



生物信息学与系统生物学

张世华

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 1012/1003
- 对课程的意见和建议

邮件:

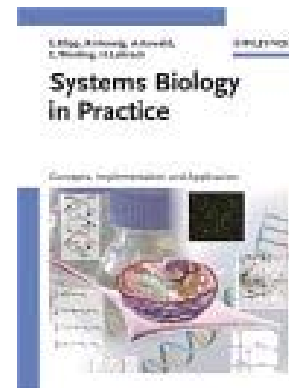
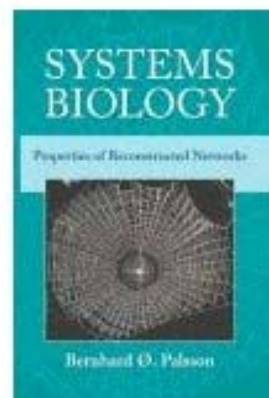
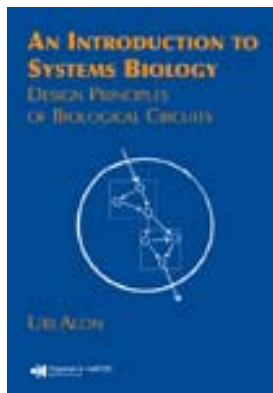
zsh@amss.ac.cn;

时间: 2012.2.22---4.27, 每周三、五下午13:30-16:10

地点: 中科院研究生院中关村教学楼N215教室

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

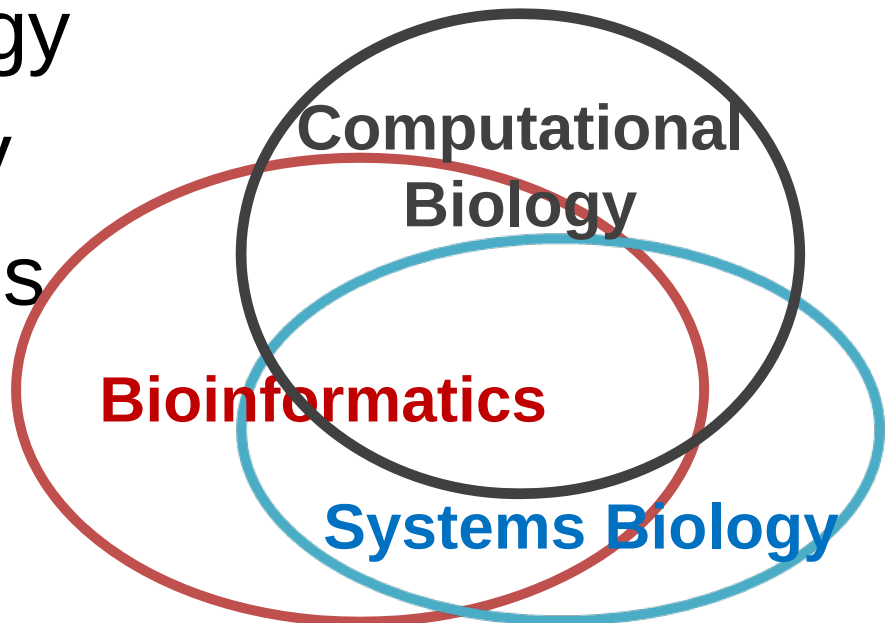


State-of-Arts

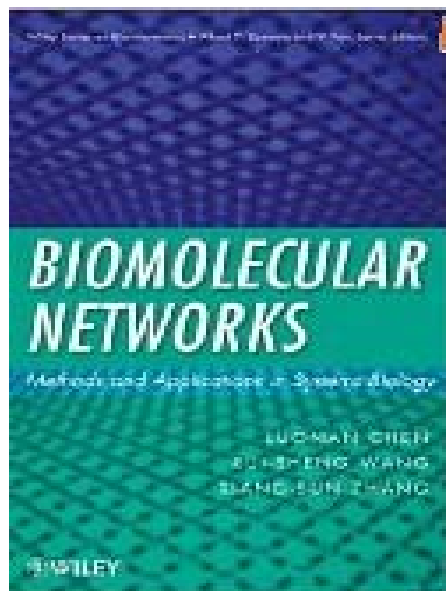
- Science, Nature, Cell, PNAS
- Molecular Systems Biology

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

- Google, Wiki



We focus on biomolecular networks as well as highly related topics



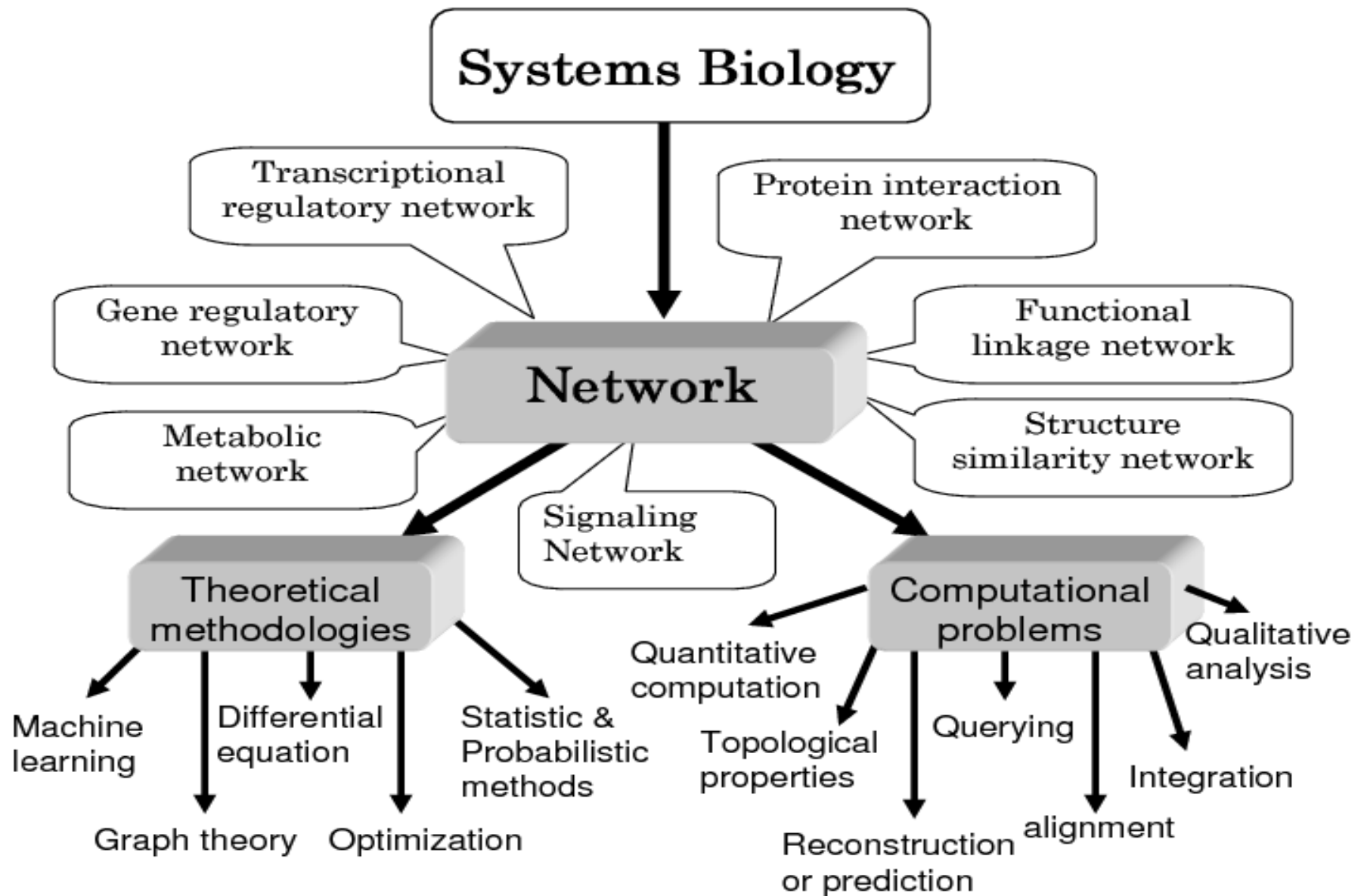
Luonan Chen, Rui-Sheng Wang, Xiang-Sun Zhang.

[Biomolecular Networks: Methods and Applications in Systems Biology.](#)

[John Wiley & Sons](#), Hoboken, New Jersey. July, 2009.

6





Network Systems Biology



大纲

如何分析网络? (张世华)

1. 生物分子网络分析
2. 生物分子网络motif, 模块分析
3. 生物分子网络比对
4. 生物活性通路和网络标记物识别

如何构建网络? (王勇)

