系统生物学

(Systems Biology)

马彬广

蛋白互作网络

（第四讲）

蛋白互作网络的定义

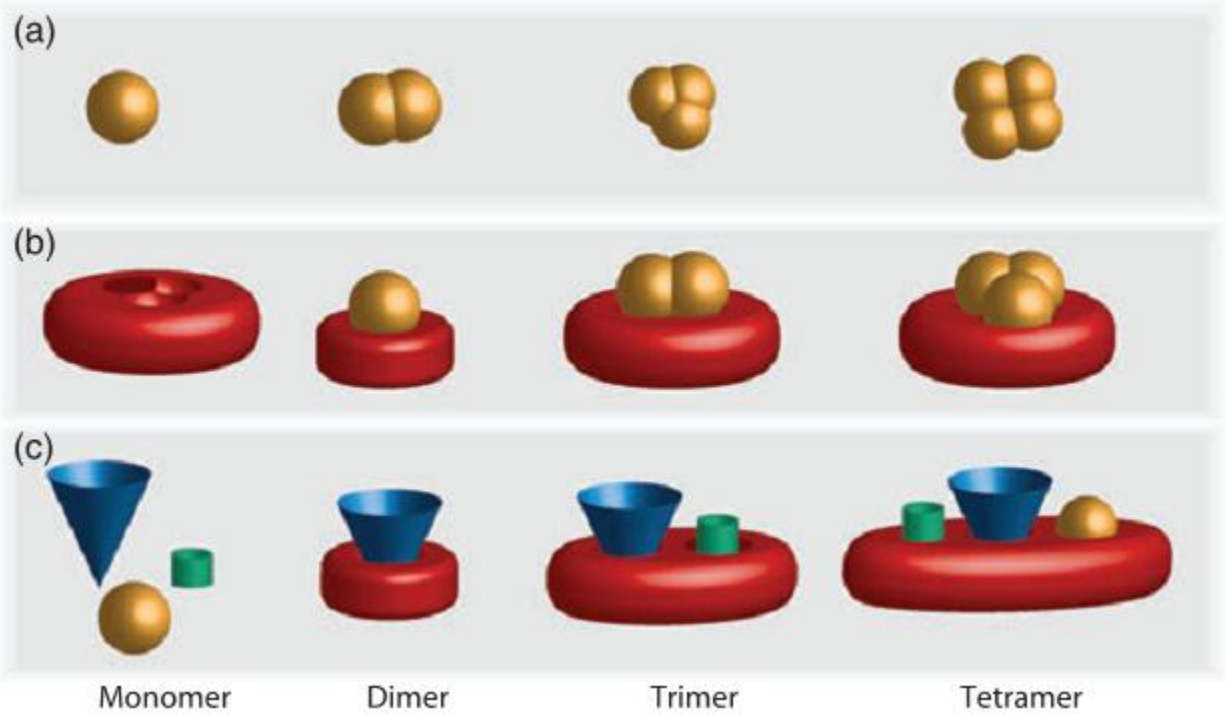
- 蛋白互作的分类：
  （1）物理互作（physical interaction），直接物理接触。
  （2）功能互作（functional interaction），逻辑关联。

- 蛋白互作交织在一起就构成蛋白互作网络。有两种类型的蛋白互作网络：binary和co-complex。

- 蛋白复合物是蛋白互作网络的功能模块，蛋白复合物可以分成永久复合物和瞬态复合物。

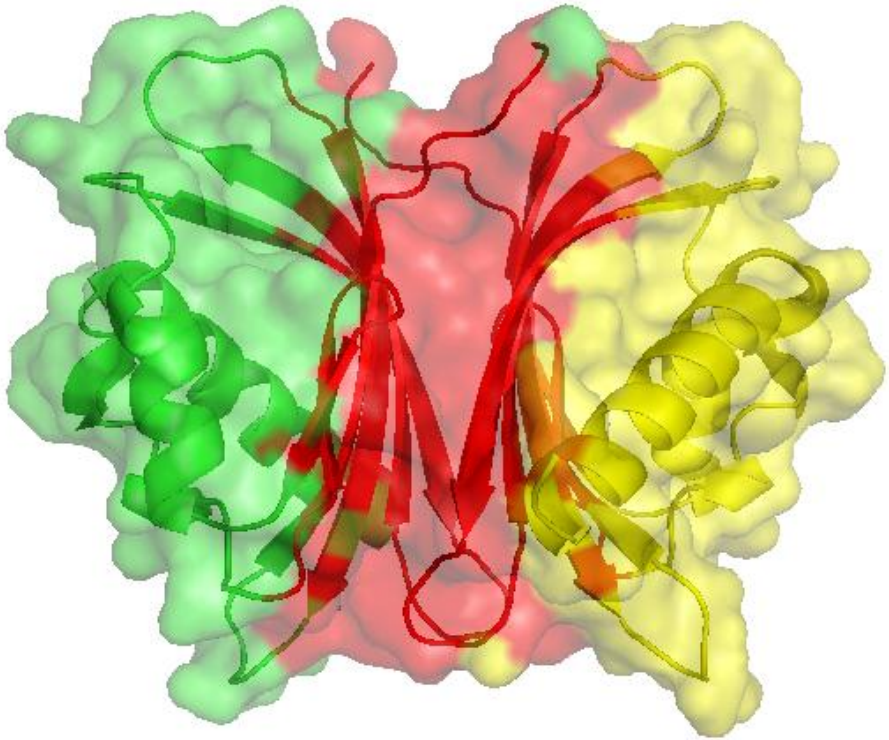
- 蛋白互作的生物学意义：（1）复合物是分子机器的基本形式，可以完成分子加工过程和作为反应的催化剂；（2）瞬时的蛋白互作在分子功能调控和信号传导中都具有重要意义。（3）蛋白互作的丢失或异常，可能导致细胞的功能失调甚至死亡。

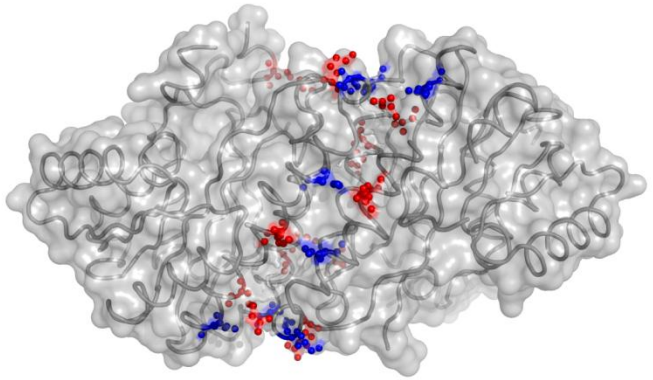
蛋白复合物的类型

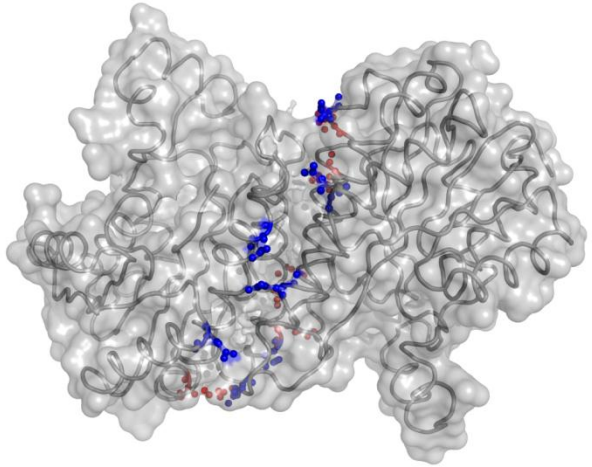


The Plant Journal (2008) 53, 610–635.

蛋白作用界面

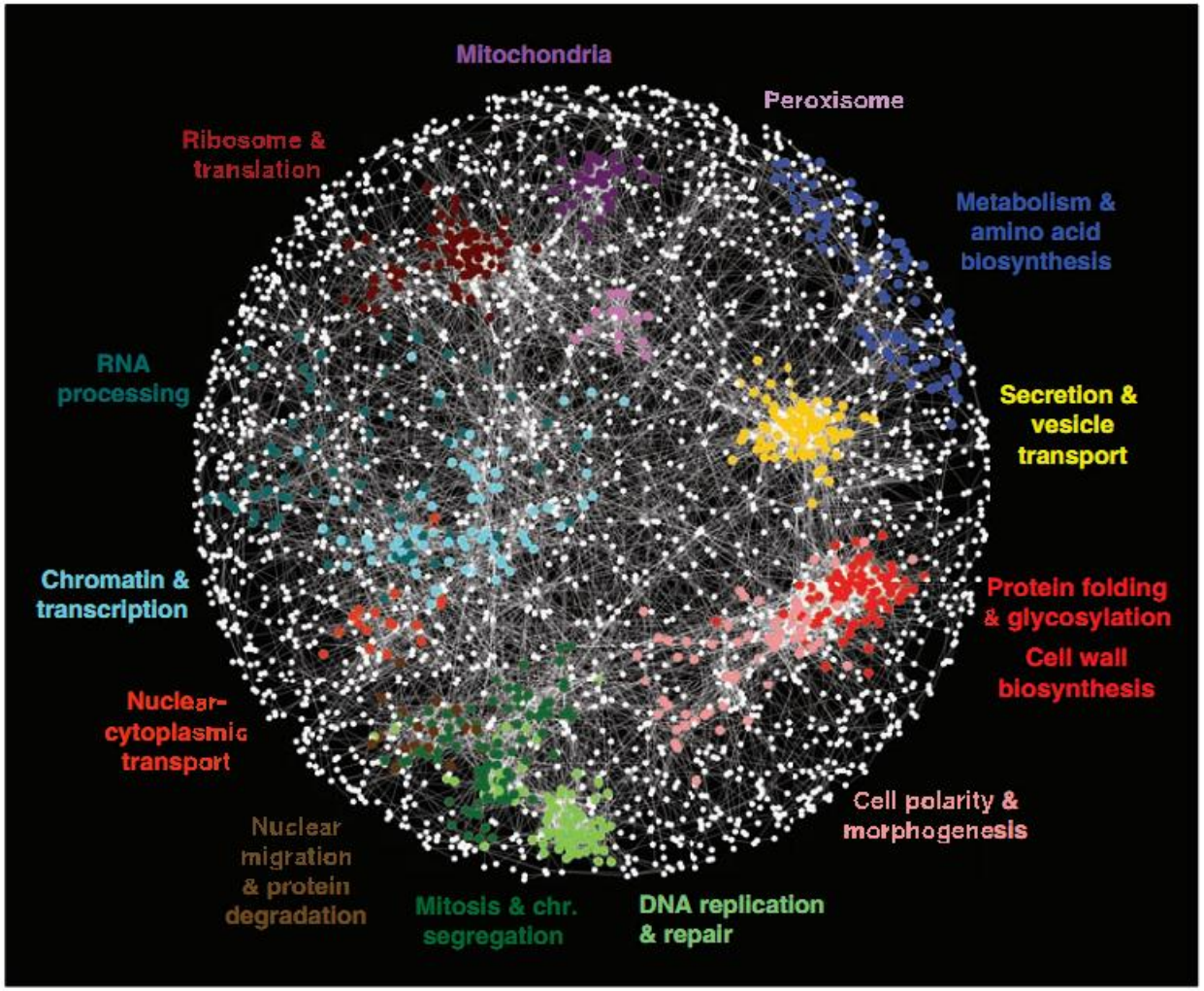






Ma B.G. et al .Structure (2010) 18, 819-828.

