View metadata, citation and similar papers at core.ac.uk
brought to you by CORE
provided by Elsevier - Publisher Connector

FEBS 29314 FEBS Letters 579 (2005) 1821–1827

Minireview

From protein networks to biological systems

Peter Uetz[a,1], Russell L. Finley Jr.[b,*]

[a] Research Center Karlsruhe, Institute of Genetics, P.O. Box 3640, D-76021 Karlsruhe, Germany
[b] Center for Molecular Medicine & Genetics, Wayne State University School of Medicine, 540 East Canfield Avenue, Detroit, MI 48201, USA

Accepted 31 January 2005

Available online 7 February 2005

Edited by Robert Russell and Giulio Superti-Furga

Abstract A system-level understanding of any biological process requires a map of the relationships among the various molecules involved. Technologies to detect and predict protein interactions have begun to produce very large maps of protein interactions, some including most of an organism's proteins. These maps can be used to study how proteins work together to form molecular machines and regulatory pathways. They also provide a framework for constructing predictive models of how information and energy flow through biological networks. In many respects, protein interaction maps are an entrée into systems biology.
© 2005 Federation of European Biochemical Societies. Published by Elsevier B.V. All rights reserved.

Keywords: Interaction map; Protein network; Pathway; System

1. Introduction

Systems Biology should give us the tools to model how genes, gene products, and other molecules work together to mediate biological processes. Use of such tools, and indeed their very development, requires, for each biological process, lists of the molecules involved and their interconnections. The genes and proteins predicted from genome sequences have provided a long list of parts (genes and gene products), and new technologies have begun to define lists of other molecules not directly encoded by the genome that are present in cells and tissues at particular times. New computational and experimental technologies have begun to produce enormous datasets representing interactions between the parts. For the moment, most of the interaction data comes from technologies to detect physical or functional interactions between genes and proteins. Here, we will review some of the sources of these data and consider how the quantity and quality of the available interaction data may impact systems-level studies.

2. Protein–protein interactions

The prominent role that protein–protein interactions play in most biological processes, combined with the fact that we know so little about the functions of most proteins, has inspired efforts to map interactions on a proteome-wide scale (e.g., for all of the proteins encoded by a genome) [1]. To date, most of the interactions that have been detected experimentally have relied on one of two technologies, the yeast two-hybrid system [2] and mass spectrometry (MS) identification of proteins that co-affinity purify (co-AP) with a bait protein [3]. The two technologies detect complementary types of interactions. Co-AP/MS identifies the constituents of multi-protein complexes but does not reveal the individual binary contacts that make up each complex. Without data on the constituent binary contacts, the possible paths of energy or information flow through the complex and its relationship to other cellular components may not be apparent. Yeast two-hybrid data, on the other hand, identifies likely binary interactions that may suggest possible paths through a pathway or complex, but cannot reveal the constituents of multiprotein complexes. Thus, both types of data will be important for understanding protein and pathway function, and ideally both approaches would be performed on a proteome-wide scale.

Yeast two-hybrid screens aiming to cover entire proteomes, or at least very large numbers of proteins, have detected thousands of interactions for a few eukaryotic model organisms (Table 1), bacteria and phage [4,5] and viruses [6]. By contrast, proteome-wide co-AP/MS screens have been conducted only in yeast (Table 1), where most of the proteome could be easily affinity tagged through the use of homologous recombination. Co-AP/MS data for other organisms is only just beginning to emerge through the use of high throughput cloning [7] and the expression of large sets of tagged proteins in tissue culture cells (e.g. [8,9]). Thus, it is likely that we will begin to see protein complex data for humans and other metazoans in similar quantities as the yeast studies have produced.

3. How complete are current protein interaction datasets?

Despite the volumes of interaction data produced, several independent analyses have shown that the data from the large two-hybrid and co-AP/MS screens is far from complete. Various authors have estimated that the roughly 6000 yeast proteins are connected by 12 000–40 000 interactions [10–12], yet the high throughput screens have detected only a small fraction of those numbers (Table 1). Another clue comes from the lack of overlap among the different datasets for a particular proteome. For example, in Table 1, the overlap among the large two-hybrid screens for yeast was only 6 interactions

* Corresponding author. Fax: +1 313 577 5218.
E-mail addresses: peter.uetz@itg.fzk.de (P. Uetz), rfinley@wayne.edu (R.L. Finley Jr.).

[1] Fax: +49 7247 82 3354.