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



Chinese

ZHANGGroup
Bioinformatics

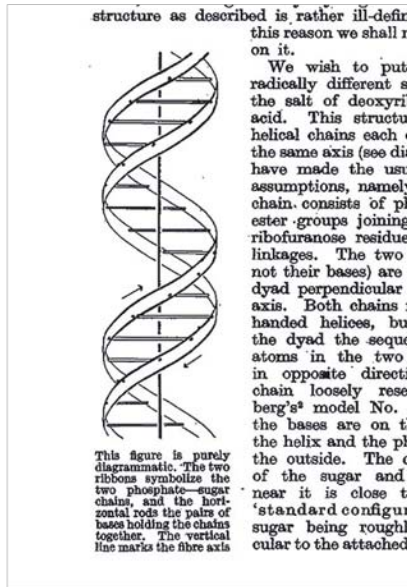


<http://zhangroup.aporc.org>
Chinese Academy of Sciences



系统生物学

《Nature》 1953

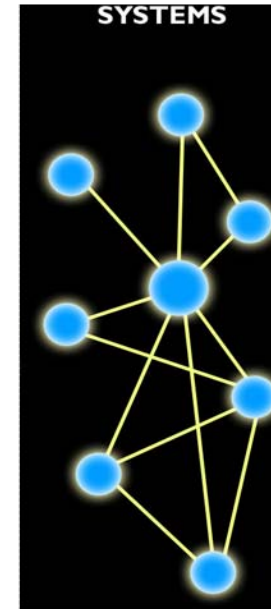


DNA双螺旋结构发现,
开启分子生物学研究

《Science》 2001

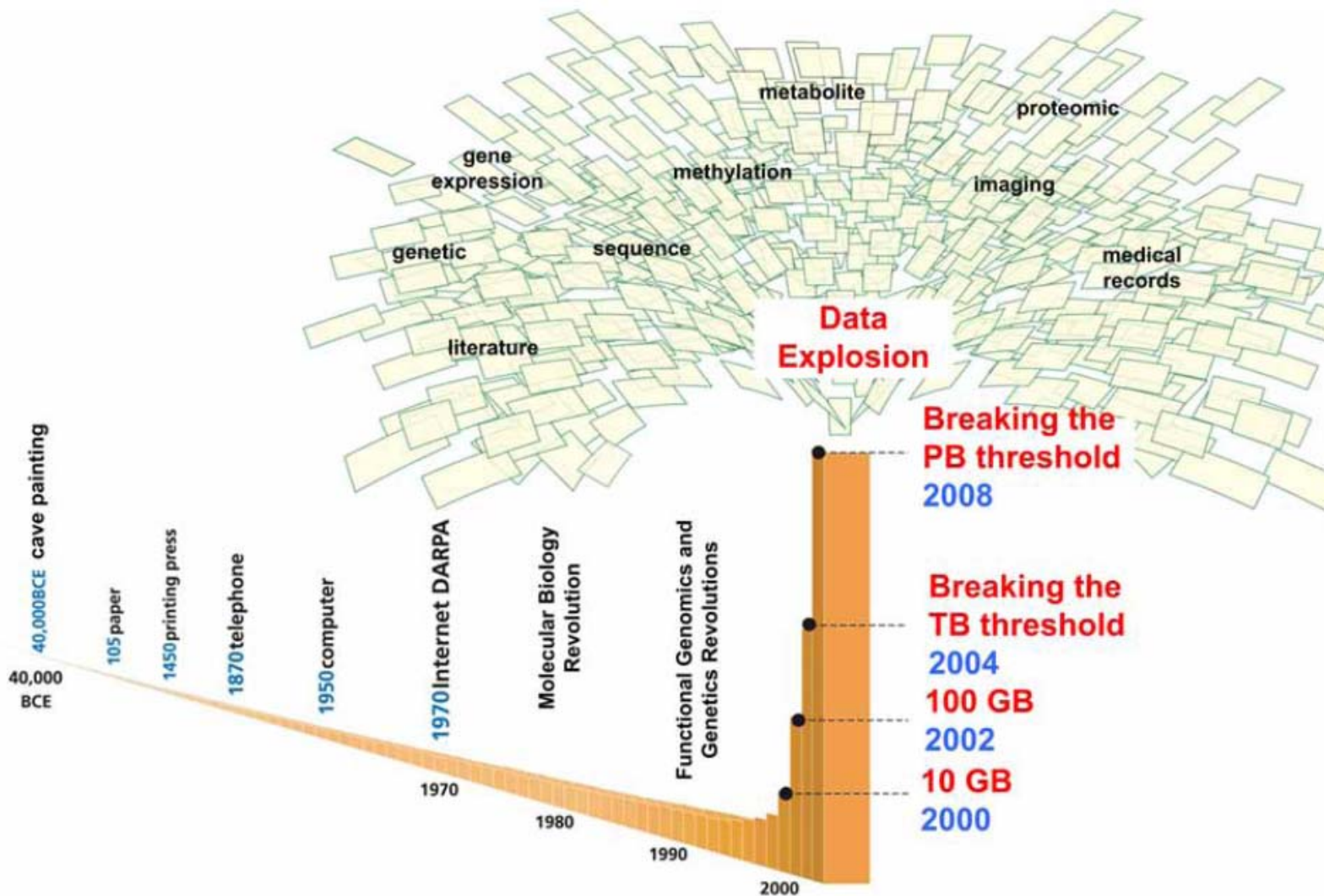


人类基因组计划完成



系统生物学





<http://www.sagebase.org>



Search:
PubMed

Limits
Advanced search
Help

systems biology

Search
Clear



PubMed

PubMed comprises more than 20 million citations for biomedical literature from MEDLINE, life science journals, and online books. Citations may include links to full-text content from PubMed Central and publisher web sites.

Using PubMed

- [PubMed Quick Start Guide](#)
- [Full Text Articles](#)
- [PubMed FAQs](#)
- [PubMed Tutorials](#)
- [New and Noteworthy](#)

PubMed Tools

- [PubMed Mobile](#)
- [Single Citation Matcher](#)
- [Batch Citation Matcher](#)
- [Clinical Queries](#)
- [Topic-Specific Queries](#)

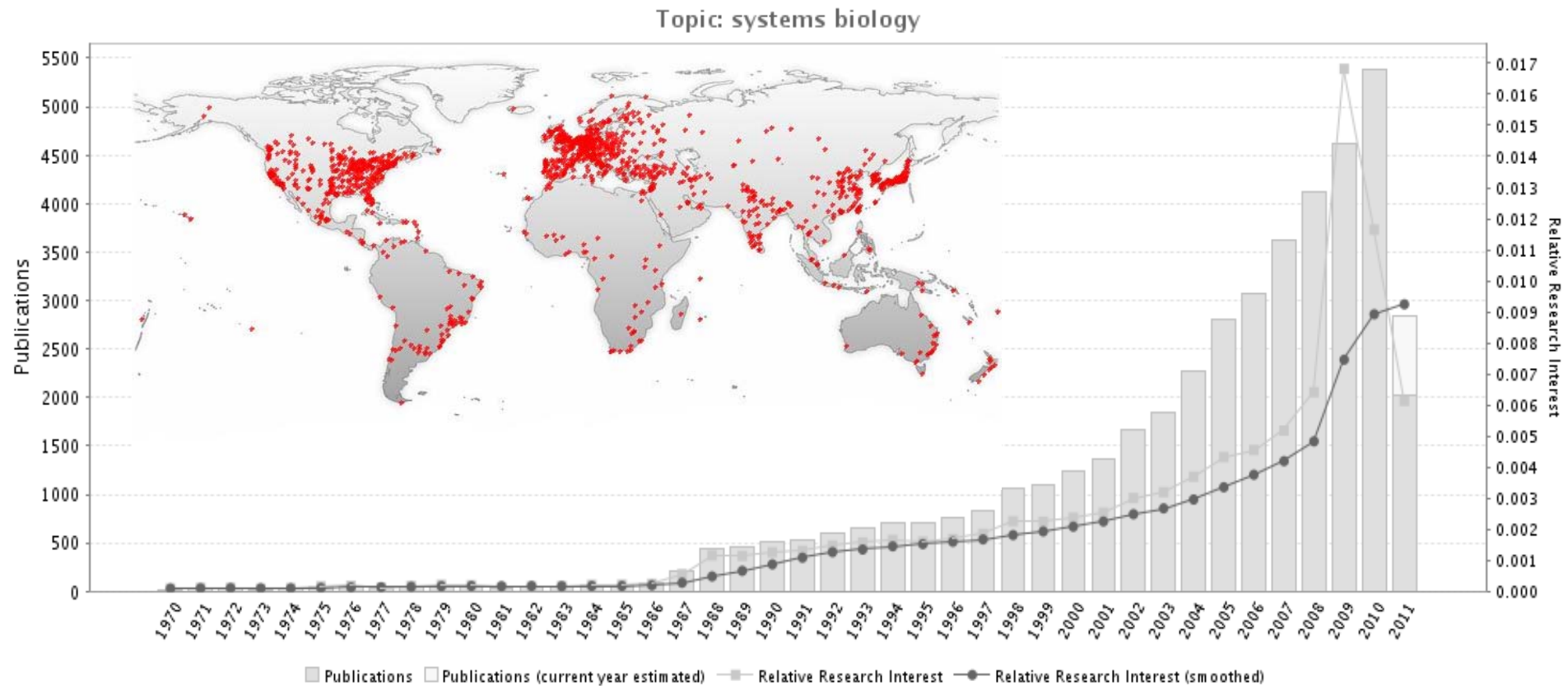
More Resources

- [MeSH Database](#)
- [Journals in NCBI Databases](#)
- [Clinical Trials](#)
- [E-Utilities](#)
- [LinkOut](#)

PubMed---Gopubmed

- ◆ 对于从事医学、生物学方面研究的人员，在NCBI上用**Pubmed**检索文章是大家最常用的一种手段。
- ◆ 如果你的检索式在PubMed的结果中出现成千上万条相关文章时，那应该是一件非常让人苦恼的事情，因为这意味着要找到你想要的信息及如同大海捞针一般。
- ◆ Gopubmed是一种Pubmed检索工具，是对于Pubmed资源的另一种组合检索方式，它比Pubmed的优点集中在于对于某一方面的研究我们能够直观得知研究时间、研究者、地域、研究过程、研究动态和各种信息之间的相关性。
- ◆ 输入你要检索的关键词，检索之后，可以根据侧边栏的what、who、where、when进行分类。这样就可以按照你的要求筛选出你最需要的文献了。<http://www.gopubmed.org/>

GoPubMed



Google Scholar

Google 学术搜索 [学术高级搜索](#)
☒ 搜索所有网页 ☐ 中文网页 ☐ 简体中文网页

学术搜索 时间不限

小提示：[只搜索中文\(简体\)结果](#)，可在 [学术搜索设置](#) 指定搜索语言

[Systems biology: a brief overview](#)

H Kitano - Science, 2002 - [sciencemag.org](#)

Abstract To understand **biology** at the system level, we must examine the structure and dynamics of cellular and organismal function, rather than the characteristics of isolated parts of a cell or organism. Properties of **systems**, such as robustness, emerge as central issues, ...

被引用次数：2107 - [相关文章](#) - [所有 58 个版本](#)

[uu.nl 中的 \[PDF\]](#)

[The systems biology markup language \(SBML\): a medium for representation and exchange of biochemical network models](#)

M Hucka, A Finney, HM Sauro, H Bolouri... - ..., 2003 - Oxford Univ Press

Abstract Motivation: Molecular biotechnology now makes it possible to build elaborate **systems** models, but the **systems biology** community needs information standards if models are to be shared, evaluated and developed cooperatively. Results: We summarize the ...

被引用次数：1398 - [相关文章](#) - [所有 59 个版本](#)

[oxfordjournals.org 中的 \[PDF\]](#)

[\[PDF\] Computational systems biology](#)

H Kitano - Nature, 2002 - [people.mbi.ohio-state.edu](#)

To understand complex biological **systems** requires the integration of experimental and computational research—in other words a **systems biology** approach. Computational **biology**, through pragmatic modelling and theoretical exploration, provides a powerful ...

被引用次数：1146 - [相关文章](#) - [HTML 版](#) - [所有 56 个版本](#)

[ohio-state.edu 中的 \[PDF\]](#)

[Evolving a lingua franca and associated software infrastructure for computational systems biology: the Systems Biology Markup Language \(SBML\) project](#)

M Hucka, A Finney, BJ Bornstein... - Systems ..., 2004 - [authors.library.caltech.edu](#)

Biologists are increasingly recognising that computational modelling is crucial for making sense of the vast quantities of complex experimental data that are now being collected. The **systems biology** field needs agreed-upon information standards if models are to be shared ...

被引用次数：145 - [相关文章](#) - [所有 19 个版本](#)

[caltech.edu 中的 \[PDF\]](#)

[A new approach to decoding life: systems biology](#)

T Ideker, T Galitski... - Annual review of genomics and ..., 2001 - [annualreviews.org](#)

Abstract **Systems biology** studies biological **systems** by systematically perturbing them (biologically, genetically, or chemically); monitoring the gene, protein, and informational pathway responses; integrating these data; and ultimately, formulating mathematical ...

被引用次数：980 - [相关文章](#) - [所有 57 个版本](#)

[buscario.com.br 中的 \[PDF\]](#)



Philip Bourne

Professor of Pharmacology, UCSD

[Computational Biology - Scholarly Communication](#)

Verified email at [ucsd.edu](#)

[Homepage](#)

Google scholar

Search Authors

[Get my own profile - Help](#)

Follow this author

[Follow new articles](#)
[Follow new citations](#)

Co-authors

Helen M. Berman
Lei Xie
Wolfgang Bluhm
Li Xie
Eric Scheeff
Ruben E. Valas
Bjoern Peters
Alessandro Sette
Andreas Prlc
Peter W. Rose
Adam Godzik
Sarah Kinnings
Ole Lund
Roland Dunbrack
Lawrence Hunter
Evangelia Petsalaki
Michael Kuhn
Rob Russell
Timothy A Holland
Yohan Kim
Herbert J. Bernstein

Citation indices

	All	Since 2007
Citations	20403	11248
h-index	40	32
i10-index	103	81

Citations to my articles



Select: [All](#), [None](#) [Export](#)

Show: 20 [1-20](#) [Next >](#)

Title / Author	Cited by	Year
The protein data bank		
☐ HM Berman, J Westbrook, Z Feng, G Gilliland, TN Bhat, H Weissig, IN ... Nucleic acids research 28 (1), 235-242	13295	2000
Protein structure alignment by incremental combinatorial extension (CE) of the optimal path.		
☐ IN Shindyalov, PE Bourne Protein engineering 11 (9), 739-747	1516	1998
The protein data bank		
☐ HM Berman, T Battistuz, TN Bhat, WF Bluhm, PE Bourne, K Burkhardt, Z Feng ... Acta Crystallographica Section D: Biological Crystallography 58 (6), 899-907	686	2002
The RCSB Protein Data Bank: a redesigned query system and relational database based on the mmCIF schema		

Welcome to Google Scholar Citations

Track citations to your publications

Check who is citing your publications. Graph your citations over time. Compute citation metrics.

View publications by colleagues

Keep up with their work. See their citation metrics.

Appear in Google Scholar search results

Create a public profile that can appear in Google Scholar when someone searches for your name.