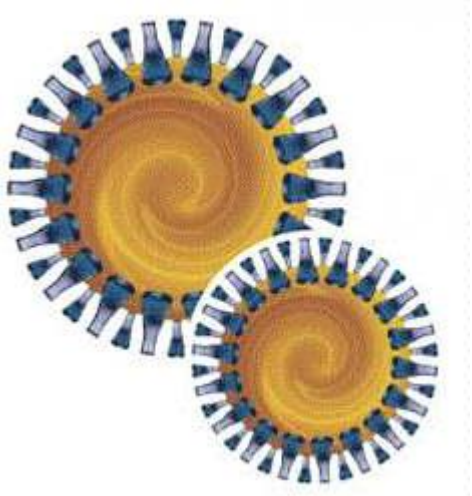
基因的功能互作网络



Science (2010), 327: 425.

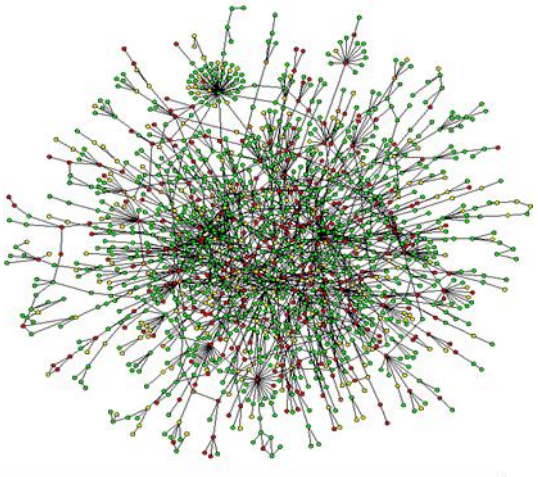
蛋白互作研究中的基本问题

- 谁跟谁发生互作？
- 为何要发生互作？
- 何时何地互作？
- 如何互作？
- 互作的结果是什么？

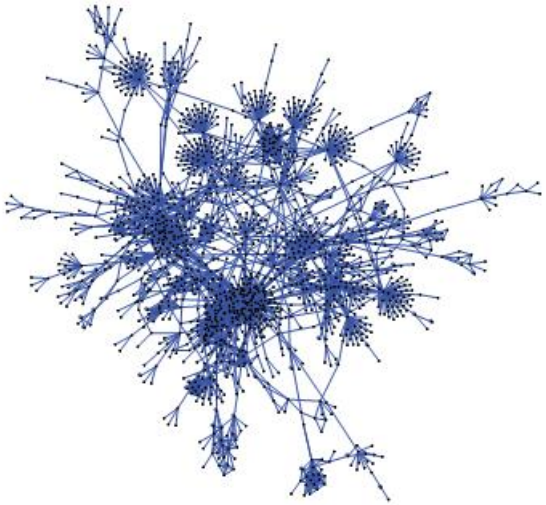


http://www.sciencemag.org/site/products/protein.xhtml

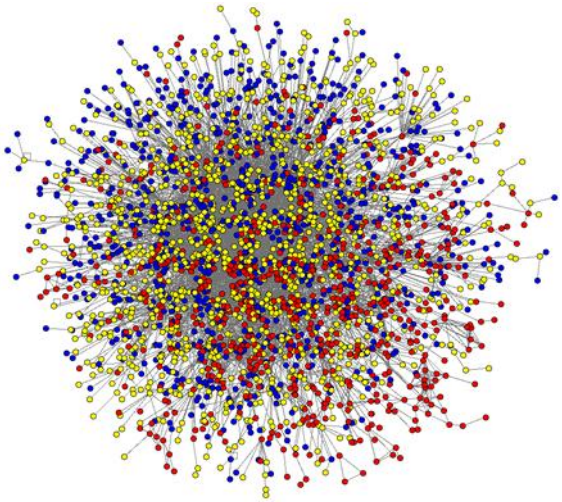
蛋白互作网络举例



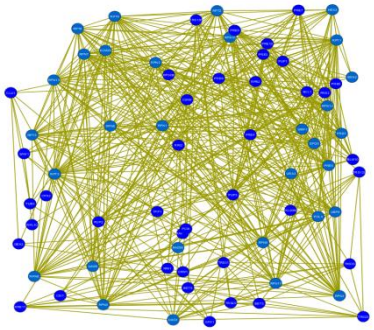
Yeast Binary



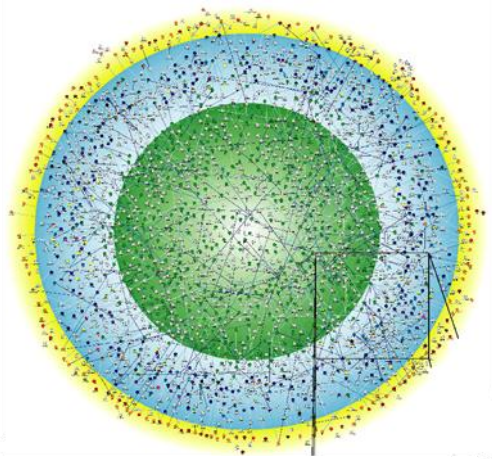
Arabidopsis Binary



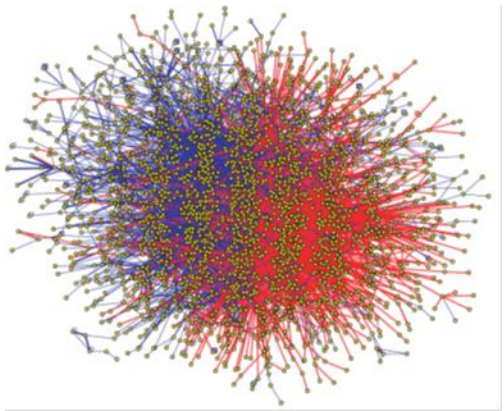
Worm Binary



Yeast Co-complex



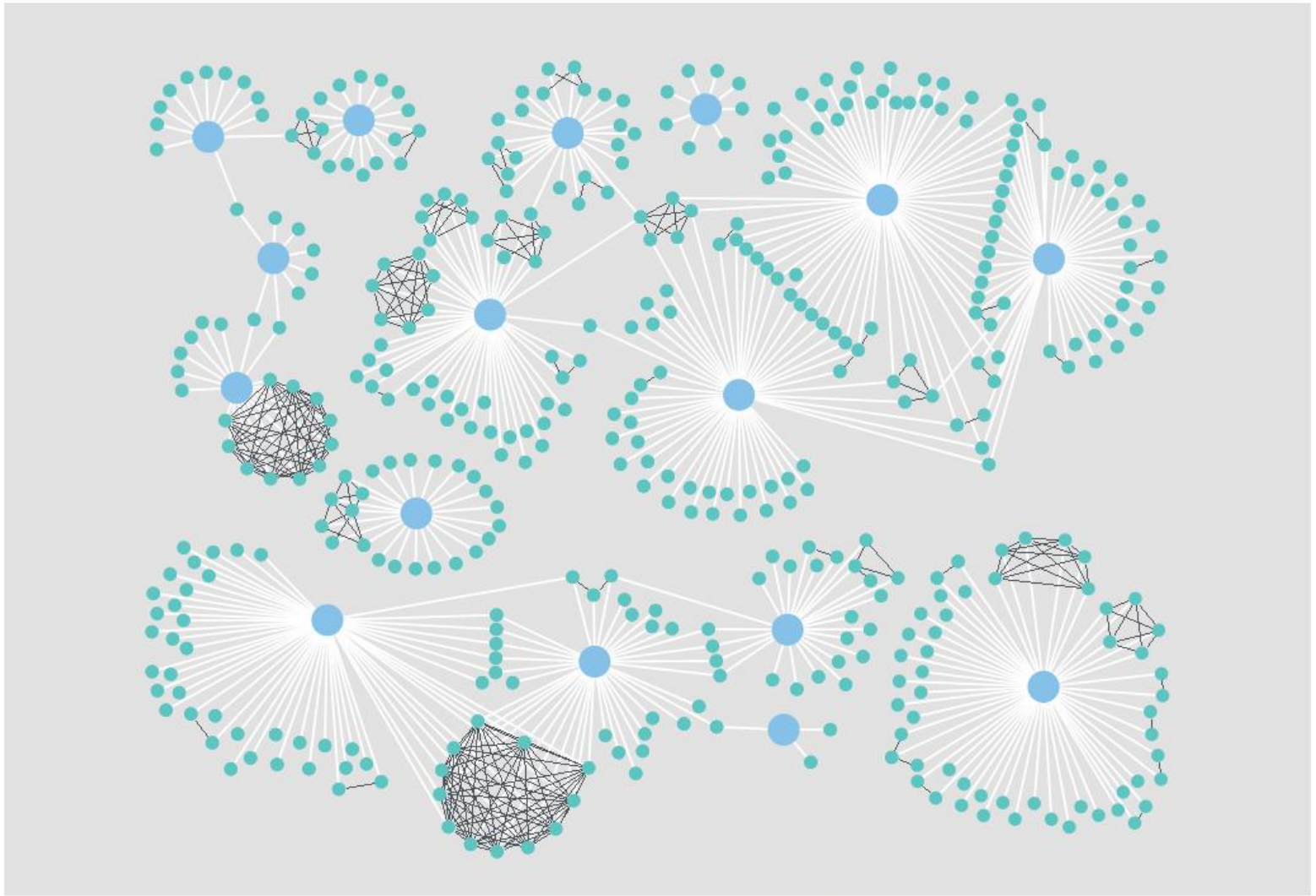
Fly Binary



Human Binary

蛋白互作网络举例

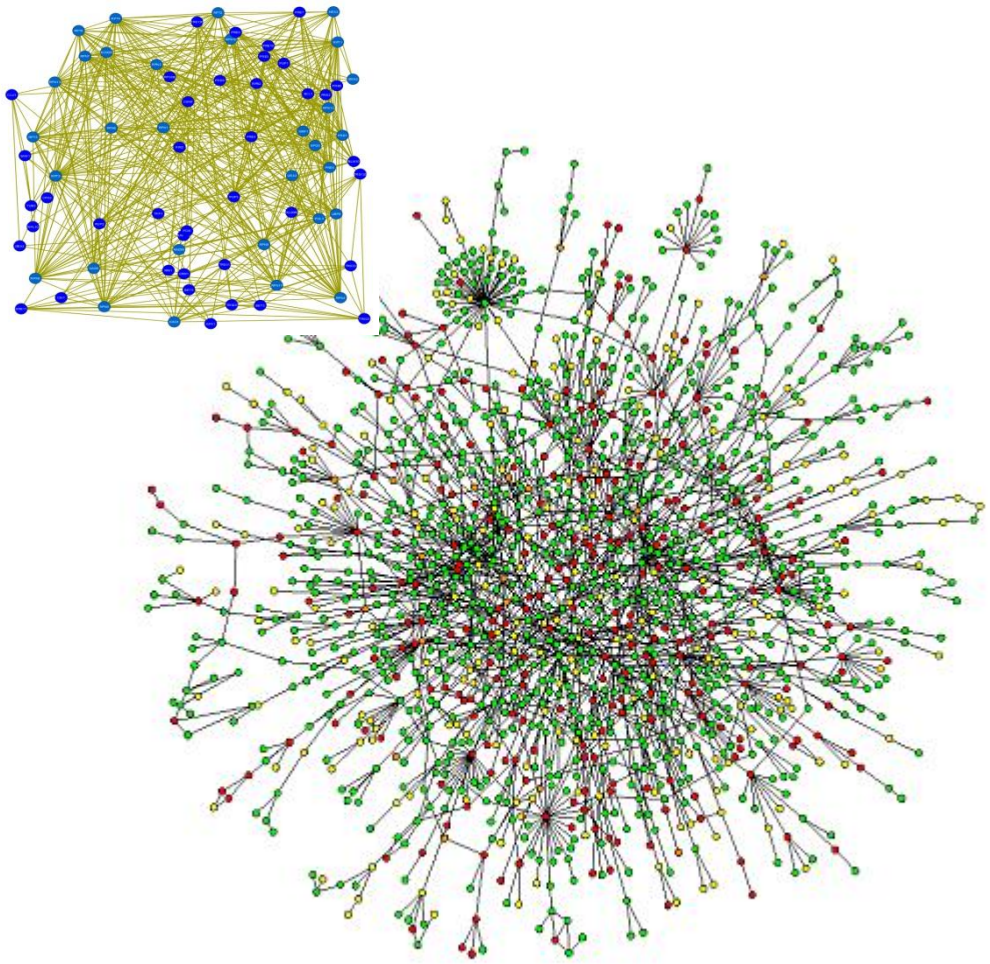
HIV病毒蛋白与人体蛋白的互作网络



Multiple replicated experiments and sophisticated statistics reveal 497 interactions between 16 HIV proteins (blue) and hundreds of human factors.

Nature (2012), 484: 271.

蛋白互作网络的特点



Yeast protein interaction network

- 蛋白为节点，互作为边。
- 节点可分类，边无方向性。
- 无标度性，结构自相似。
- 中心性， 存在Hub节点。
- 模块性，存在团簇和模体。

蛋白互作的实验检测方法

体内

- 酵母双（多）杂交
- 细菌双杂交
- 噬菌体展示技术
- 荧光互补或共振能量转移
- 合成致死
- 等

体外

- 免疫共沉淀（Pull-down）
- 亲和纯化质谱
- 化学交联
- 亲和印记、层析
- 蛋白质芯片
- 复合物结晶

蛋白互作的实验检测方法

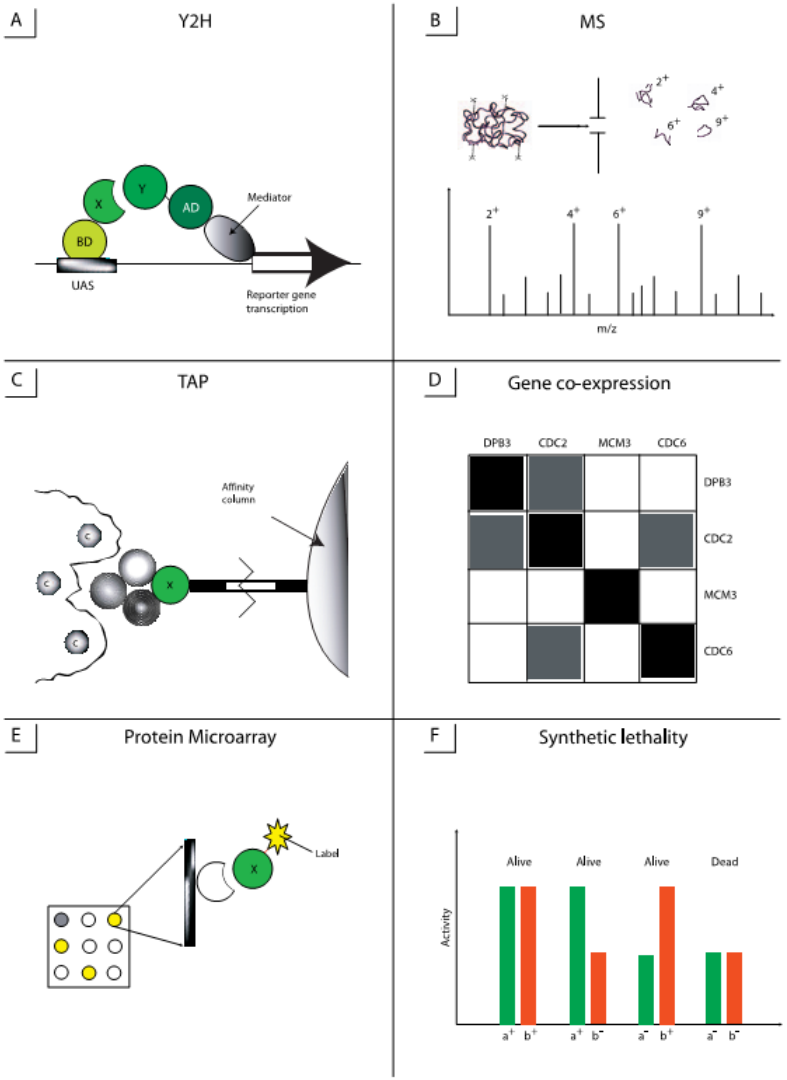
Table 1. Different Experimental Methods Measuring Protein Interactions

Method	High-Throughput Approach	Living Cell Assay	Type of Interactions	Type of Characterization
Y2H [47,48]	+	In vivo	Physical interactions (binary)	Identification
Affinity purification–MS [61]	+	In vitro	Physical interactions (complex)	Identification
DNA microarrays/Gene coexpression [113]	+	In vitro	Functional association	Identification
Protein microarrays [114–116]	+	In vitro	Physical interaction (complex)	Identification
Synthetic lethality [85,86]	+	In vivo	Functional association	Identification
Phage display [117]	+	In vitro	Physical interaction (complex)	Identification
X-ray crystallography, NMR spectroscopy [84]	–	In vitro	Physical interactions (complex)	Structural and biological characterization
Fluorescence resonance energy transfer [89]	–	In vivo	Physical interaction (binary)	Biological characterization
Surface plasmon resonance [91]	–	In vitro	Physical interaction (complex)	Kinetic, dynamic characterization
Atomic force microscopy [93]	–	In vitro	Physical interaction (binary)	Mechanical, dynamic characterization
Electron microscopy [118]	–	In vitro	Physical interaction (complex)	Structural and biological characterization

High-throughput techniques are indicated with pluses (second column), and those which can provide information on interactions in vivo are shown in the third column. Fourth column indicates whether the method supplies data on physically interacting proteins in a complex ("complex") or only pairwise interactions ("binary"). Methods inferring interactions through functional association are shown as well. The type of protein interaction characterization is shown in the last column.
doi:10.1371/journal.ppat.0030042.t001

