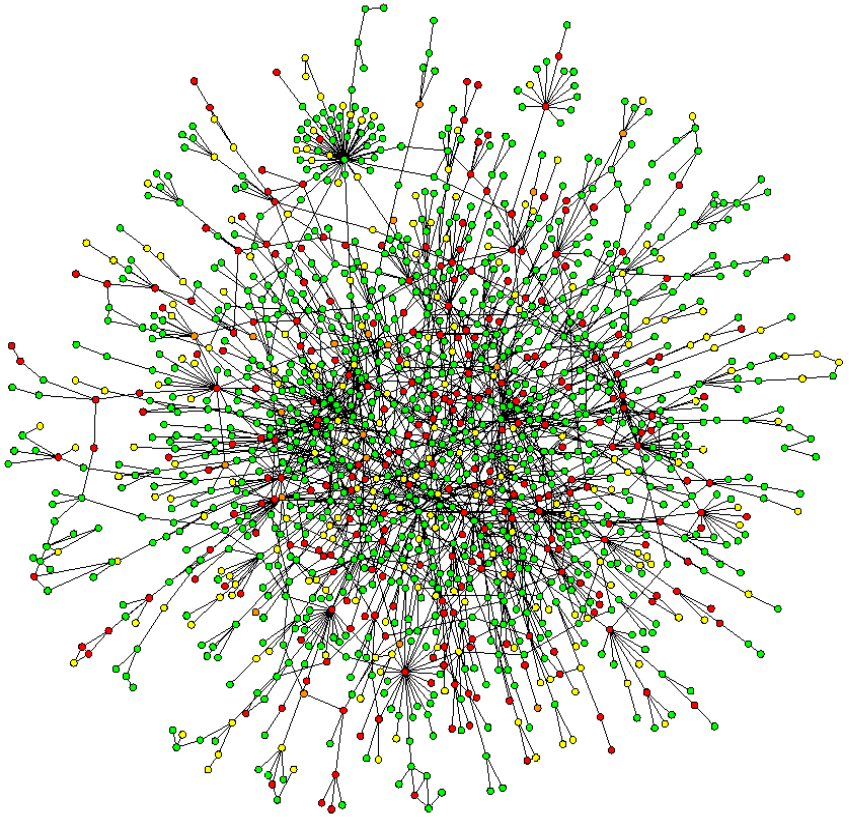
YBR118W←→YAL003W PF03144 PF00009 ← → PF00736
YBR118W←→YAL003W PF03143 PF00009 ← → PF00736
YBR118W←→YAL003W PF03143 PF03144 ← → PF00736

YPR080W←→YAL003W PF03144 PF00009 ← → PF00736
YPR080W←→YAL003W PF03143 PF00009 ← → PF00736
YPR080W←→YAL003W PF03143 PF03144 ← → PF00736