0014-5793/$30.00 © 2005 Federation of European Biochemical Societies. Published by Elsevier B.V. All rights reserved.
doi:10.1016/j.febslet.2005.02.001

Table 1
Large protein interaction screens for eukaryotes

Organism (genes)	Method	Interactions[a]	Proteins	Reference
Yeast (~6000)	Yeast two-hybrid	967	1004	[63]
	Yeast two-hybrid	4549	3278	[13]
	Yeast two-hybrid	420	271	[64,65]
	Co-AP/MS	9421	1665	[66]
	Co-AP/MS	3878	1578	[67]
Drosophila (~14 000)	Yeast two-hybrid	20 405	7048	[49]
	Yeast two-hybrid	1814	488	[14]
Worm (~20 000)	Yeast two-hybrid	4027	1926	[68]

[a]For two-hybrid screens, the approximate number of unique binary interactions is shown. For co-AP/MS screens, the approximate number of binary interactions that would result if each bait protein contacted every protein that co-purified with it (the "hub and spoke" model) is shown. Data can be retrieved from one of the databases cited [42–44].

[13] while the overlap between the screens for *Drosophila* was a measly 28, or less than 2% of the smallest data set [14]. The co-AP/MS data is not much different. For example, when results from the two large-scale studies are compared, the number of interactions common to both datasets is less than 9% of the total in both datasets [15]. Data from the high throughput screens also fails to overlap significantly with published "low throughput" studies, which are generally considered to be less subject to false positives and false negatives. Such analyses have led some authors to estimate false negative rates as high as ~85% in large yeast two-hybrid screens and 50% in co-AP/MS screens [16,17]. These results suggest that many more interactions could be detected by more exhaustive application of these technologies. In addition, there is a need for improved or new high throughput technologies to identify interactions that may be difficult to detect with two-hybrid or co-AP/MS, such as interactions involving membrane proteins.

4. Physical and functional interactions

Comparison of the data from yeast two-hybrid and co-AP/MS provides an example of an important distinction between two types of interaction data: physical interactions (A touches B) and functional interactions (A functions with B in some biological process). A functional relationship may or may not correspond to a direct physical interaction. Thus, physical and functional interactions are two distinct though partially overlapping types of interactions and the distinction is likely to be important for the development of systems-level models of protein networks and pathways. Yeast two-hybrid is an experimental approach to detect physical interactions. Co-AP/MS detects group of proteins in stable complexes, implying that they function together. Another example of a functional but not necessarily physical interaction is a genetic interaction, in which the combination of alleles of two different genes has specific phenotypic consequences. This is often taken to suggest that the two genes function in the same or parallel pathways affecting a particular biological process. Thus, a genetic interaction is a measured functional interaction that may or may not correspond to a physical interaction, but that could be usefully represented as a connection between the two genes or gene products. Ongoing large-scale screens in yeast have mapped thousands of genetic interactions [18]. Combination of genetic and physical interaction data is a powerful approach to mapping pathways [18,19].

5. Predicted and experimentally measured interactions

The increasing use of computational approaches to predict protein interactions has led to additional large datasets (e.g. [20]) and to another distinction between two types of interaction data: experimentally measured and predicted. Predicted interactions can also be classified as either physical or functional. Gene expression profiles have been used, for example, to infer functional interactions among gene products, based on the assumption that proteins that function together in the same pathway or complex should be frequently expressed together; which is supported by data for stable protein complexes [21,22]. Similarly, genes whose coexpression profiles are conserved through evolution are often functionally related [23,24], as are genes that are co-conserved from species to species [25–27]. The functional links between proteins in each of these cases may be direct or very general; they may suggest roles in the same pathways, or in distinct cellular systems that are concomitant but that have very few direct molecular connections. Genetic interactions have also been predicted based on physical interactions, gene expression, protein localization, and other experimental data [28,29]. Numerous methods for predicting physical protein–protein interactions have also been developed [22,30–35]. One very powerful approach takes advantage of the large number of experimentally measured interactions available for organisms like yeast and *Drosophila* to predict interactions in other organisms [36]. Simply put, the approach predicts that two proteins will interact if their orthologs were shown to interact; such conserved interactions have been referred to as interologs [37–39]. This approach has been used, for example, to predict 70 000 interactions involving proteins encoded by a third of the human genes [40]. Several other studies have begun to effectively integrate genomic and proteomic data to make increasingly accurate interaction predictions [33,41]. The further development and use of in silico approaches to map interactions seems particularly important in light of the shortcomings of high throughout experimental detection systems.

6. Protein interaction maps

The wealth of data from high throughput screening and other studies has begun to be consolidated into centralized, standardized databases. Three of the largest public database repositories for interaction data are BIND, DIP, and IntAct