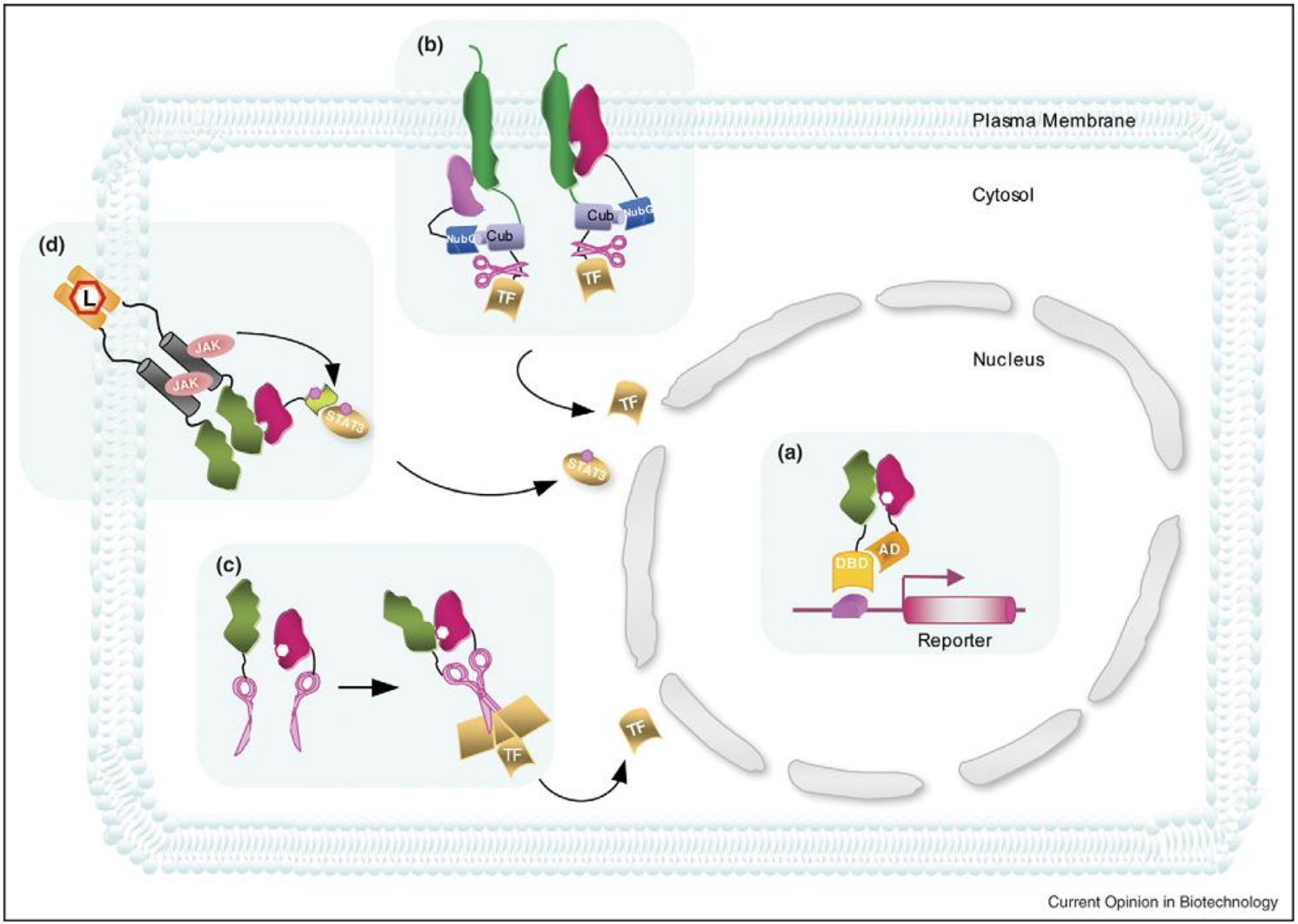
PLoS Comput Biol 3(3): e42. doi:10.1371/journal.pcbi.0030042

蛋白互作的几种主要的高通量方法



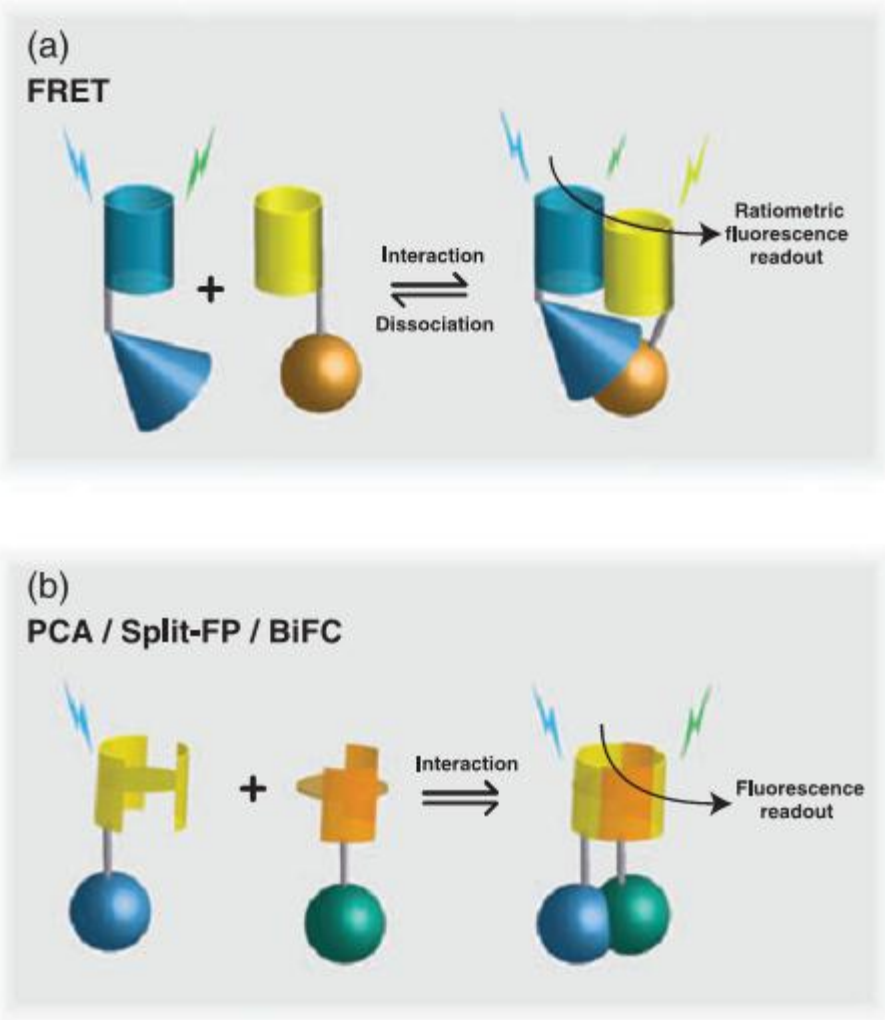
PLoS Comput Biol 3(3): e42. doi:10.1371/journal.pcbi.0030042

酵母双杂交实验方法简介



Current Opinion in Biotechnology 2008, 19:316–323.

荧光共振能量转移和荧光互补方法



The Plant Journal (2008) 53, 610–635.