Shihua Zhang

Assistant Professor of Academy of Mathematics and Systems Science, Chinese Academy of Science

Bioinformatics - Computational Biology - Network Biology - Algorithm - Data Mining

Verified email at amss.ac.cn

[Homepage](#)

Google scholar

[Get my own profile](#) - [Help](#)

Follow this author

[Follow new articles](#)
[Follow new citations](#)

Co-authors

No co-authors

Citation indices

	All	Since 2007
Citations	592	591
h-index	13	13
i10-index	14	14

Citations to my articles



Select: All, None

Show: 20 [Next >](#)

Title / Author	Cited by	Year
☐ Identification of overlapping community structure in complex networks using fuzzy c-means clustering S Zhang, RS Wang, XS Zhang Physica A: Statistical Mechanics and its Applications 374 (1), 483-490	141	2007
☐ Quantitative function for community detection Z Li, S Zhang, RS Wang, XS Zhang, L Chen Physical Review E 77 (3), 036109	67	2008
☐ Discovering functions and revealing mechanisms at molecular level from biological networks S Zhang, G Jin, XS Zhang, L Chen Proteomics 7 (16), 2856-2869	57	2007
☐ Alignment of molecular networks by integer quadratic programming Z Li, S Zhang, Y Wang, XS Zhang, L Chen Bioinformatics 23 (13), 1631-1639	55	2007
☐ Uncovering fuzzy community structure in complex networks S Zhang, RS Wang, XS Zhang	41	2007

h-index

h-index，又称为h指数或h因子（h-factor），是一种评价学术成就的新方法。h代表“高引用次数”（high citations），一名科研人员的h指数是指他至多有h篇论文分别被引用了至少h次。h指数能够比较准确地反映一个人的学术成就。一个人的h指数越高，则表明他的论文影响力越大

什么是系统生物学？

- 内涵？

- 外延？

- 问题？

Systems Biology

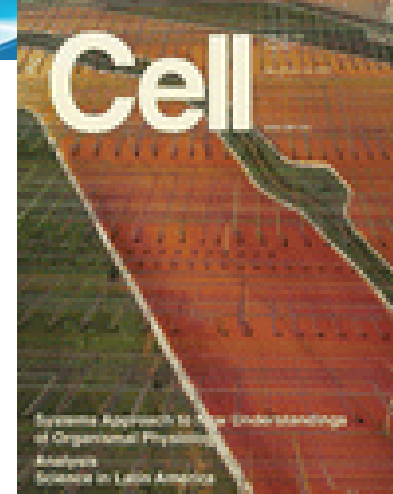
- **Networking** the whole biological system, rather than studying its isolated parts.
- **Integrating** large amounts of data in the context of biological network (Sequence, structure, function, gene expression, protein expression, protein interaction, protein-DNA interaction, and literature data).

Systems Biology—Procedure

- 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
 POST-PUBLICATION PEER REVIEW

What is F1000? Take the tour.

[Evaluated Articles](#)
[Rankings](#)
[Reports](#)
[Faculty](#)
[Naturally Selected](#)

[MyF1000](#)
[About](#)
[Free Trial](#)
[Subscribe](#)

10% IVIG: Good efficacy and tolerability at high infusion rates

CSL Behring

Biotherapies for Life™

[Learn More](#)


ADVERTISEMENT

Search

BIOLOGY AND MEDICINE ARTICLES SELECTED BY 10,000 EXPERTS



☒ F1000
 ☐ PubMed

[Advanced Search](#)

What is F1000?



MYF1000



Customise alerts from your specialty and save evaluations and articles to your personal library

Browse

FIND YOUR FACULTY

[Anesthesiology & Pain Management](#)

[Biochemistry](#)

[Bioinformatics & Computational Biology](#)

[Biotechnology](#)

[Cancer Biology](#)

[Cardiovascular Biology](#)

[Gastroenterology & Hepatology](#)

[Gastrointestinal Biology](#)

[Genomics & Genetics](#)

[Hematology](#)

[Immunology](#)

[Infectious Diseases](#)

[Metabolic & Endocrine Science](#)

[Otolaryngology](#)

[Pharmacology & Drug Discovery](#)

[Physiology](#)

[Plant Biology](#)

[Psychiatry](#)

[Public Health & Epidemiology](#)

[Renal Biology](#)



BIO-RAD

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 did they study?

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

Methodology

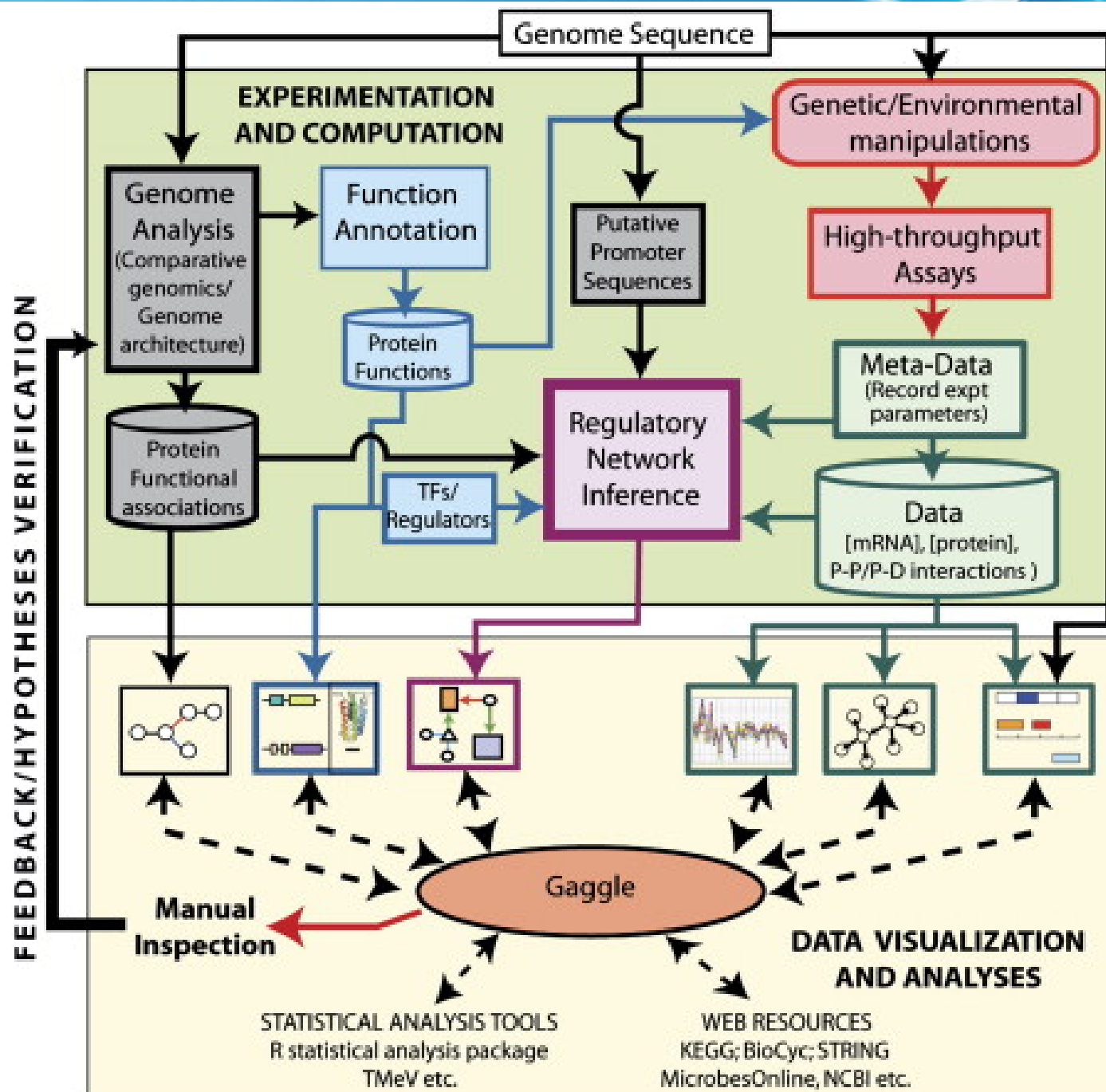
- Experiments:

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

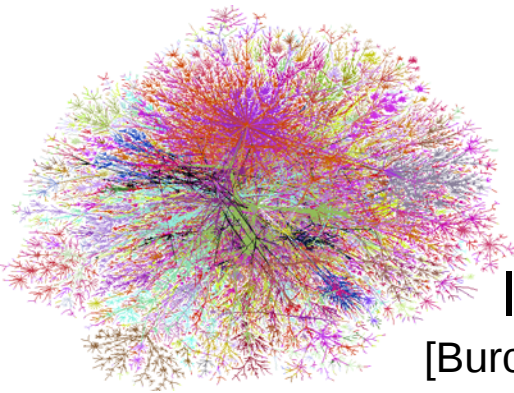




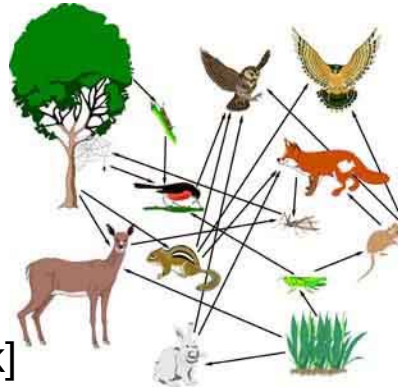
网络生物学



Networks as a universal language



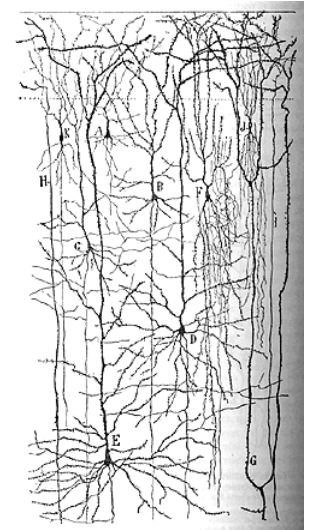
Internet
[Burch & Cheswick]



Food Web



Electronic
Circuit



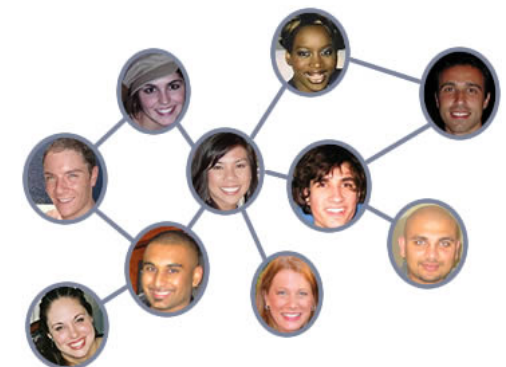
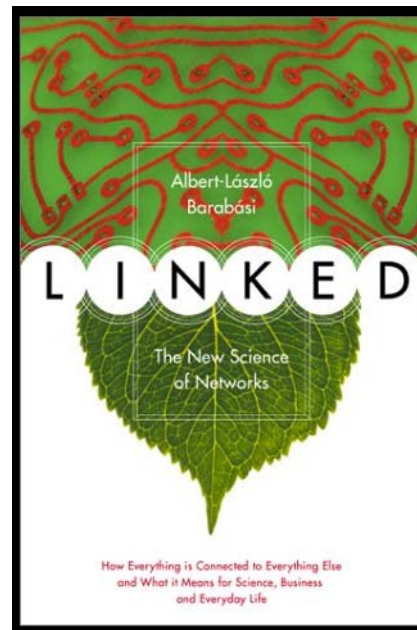
Neural Network
[Cajal]



Disease
Spread
[Krebs]



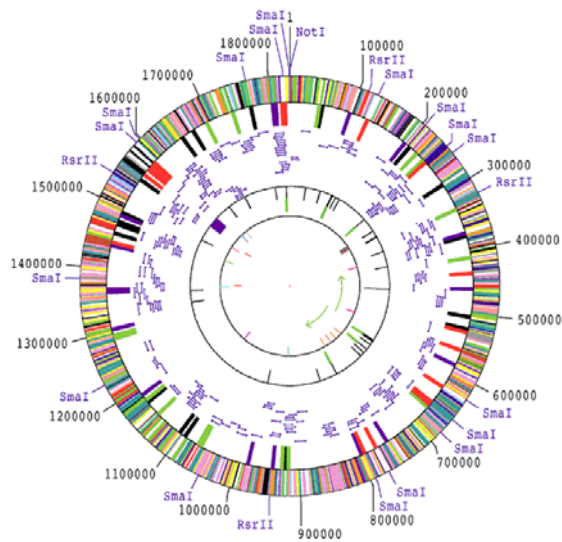
Protein
Interactions
[Barabasi]