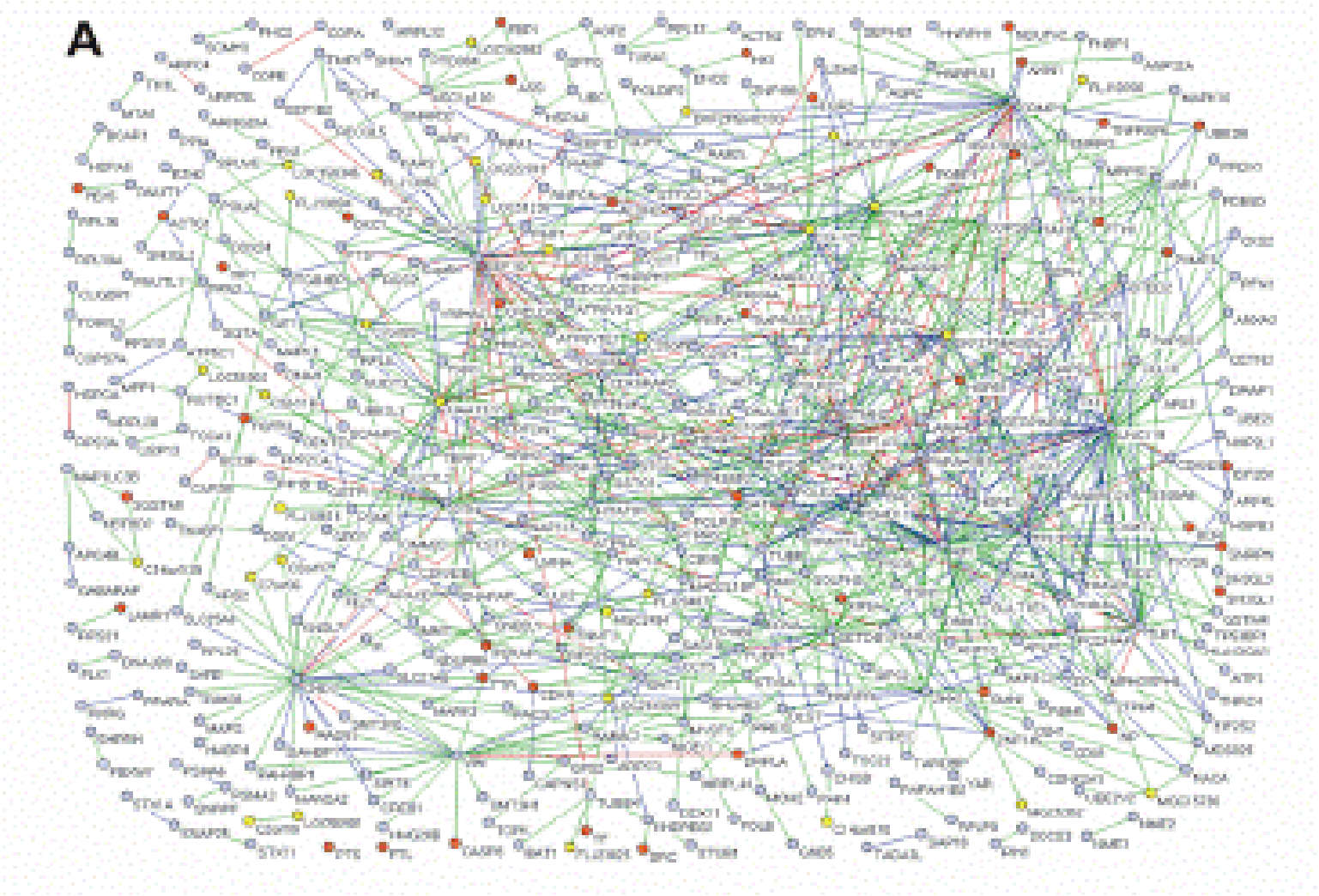
Protein interaction

Protein Interaction Network

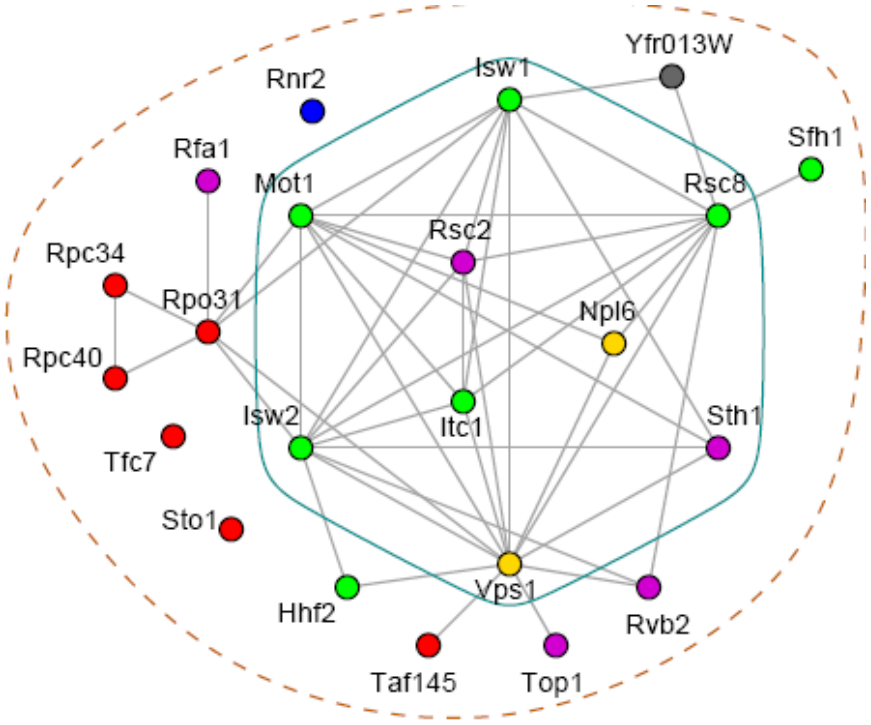


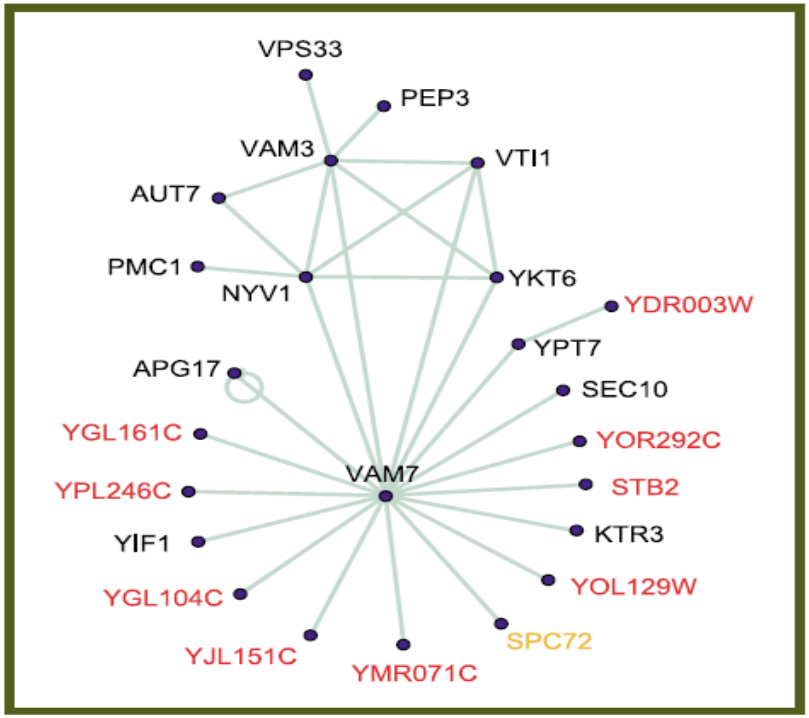
Yeast protein interaction network

Protein interaction network in Human by yeast two-hybrid method, 2006. Protein number: 1705, links：3186



Representation of PPI network





Node: proteins, Edge: interaction relationships, Undirected, Weighted (Binary or Strengthed)

生物分子网络研究的科学问题

- **如何构建网络？**

建模，数据处理、集成

- **如何分析网络？**

静态：结构，功能等

动态：不同条件，进化