化学交联法捕获蛋白瞬时相互作用

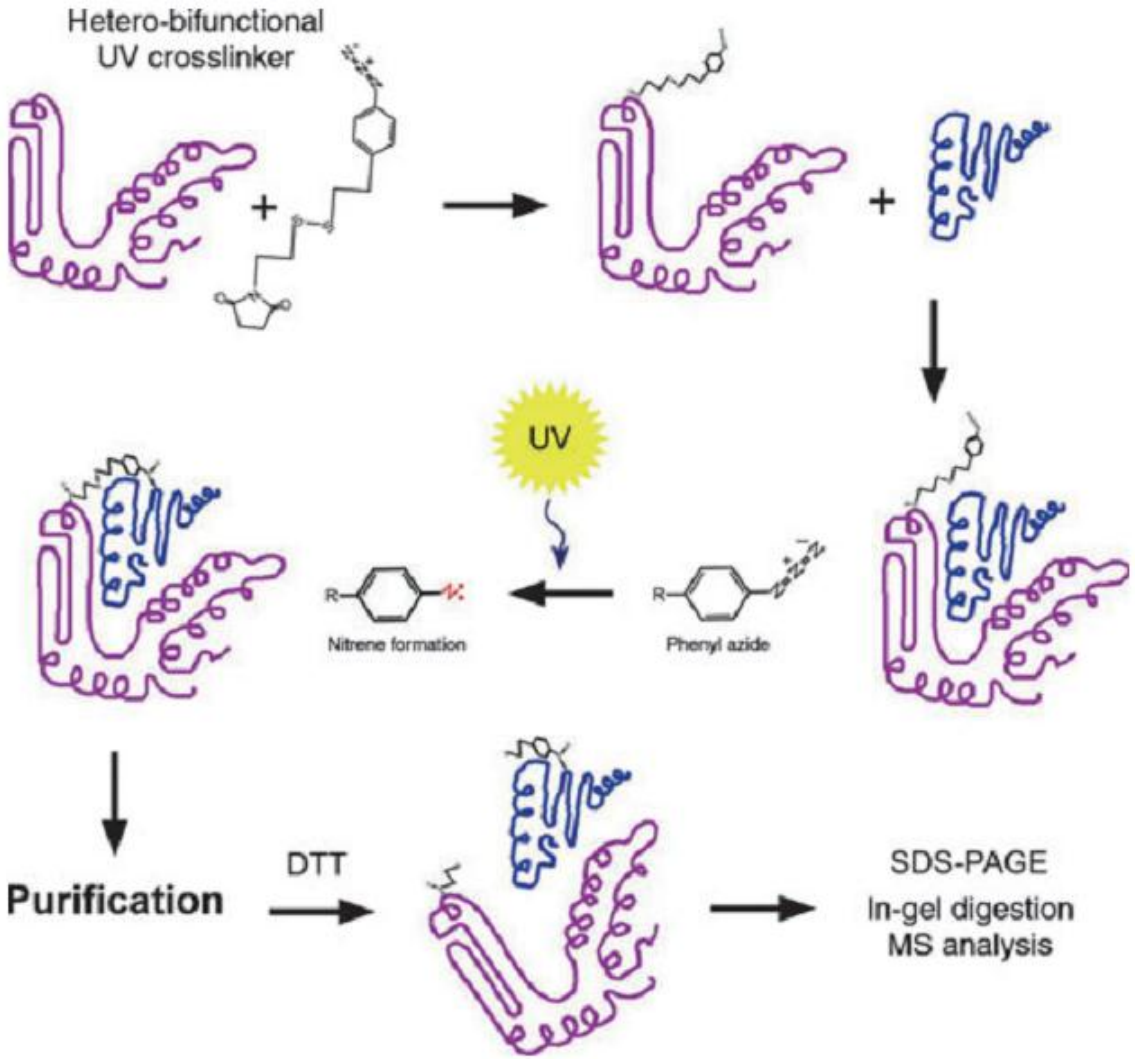
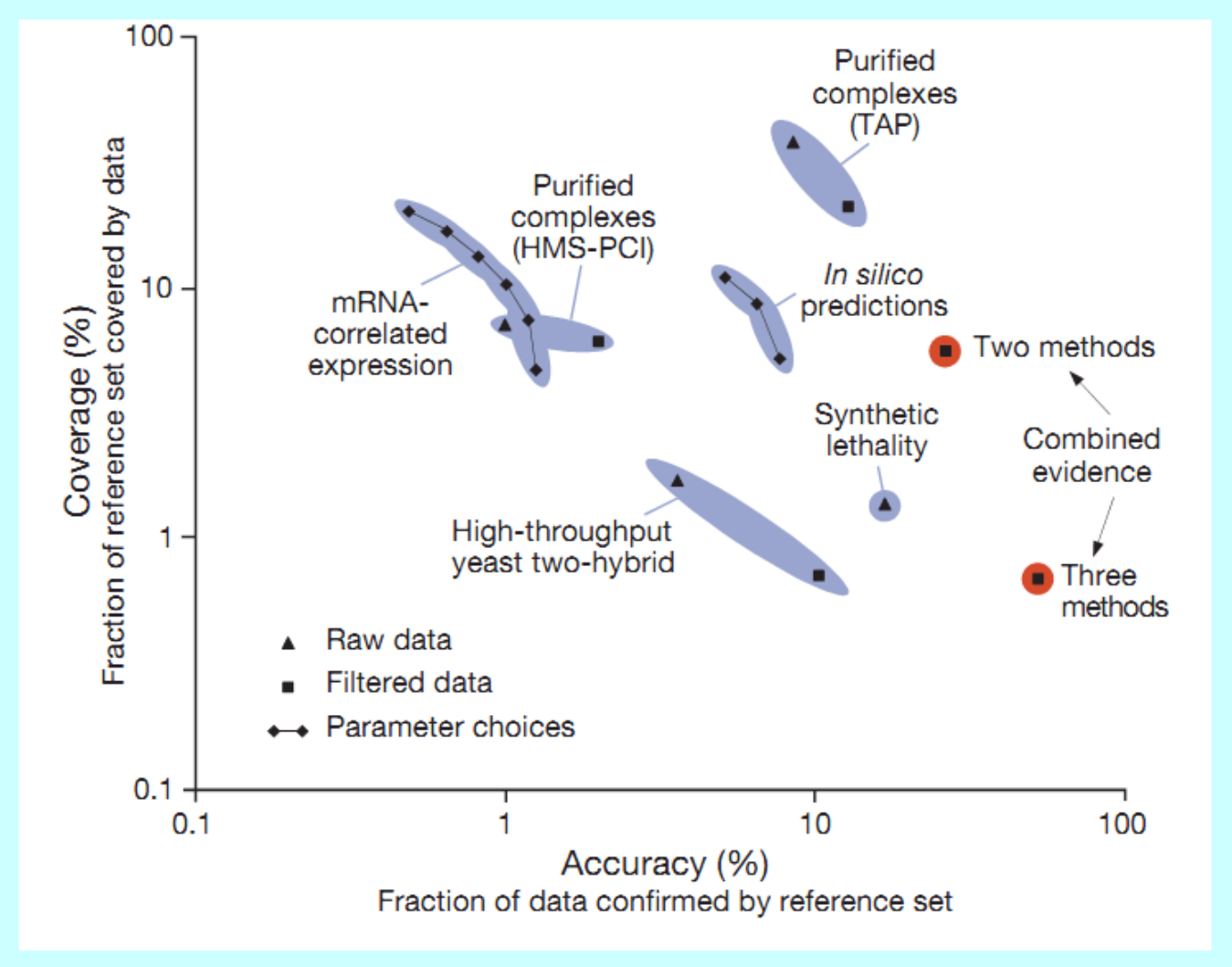


Figure 4. A schematic model for cross-linking cohort proteins that interact transiently, using a hetero-bifunctional, UV-activated, cleavable cross-linker.
In the first step, the maleimide function of the cross-linker reacts with a –SH group exposed on the surface of protein 1. Protein 1 then binds to and forms a complex with protein 2. Upon UV irradiation, the phenyl-azide function of the cross-linker is converted to a highly reactive nitrene which reacts with a primary amine group on the surface of protein 2. The cross-linking stabilizes the previously labile complex so that it can be purified. After purification, the sulfhydryl reducing reagent DTT is added, which cleaves the disulfide bond of the cross-linker and reduces both Cys. The two proteins are then separated by SDS-PAGE, and identified by MS analysis of the tryptic peptides.

The Plant Journal (2008) 53, 597–609.

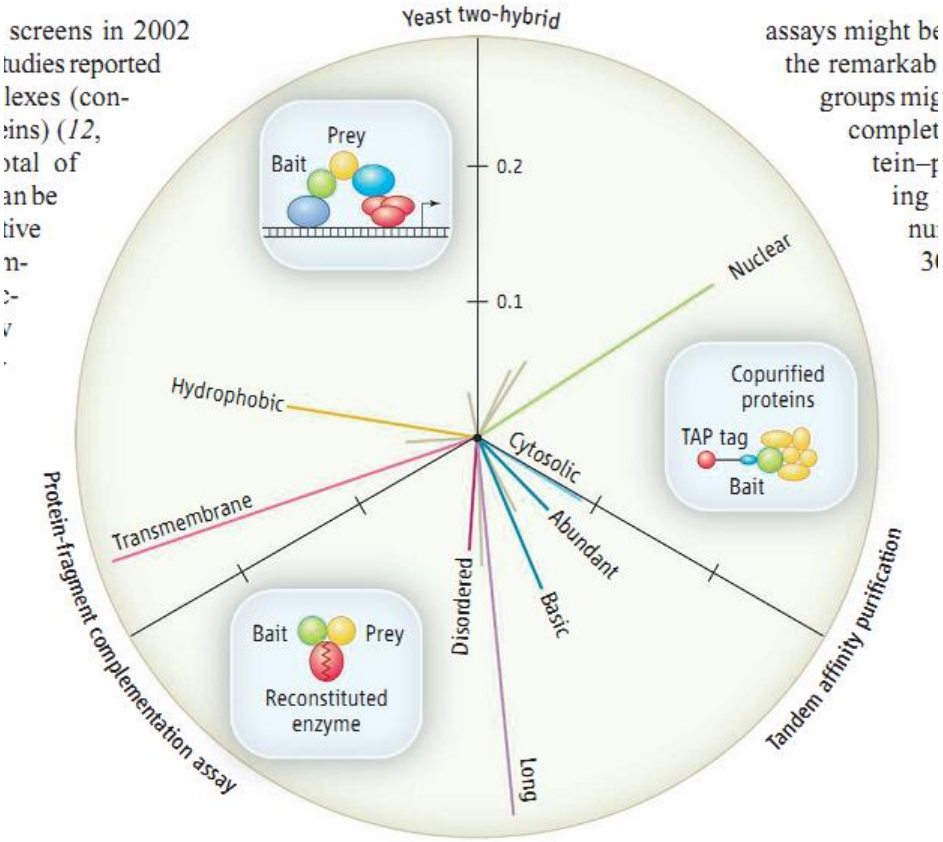
大规模比较显示蛋白互作实验结果重合不高



Nature (2002), 417: 399

蛋白互作的实验检测方法互补而非可比

Science (2008), 322: 56



Protein preferences. The three methods shown for detecting protein interactions function in fundamentally different ways and hence have different physiochemical constraints. Of 15 protein features tested for biases between the sets of proteins for which interactions were identified by each assay, 8 differed significantly (false-positive rate of <0.001): presence of transmembrane helices, hydrophobicity, nuclear and cytosolic localization, abundance, predicted isoelectric point, length, and intrinsic disorder. The other seven protein features are shown in gray. The average normalized scores (*Z* scores) for each of these features are shown, projected onto a plane in which each axis corresponds to one of the three methods for detecting interactions. The length of each line thus represents the strength of the bias.

蛋白互作数据库

通用的

DIP: Database of Interacting Proteins. website

IntAct: EMBL-EBI Protein Interaction. website

STRING: Known and Predicted Protein-Protein Interactions. website

MIND: the Molecular Interactions Database. website

HINT: *H*igh-quality *INT*eractomes. website

物种特异的

MIPS: The Mammalian Protein-Protein Interaction Database. website

HPRD: Human Protein Reference Database. website

SGD, wormbse, flybase, etc.

蛋白互作的理论预测方法

- 基于序列相似性的同源映射方法
- 基于序列统计特征的机器学习方法
- 基于共表达模式的预测方法
- 基因共进化模式的预测方法
- 基于基因组定位的预测方法
- 基于亚细胞定位的预测方法
- 基于蛋白质家族和结构域信息的预测方法
- 基于蛋白3D空间结构的同源模建方法
- 等