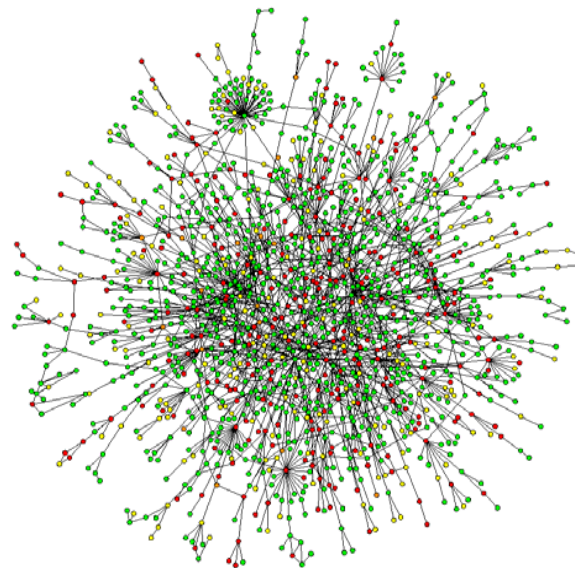
Social Network

生物分子网络

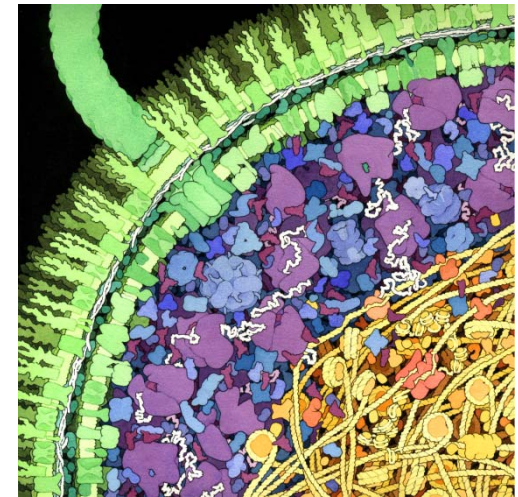
- 图作为基本工具用来强调相互作用并直观表示复杂的生物系统
- 节点代表生物分子，边代表他们之间在生命过程中的某种关系



1D: Complete
Genetic Partslist



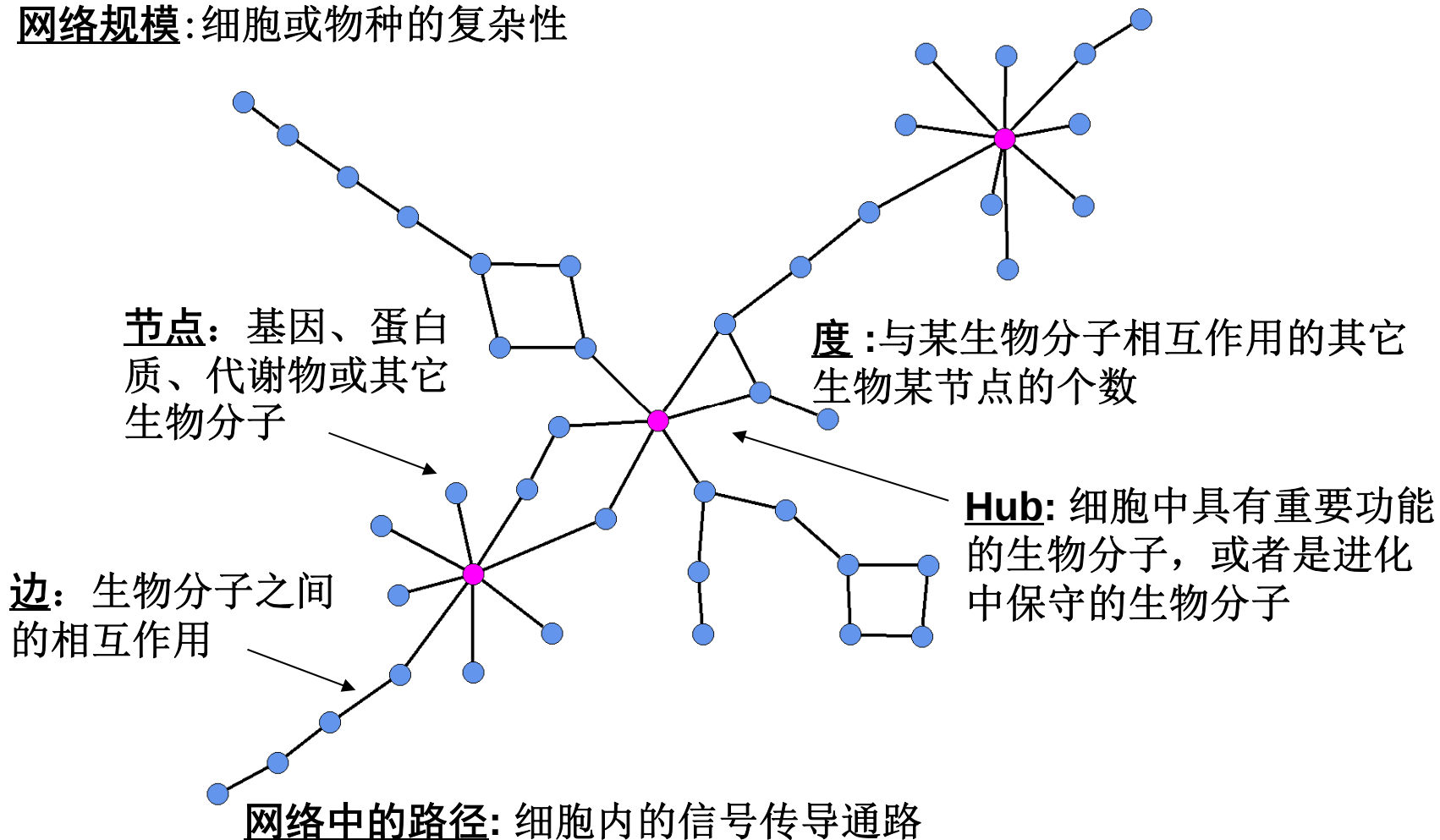
~2D: Bio-molecular
Network



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





At molecular level

Emerging area

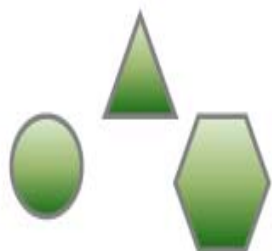
Network Systems Biology?

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



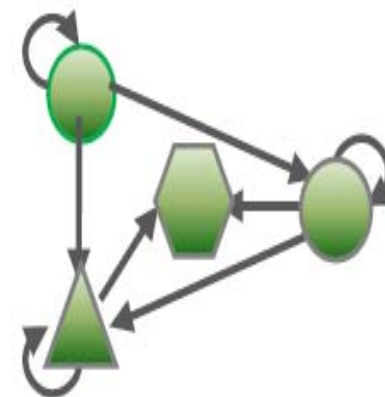
Interactions



Pathways



Networks



Individual molecules $\Rightarrow$ Pairwise interactions $\Rightarrow$ Local structures $\Rightarrow$ Global networks

Local

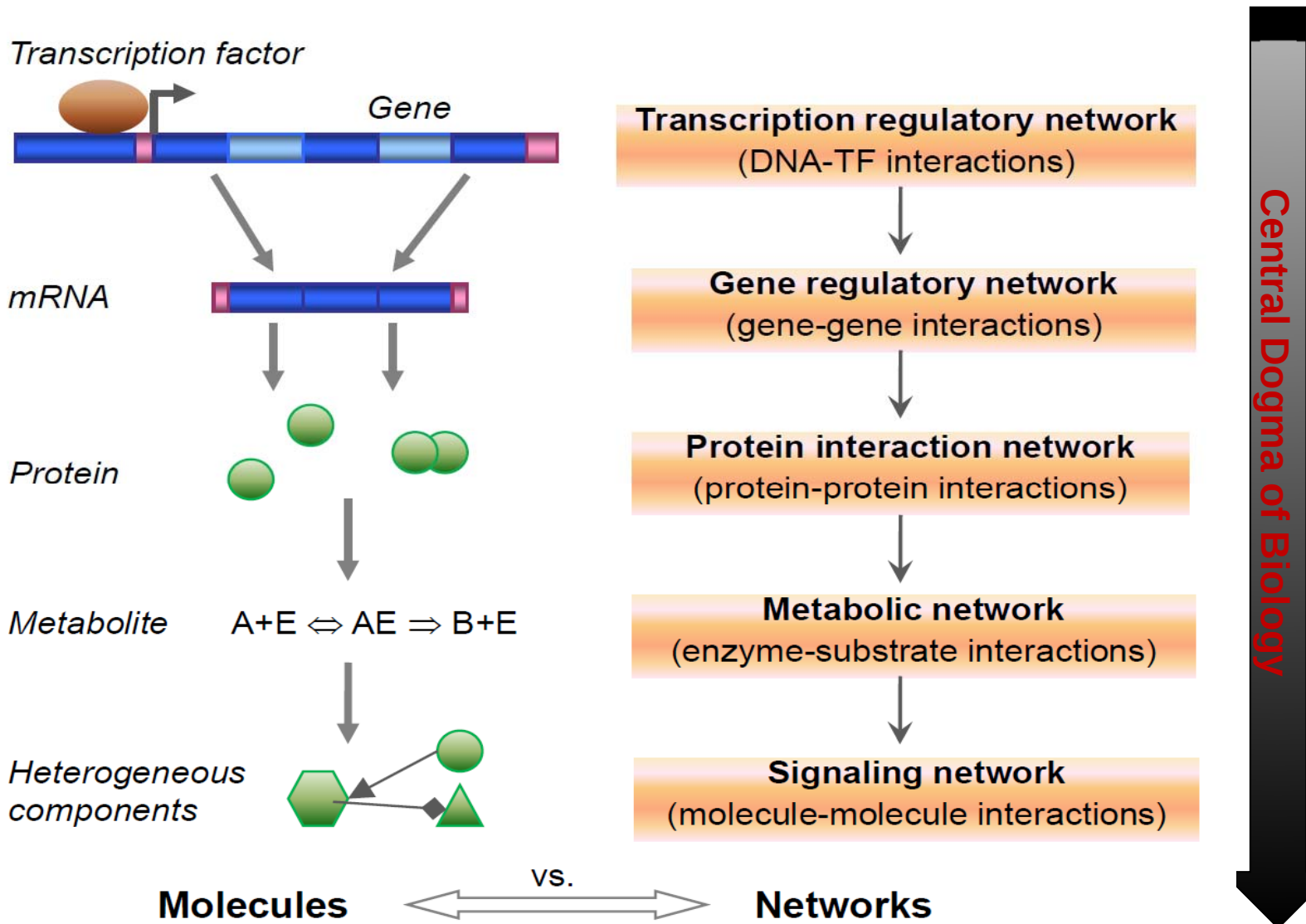
Global

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.

Biological networks

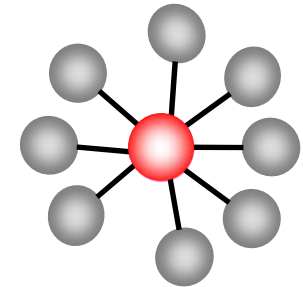


Types of Biological networks

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.