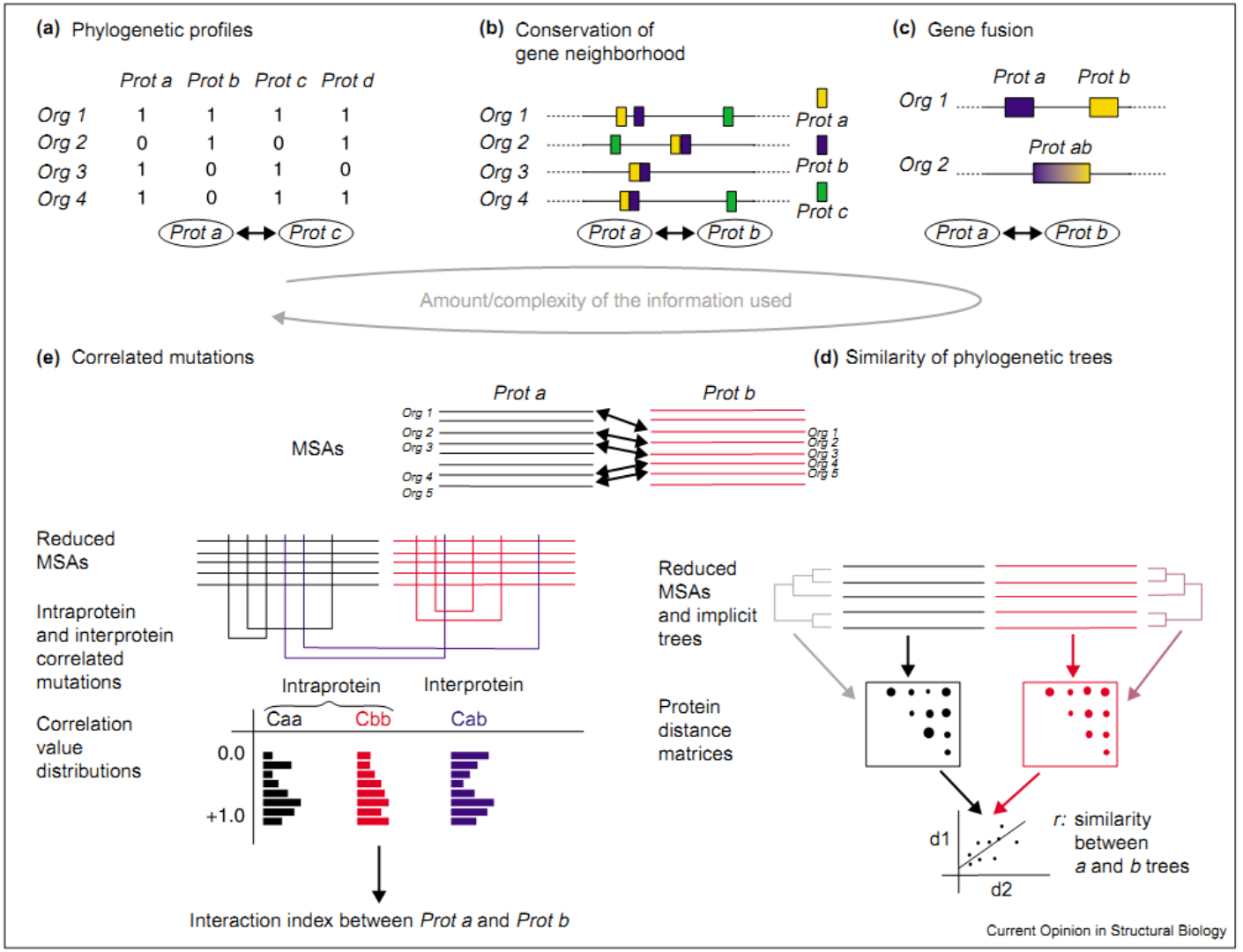
Wiki: protein–protein interaction prediction.

蛋白互作的理论预测方法

Table 1. Different Prediction Methods

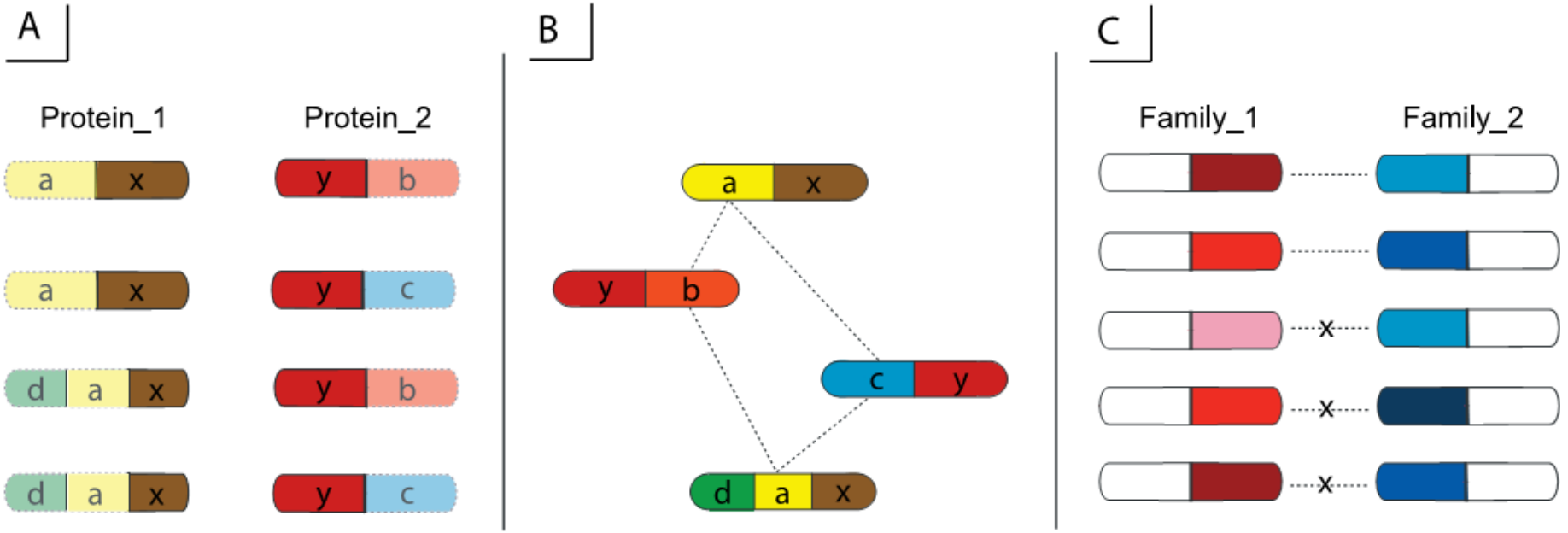
Method Name	Protein/Domain Interaction	Physical Interaction/ Functional Association
Gene co-expression	P	F
Synthetic lethality	P	F
Gene cluster and gene neighbor	P	F
Phylogenetic profile	P, D	F
Rosetta Stone	P	F
Sequence co-evolution	P, D	F
Classification	P, D	P
Integrative	P, D	P
Domain association	D	P
Bayesian networks	P, D	F, P
Domain pair exclusion	D	P
p-Value	D	P

基于进化的蛋白互作预测方法



Current Opinion in Structural Biology 2002, 12:368–373.