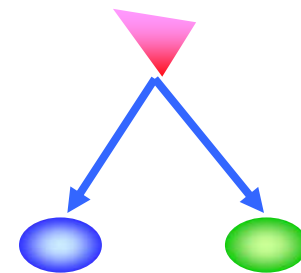


Undirected

Protein-protein interaction network: Physical or genetic network

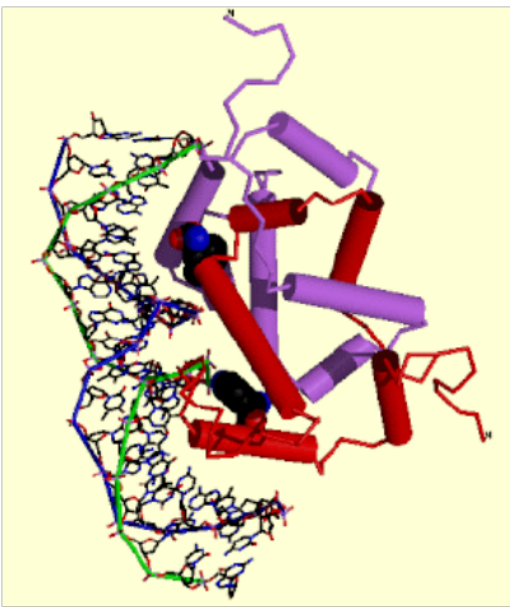
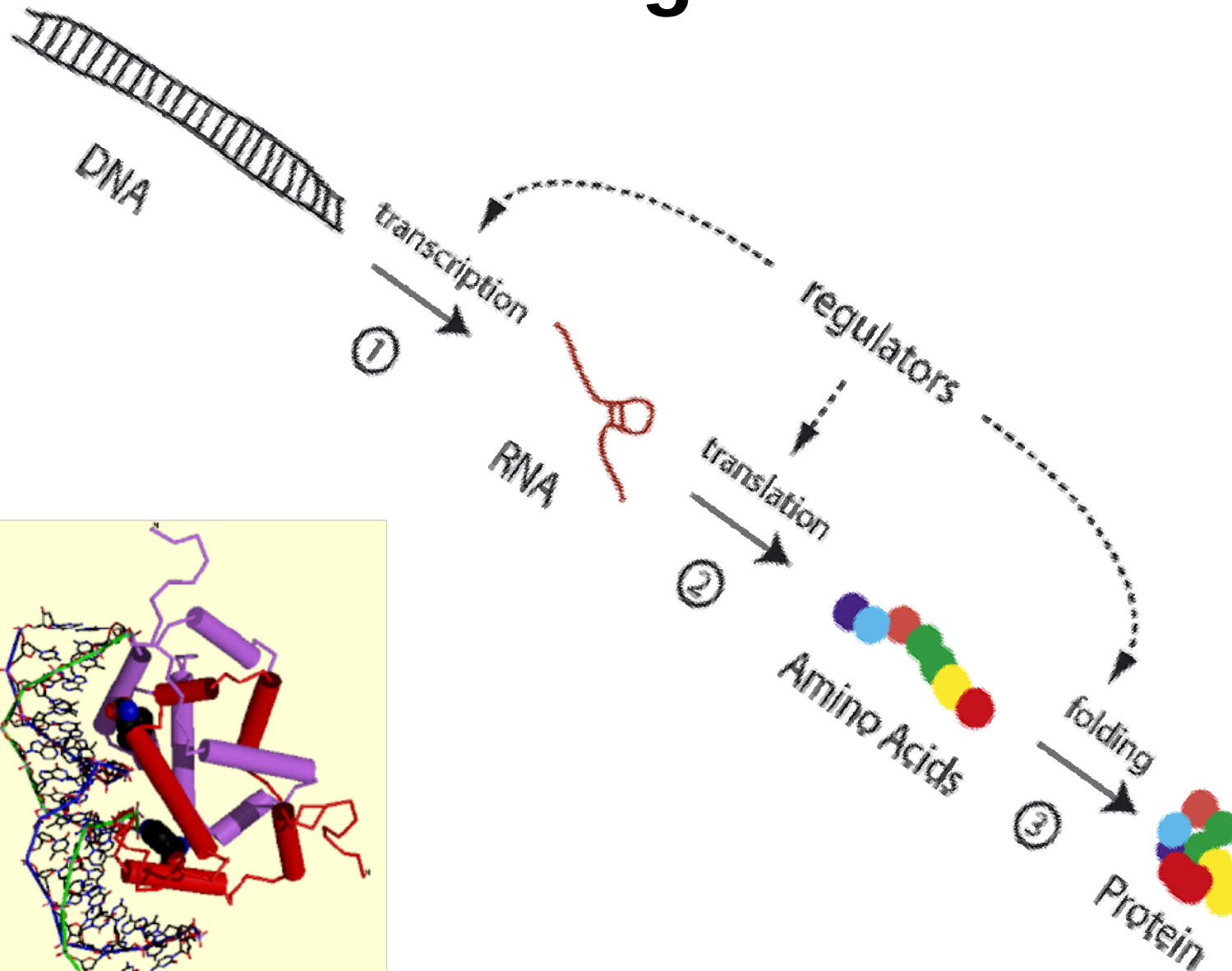
Metabolic network: Network of interactions between proteins, Synthesize / break down molecules (enzymes, cofactors)

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

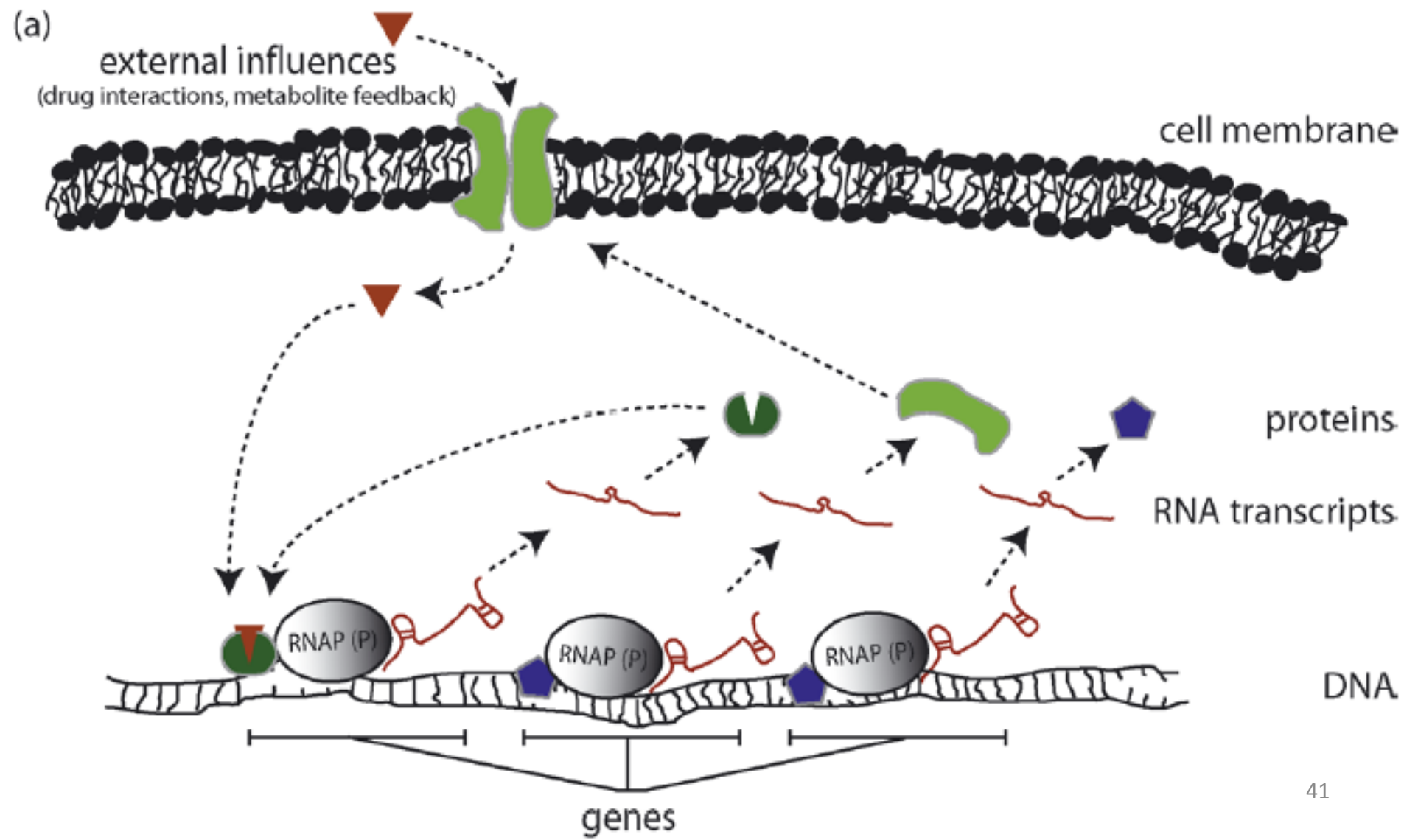


Directed

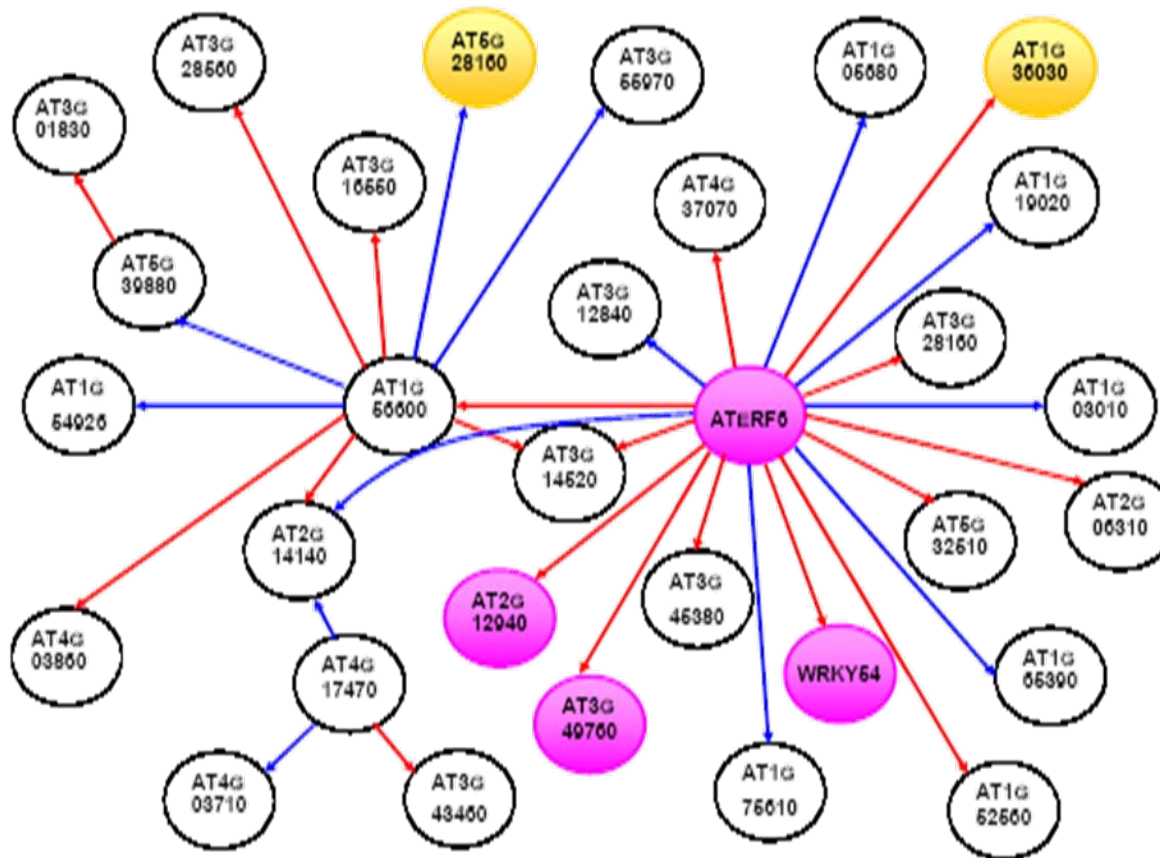
Regulation



Direct and indirect regulation



Genetic network



Node: gene

Edge: causality
regulatory
relationships

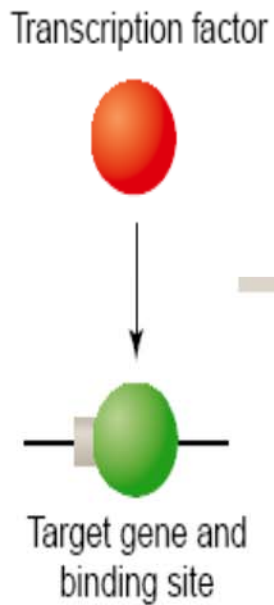
Directed

Edge weighted

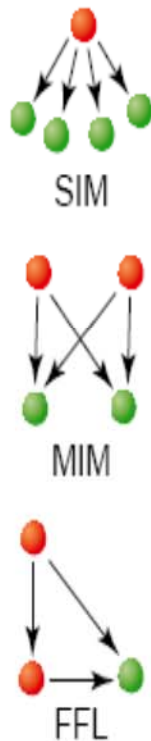
Genetic network consists of set of genes, proteins, small molecules, and their mutual regulatory interactions.