蛋白互作与结构域互作的关系

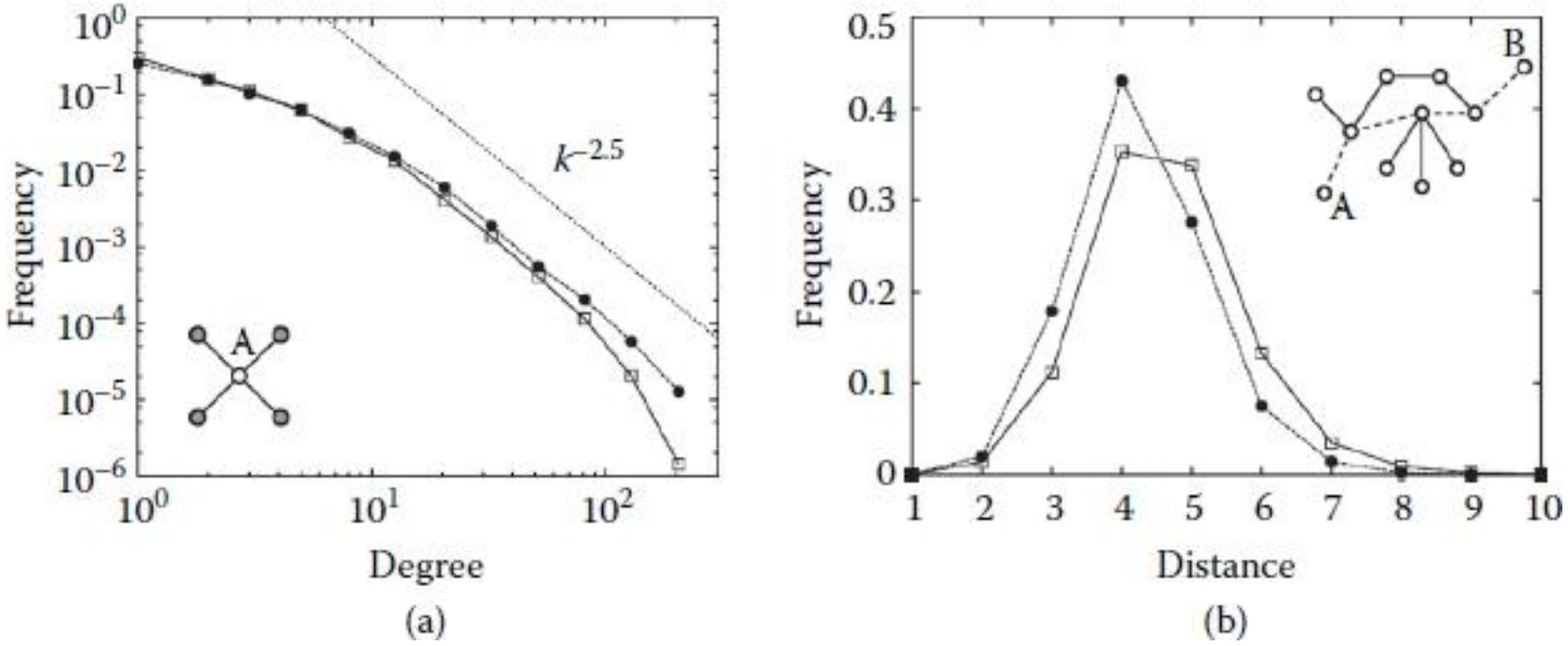


doi:10.1371/journal.pcbi.0030043.g002

Figure 2. Strategies to Predict Domain Interactions from Protein Interactions

PLoS Comput Biol 3(4): e43. doi:10.1371/journal.pcbi.0030043

蛋白互作网络的拓扑参数



人类蛋白互作网络的连接度分布和路径长度分布

http://www.ncbi.nlm.nih.gov/books/NBK56024/

蛋白互作网络中的模块性

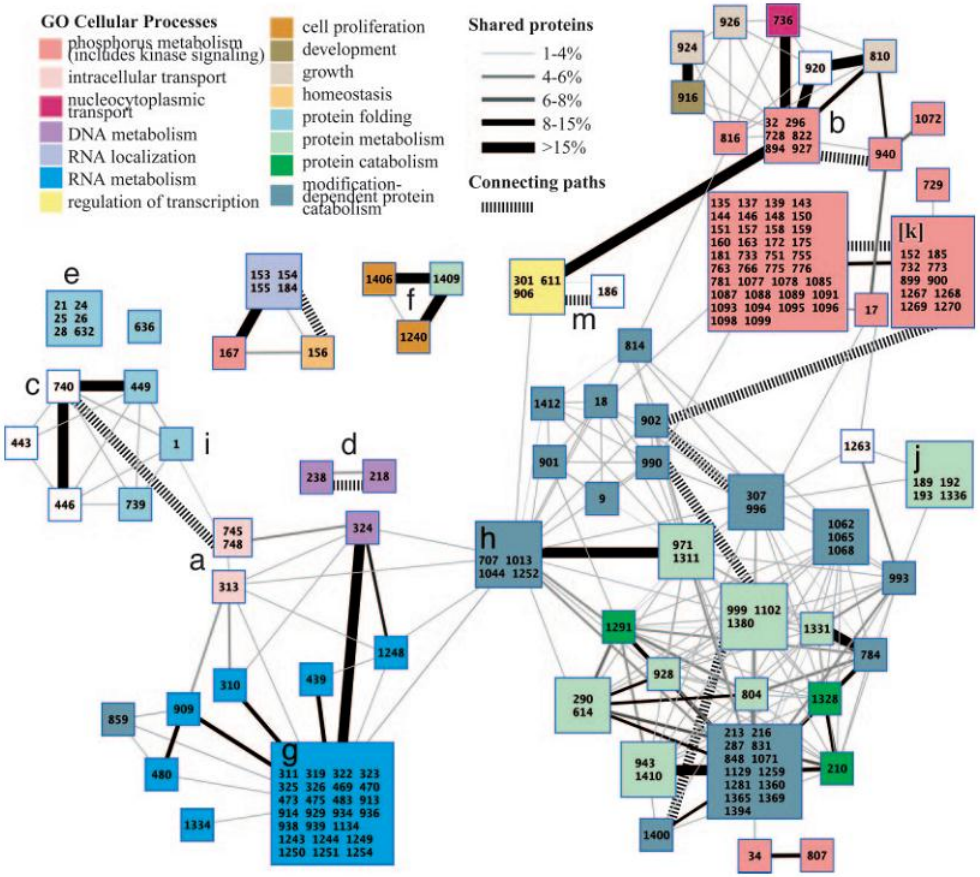
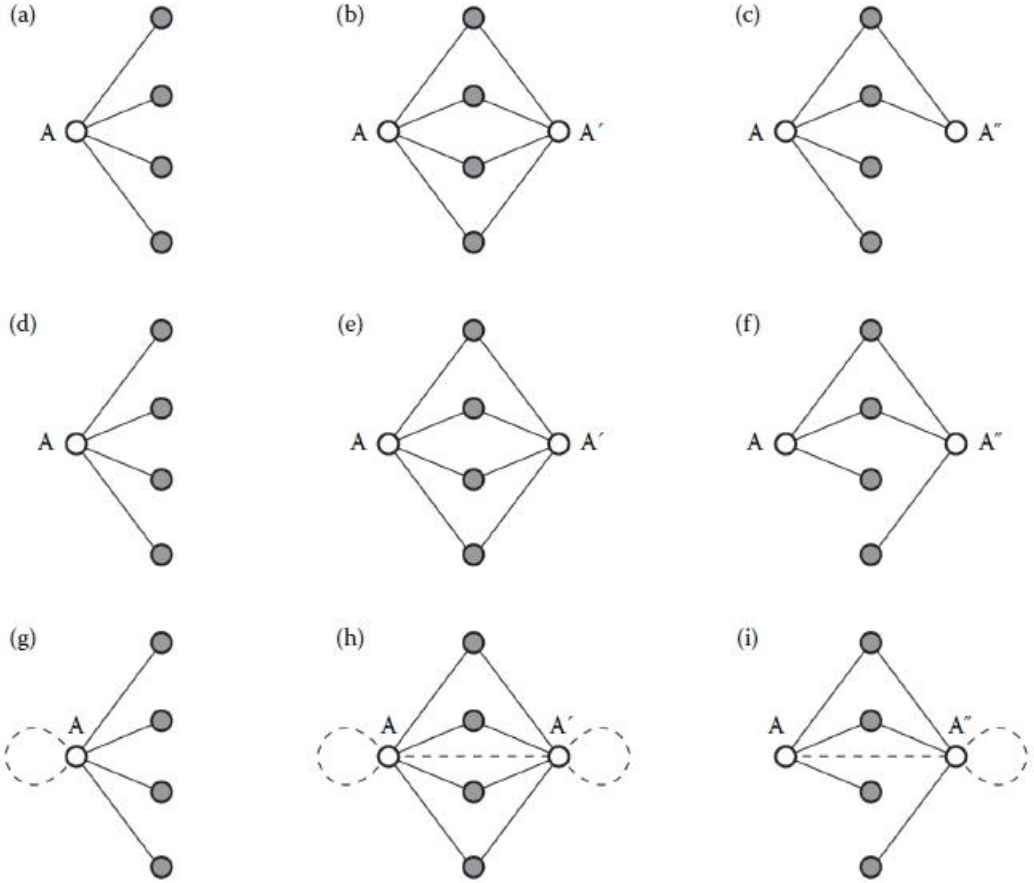


Fig. 3. Modular structure of conserved clusters among yeast, worm, and fly. Multiple network alignment revealed 183 conserved clusters, organized into 71 network regions represented by colored squares. Regions group together clusters that share >15% overlap with at least one other cluster in the group and are all enriched for the same GO cellular process ($P < 0.05$ with the enriched processes indicated by color). Cluster ID numbers are given within each square; numbers are not sequential because of filtering. Solid links indicate overlaps between different regions, where thickness is proportional to the percentage of shared proteins (intersection/union). Hashed links indicate conserved paths that connect clusters together. Labels a–k and m mark the network regions exemplified in Fig. 2.