Transcriptional regulatory network

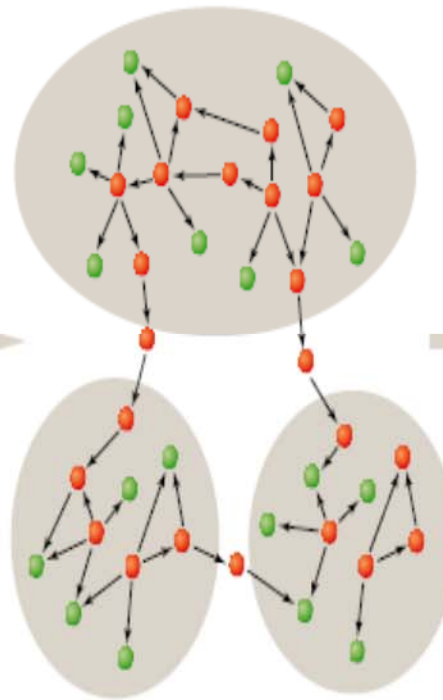
(a) Basic unit



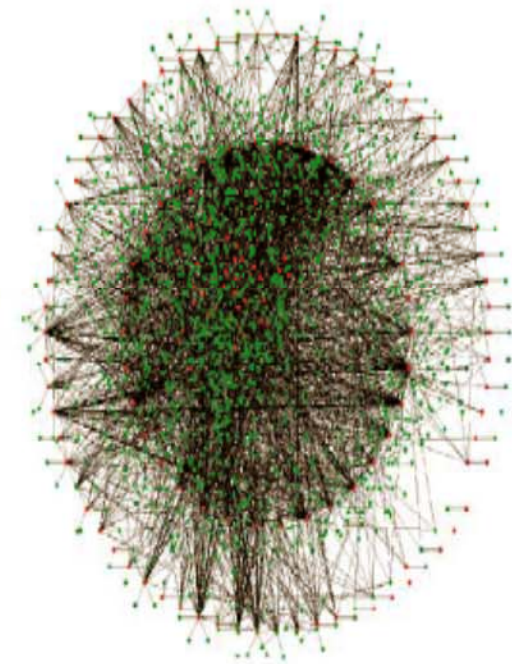
(b) Motifs



(c) Modules



(d) Transcriptional regulatory network



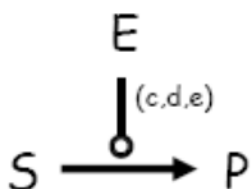
Current Opinion in Structural Biology

Subset of genetic network

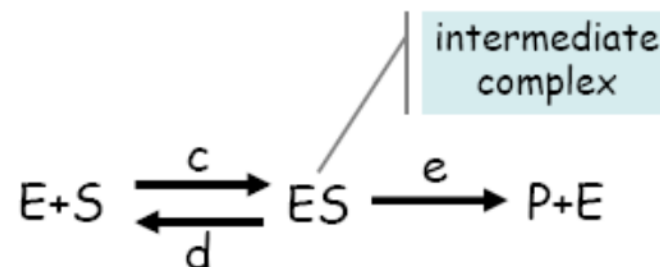
Node: TF and genes, Edge: regulation relationships
Directed, Edge weighted

Enzymatic Reactions

Reaction View

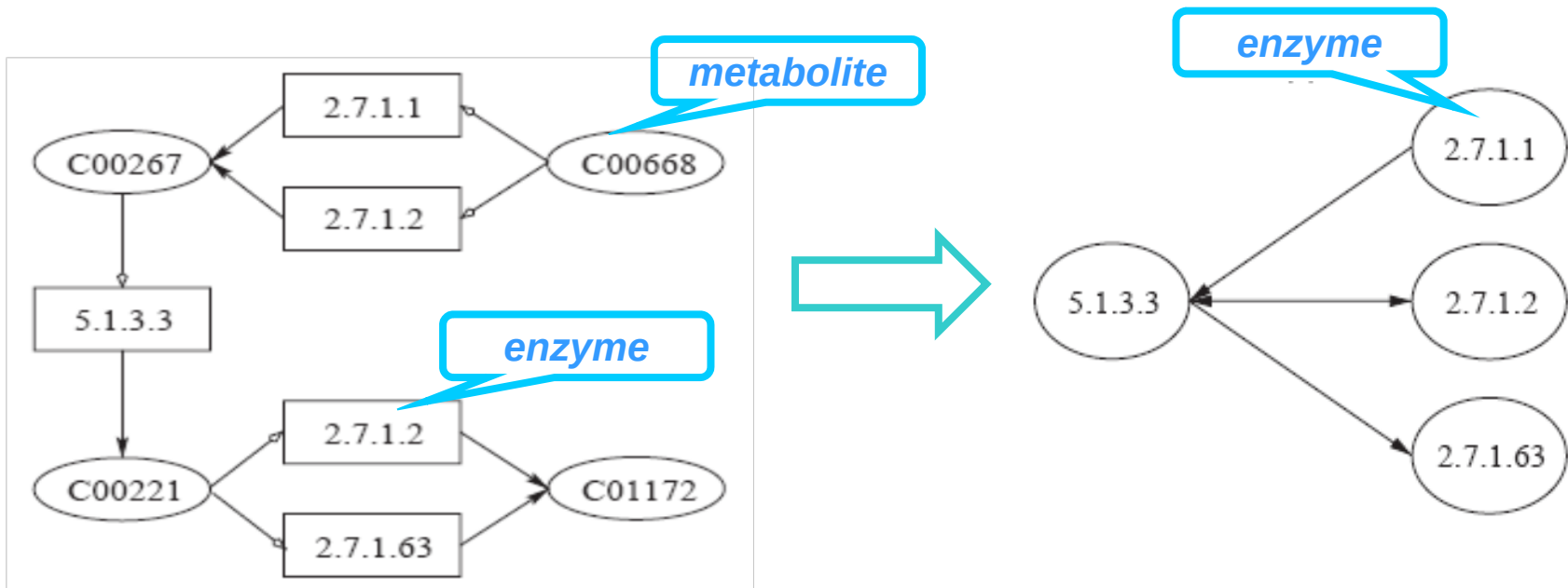


≡

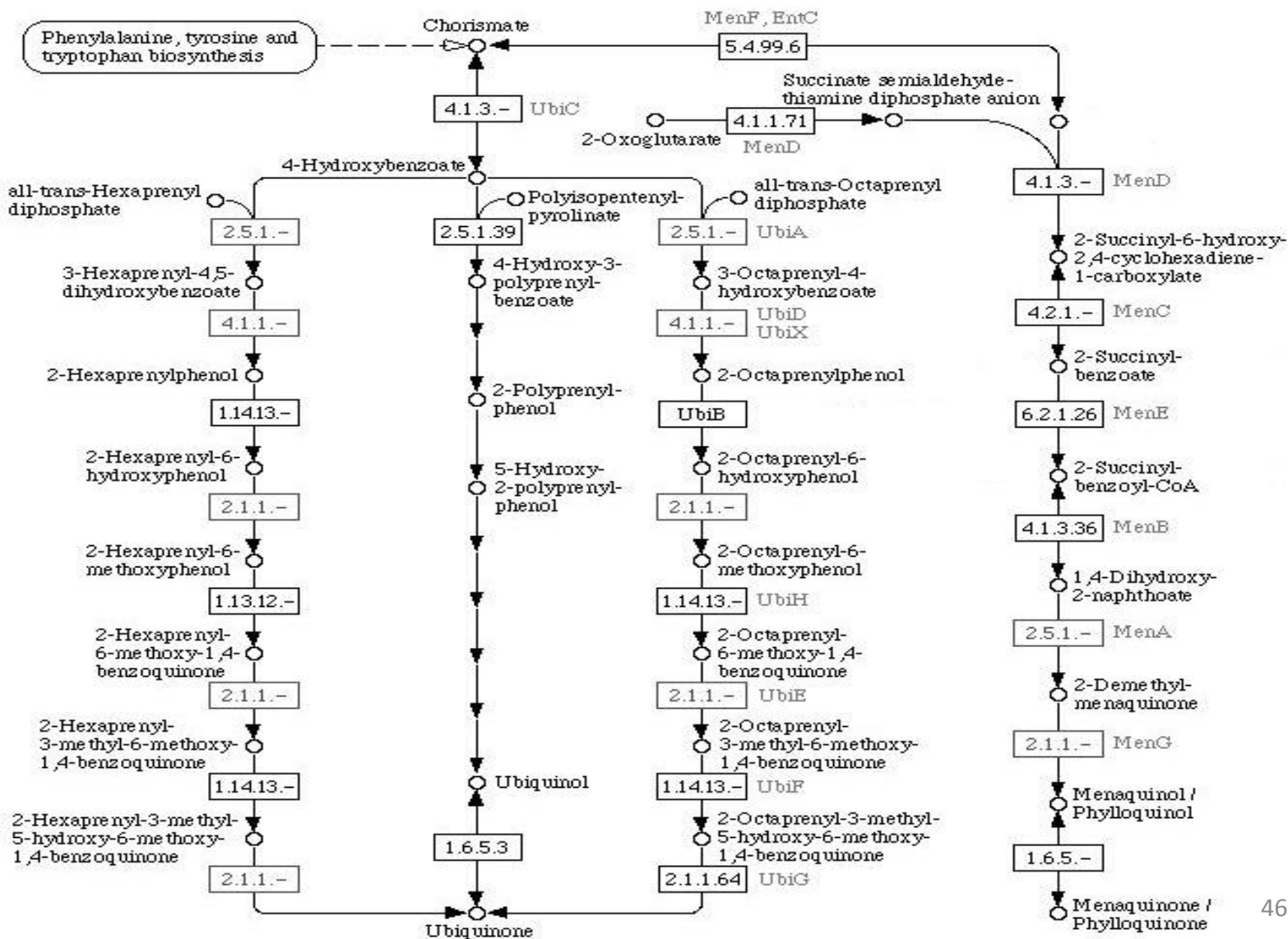


- **Definition of Metabolic network $G(V,E)$:**
 - For every enzyme z_i in Z - a node v_i exists
 - (v_i, v_j) in E iff z_j consumes the product of z_i
酶A的一个产物是酶B 的一个底物