PNAS (2005), 102: 1974

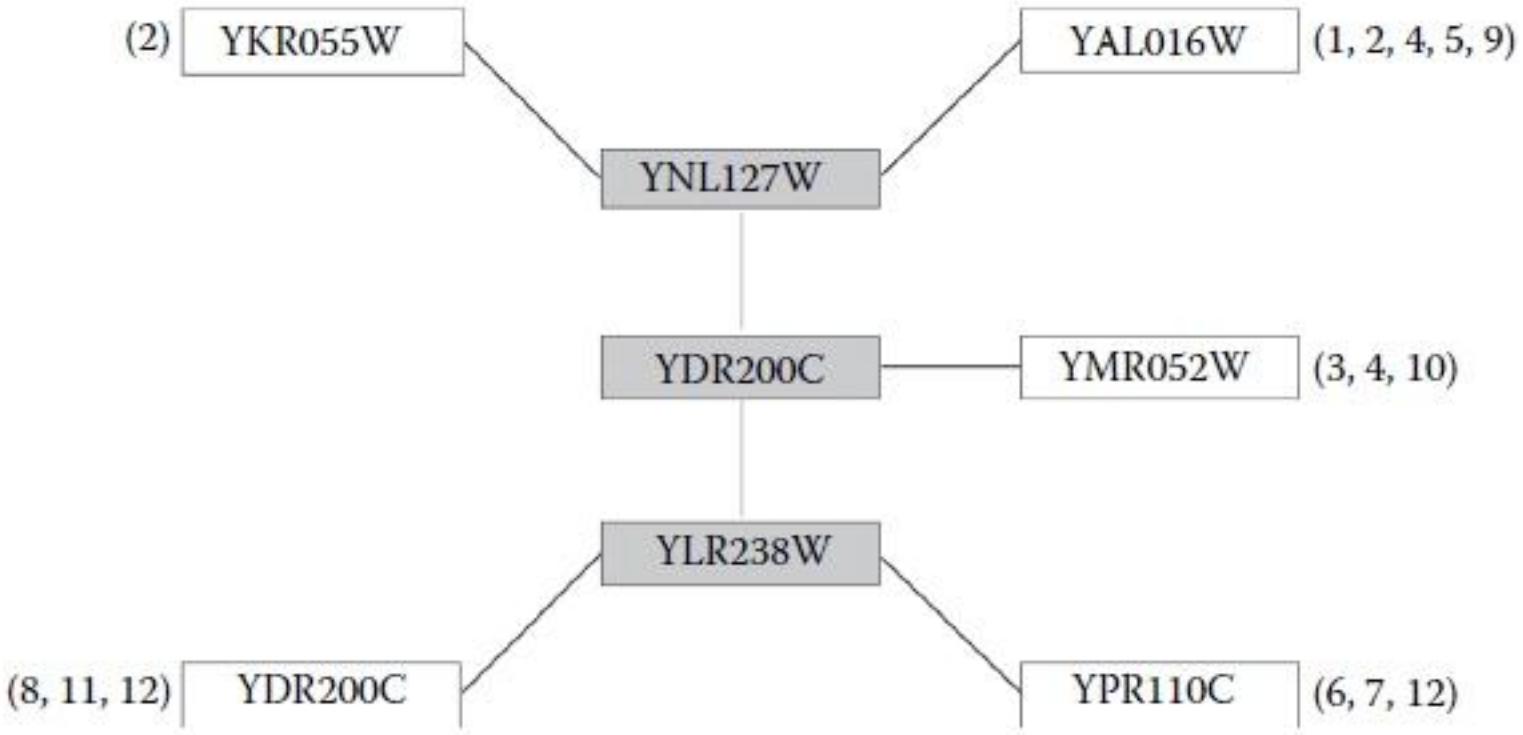
蛋白互作网络的进化



“基因复制-分化”模型示意图

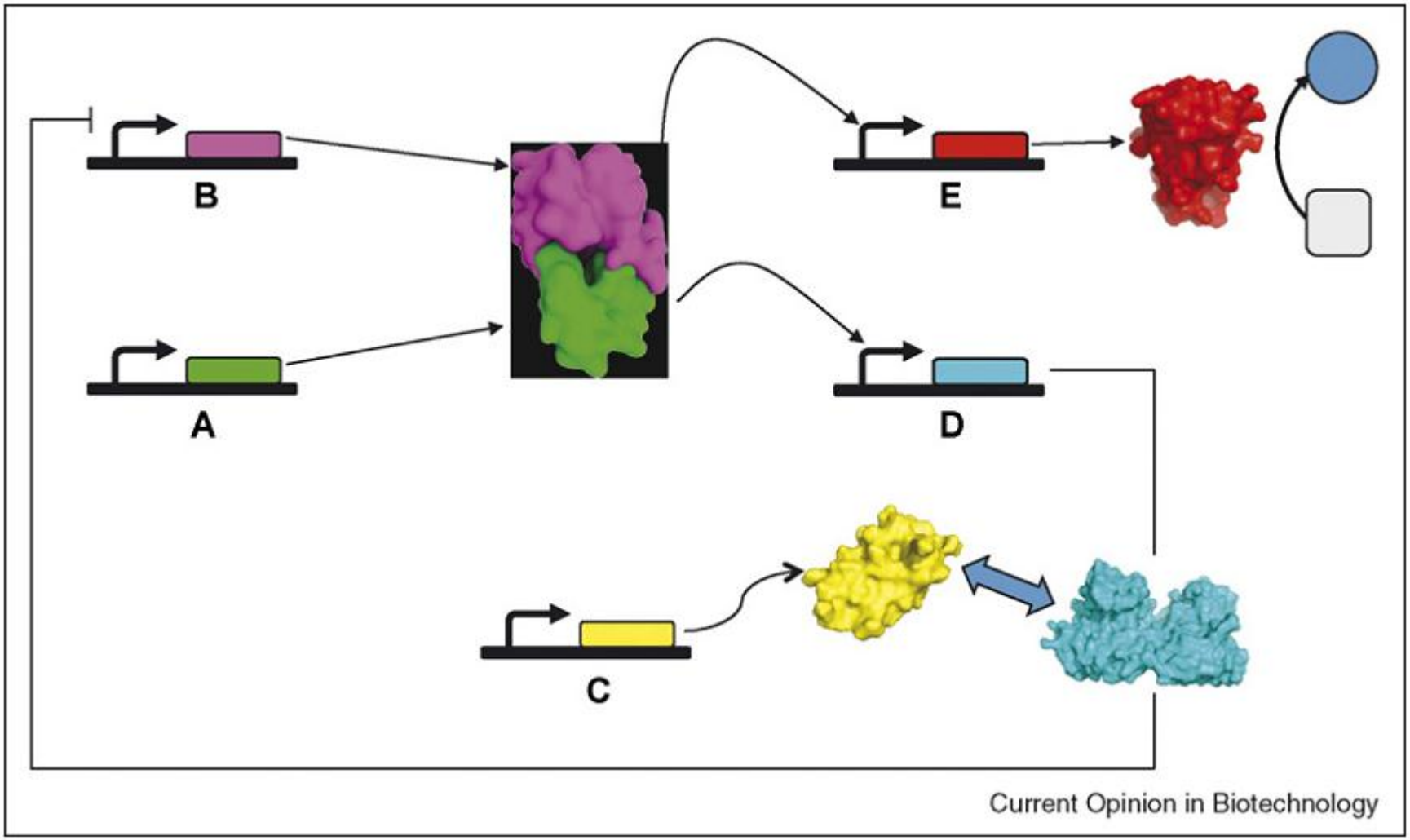
http://www.ncbi.nlm.nih.gov/books/NBK56024/

蛋白互作网络用于蛋白功能预测



http://www.ncbi.nlm.nih.gov/books/NBK56024/

蛋白互作与其它分子过程的关联



Current Opinion in Biotechnology 2008, 19:396–403.

酵母细胞周期中的蛋白互作

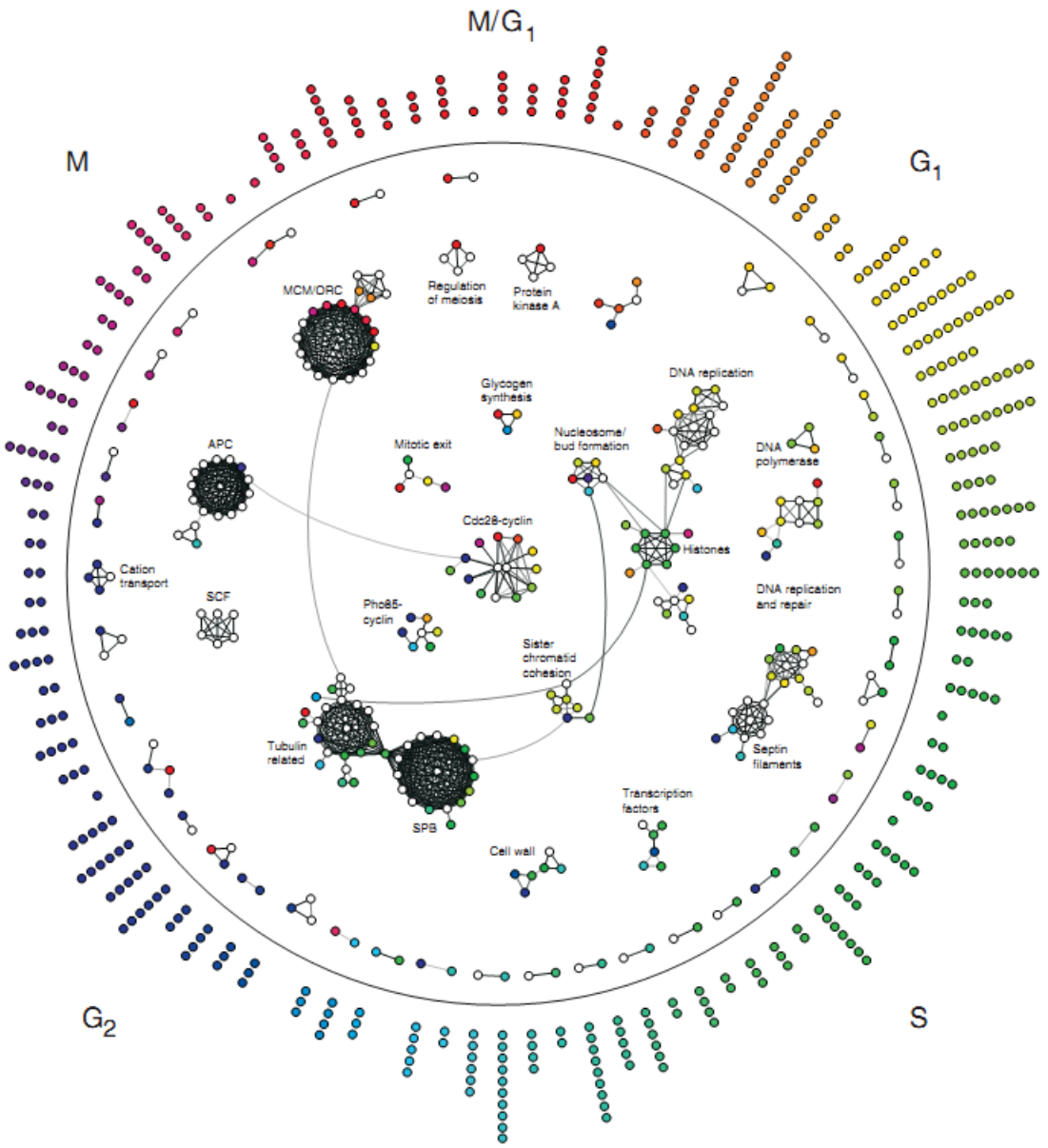
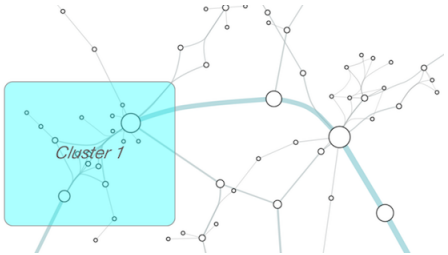


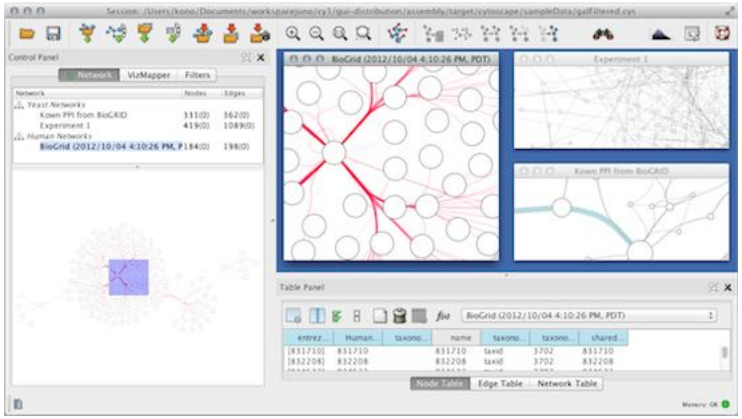
Fig. 1. Temporal protein interaction network of the yeast mitotic cell cycle. Cell cycle proteins that are part of complexes or other physical interactions are shown within the circle. For the dynamic proteins, the time of peak expression is shown by the node color; static proteins are represented by white nodes. Outside the circle, the dynamic proteins without interactions are both positioned and colored according to their peak time and thus also serve as a legend for the color scheme in the network. More detailed versions of this figure (including all protein names) and the underlying data are available online at www.cbs.dtu.dk/cellcycle.

Cytoscape软件介绍

Cytoscape is an open source bioinformatics software platform for visualizing molecular interaction networks and integrating with gene expression profiles and other state data. Additional features are available as plugins. Plugins are available for network and molecular profiling analyses, new layouts, additional file format support and connection with databases and searching in large networks. Plugins may be developed using the Cytoscape open Java software architecture by anyone and plugin community development is encouraged. (from wiki)

Cytoscape





Current Version: 3.x

http://www.cytoscape.org/