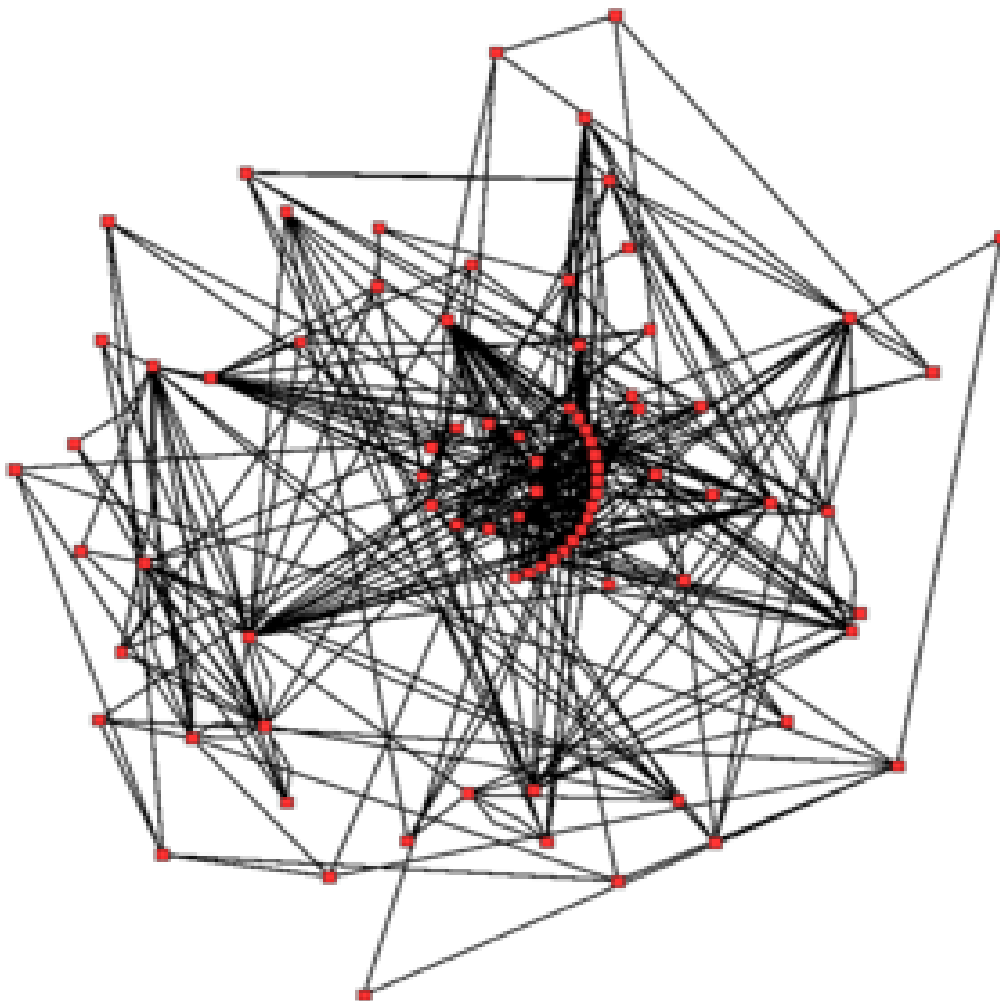
■ **Example:**



UBIQUINONE BIOSYNTHESIS



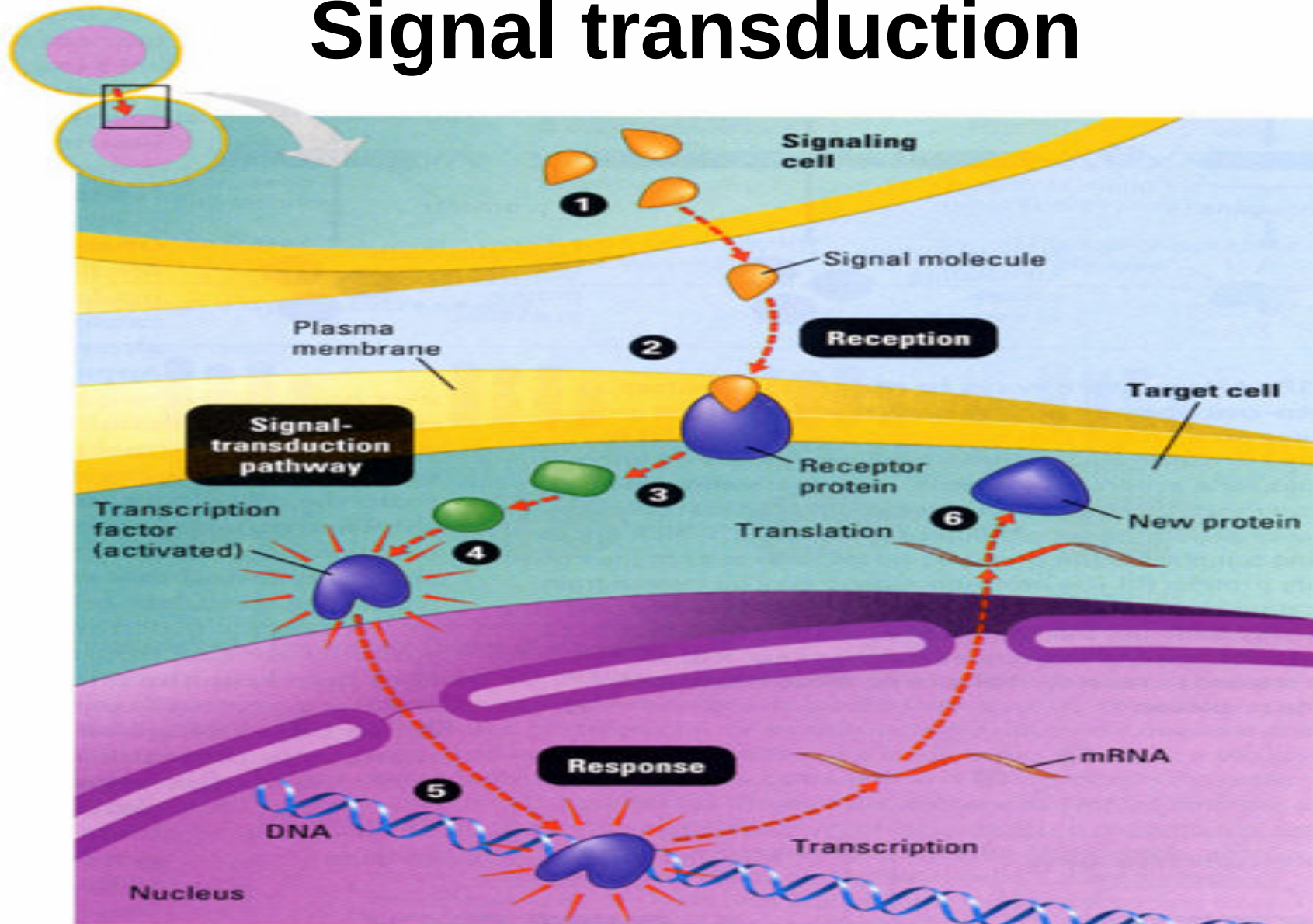
Metabolic network



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

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

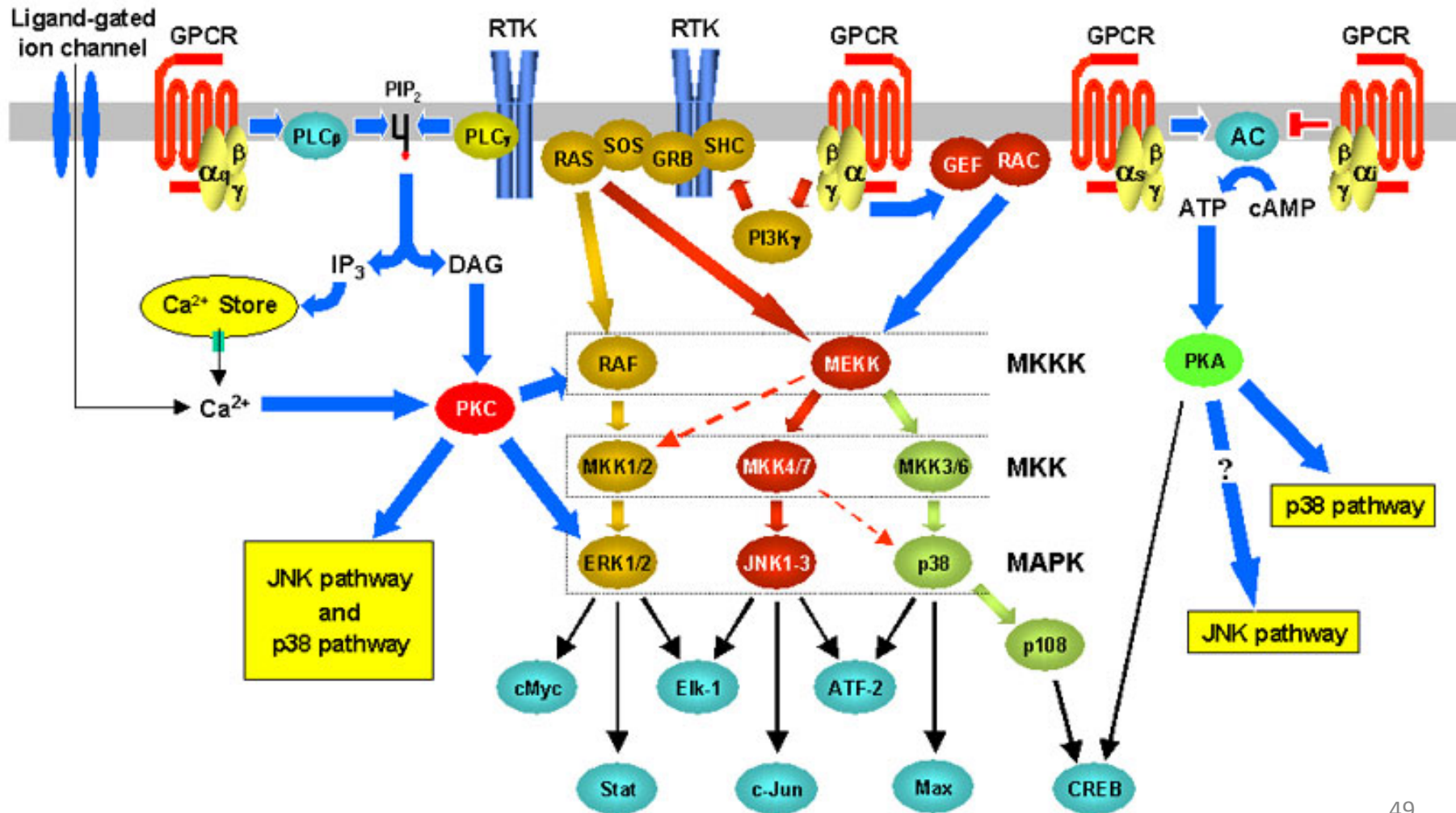
Signal transduction



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



Protein Interaction

Proteins:

Domains:

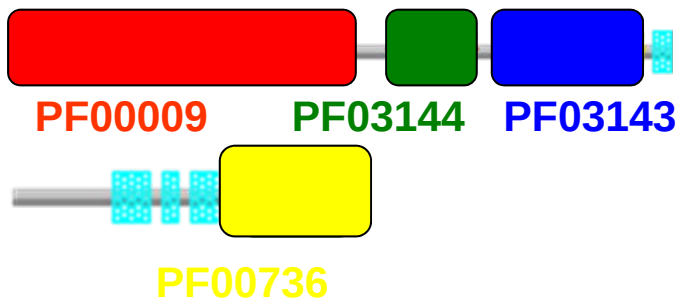
ORFs:

Interactions:

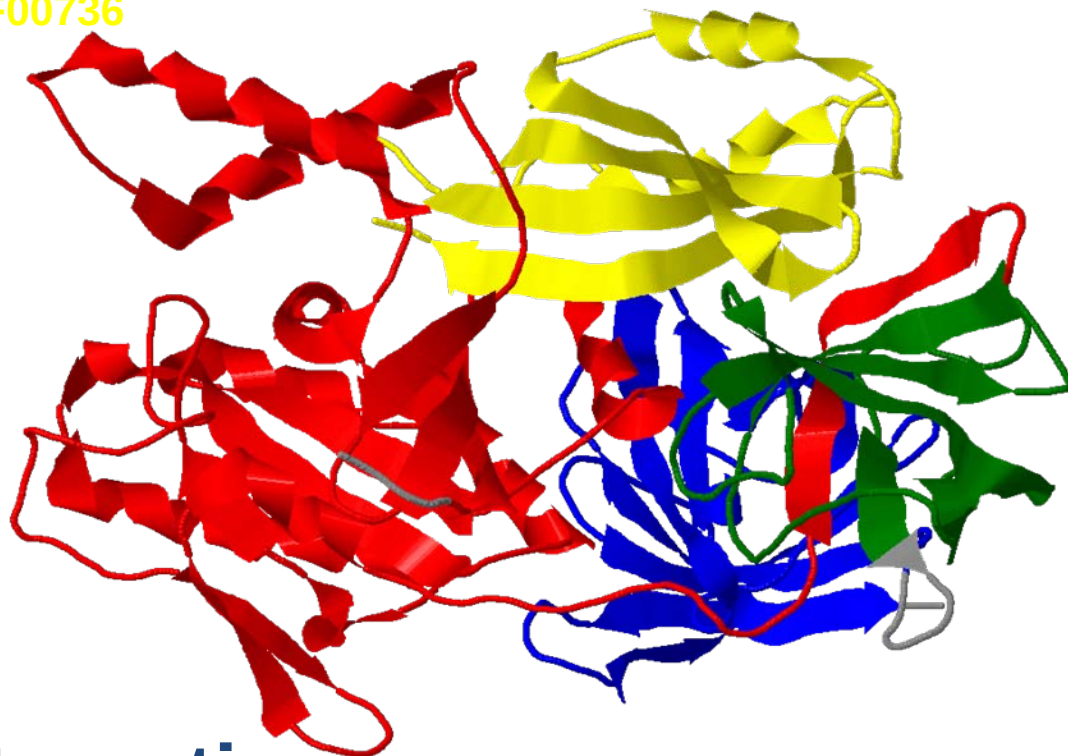
P02994



P32471

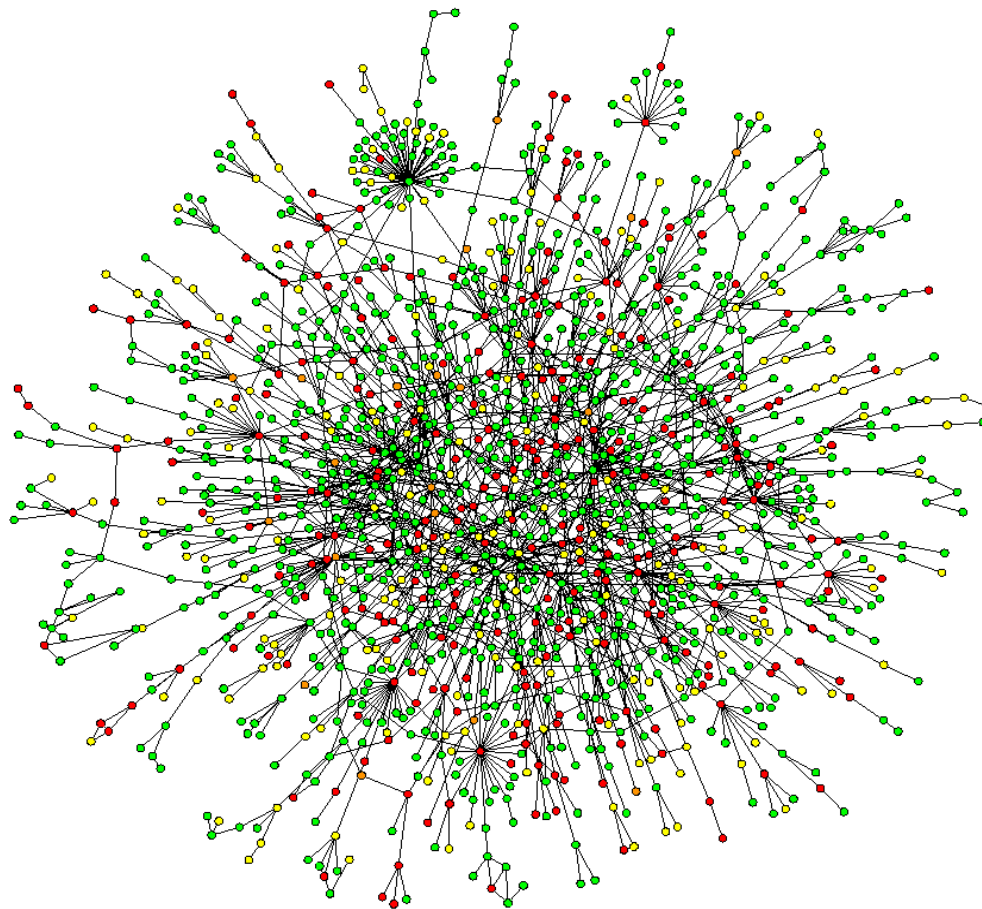


YBR118W ↔ YAL003W PF03144 PF00009 ↔ PF00736
 YBR118W ↔ YAL003W PF03143 PF00009 ↔ PF00736
 YBR118W ↔ YAL003W PF03143 PF03144 ↔ PF00736
 YPR080W ↔ YAL003W PF03144 PF00009 ↔ PF00736
 YPR080W ↔ YAL003W PF03143 PF00009 ↔ PF00736
 YPR080W ↔ YAL003W PF03143 PF03144 ↔ PF00736



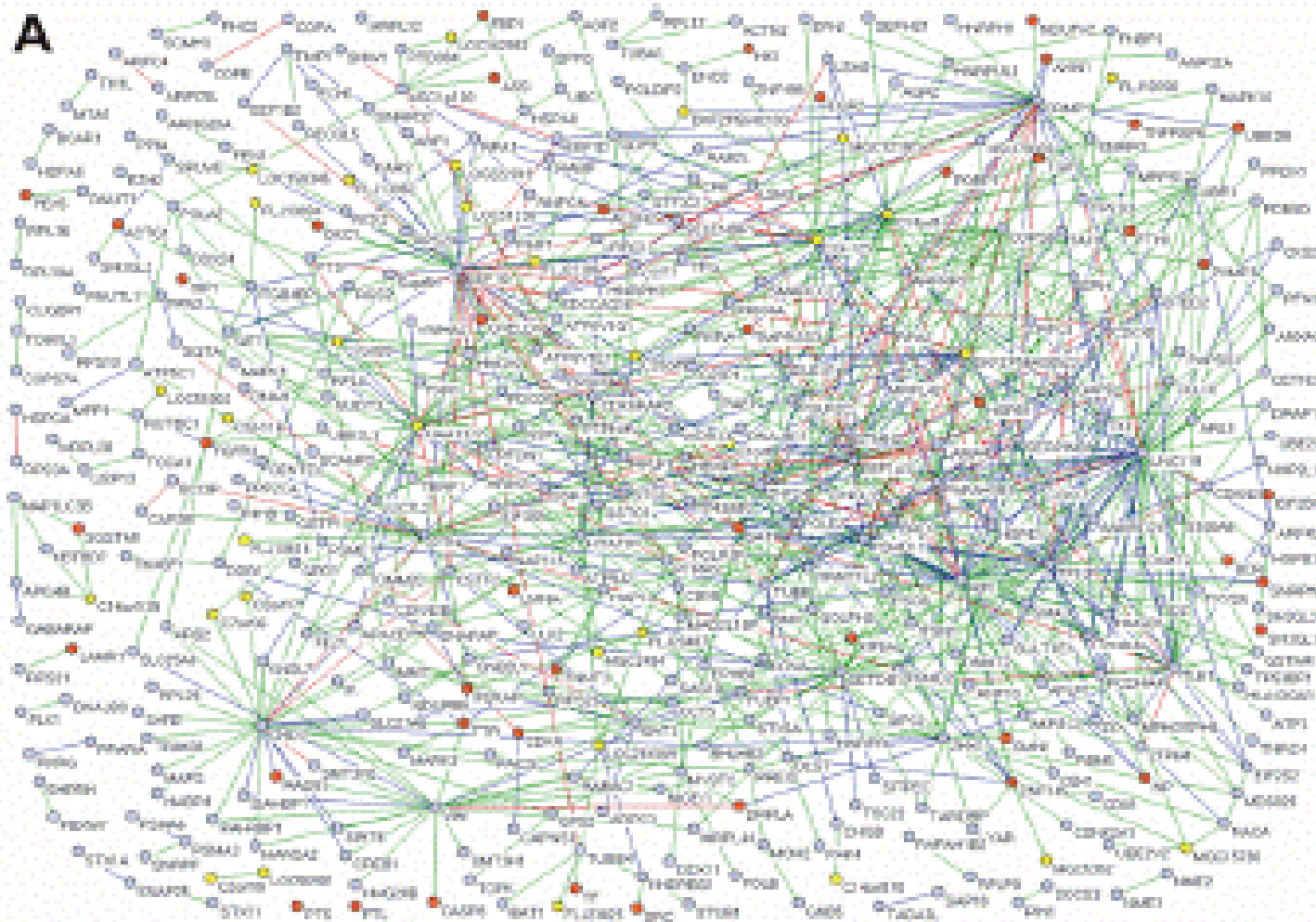
Protein interaction

Protein Interaction Network

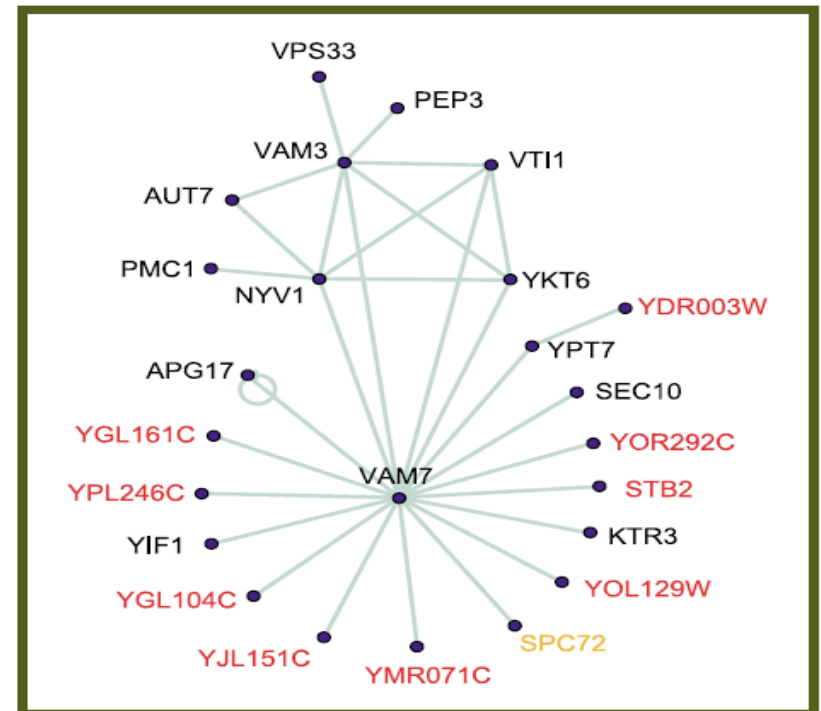
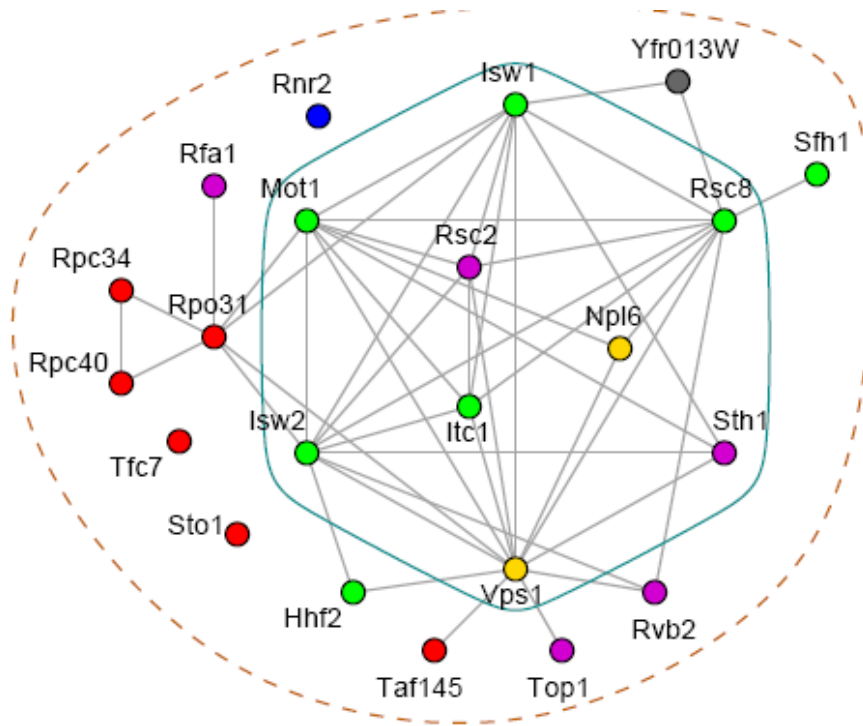


Yeast protein interaction network

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 Strengthened)

生物分子网络研究的科学问题

- 如何构建网络? (王勇)

建模, 数据处理、集成

- 如何分析网络? (张世华)

静态: 结构, 功能等

动态: 不同条件, 进化

致谢！

该PPT课件的制作基于2010、2011版本，并有适当增删。

感谢章祥淼研究员、吴凌云和王勇博士以及整个课题组！

感谢许多工作的原始作者，限于时间和篇幅文献和作者未能及时标注。