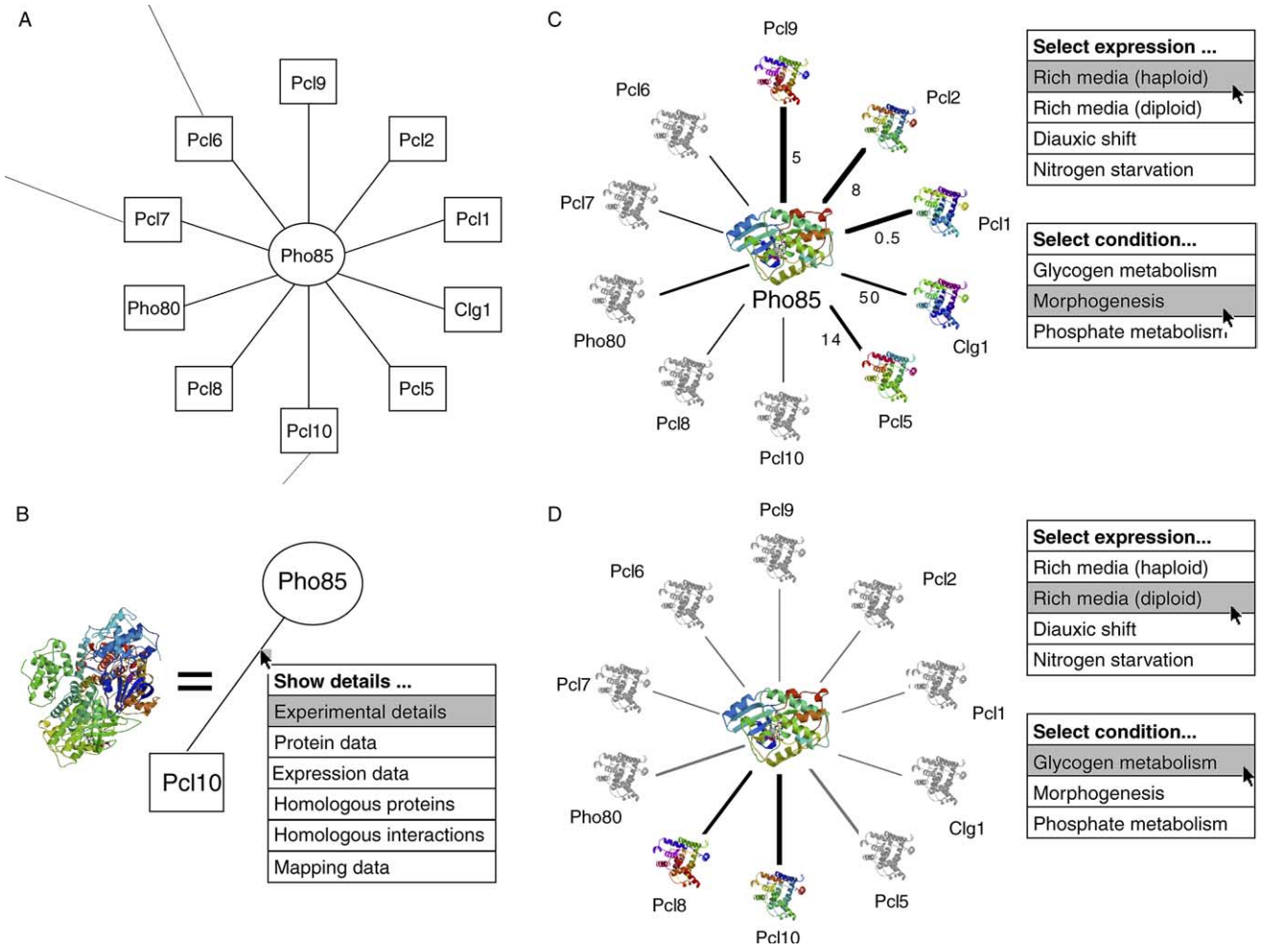

Fig. 1. Interaction maps today and tomorrow. (A) Typical representation of a protein–protein interaction map. (B) Proteins are usually shown as nodes (e.g., circles and boxes) and interactions as edges (lines) connecting them. This organizes information so that attributes of both proteins and interactions are easily accessible, for example, by hypertext links, but fails to capture structural information. However, additional information could be made easily accessible through pop-up menus by clicking on the edges (like here) or the nodes. (C) and (D) Ideally, a protein interaction map visualization tool would allow the structures of proteins and interaction interfaces to be expanded and browsed, and would also provide access to more global interaction attributes (e.g., conditions under which a certain set of interaction can be found, experimental conditions, dissociation constants, expression levels, etc.). Based on interactions published by Measday et al. [69], Uetz et al. [63], and reviewed in [70].

[42–44]. These allow researchers online access to browse and download data in a standardized format [45]. Sets of interaction data can be viewed as graphs or maps in which each gene/protein is a node and each interaction is a line connecting two nodes (Fig. 1). The importance of this view has led to use of the term "interaction map" to refer generically to interaction datasets. The map view provides not only an intuitive interface for biologists to explore the data, but also a formal mathematical framework for computational biologists to explore the properties of interaction networks. However, before interaction maps can be used to represent biological networks, their limitations must be considered.

In addition to the problem of false negatives discussed previously, most interaction maps and particularly those from high throughput screens have false positives. Estimates of false positive rates vary widely, in part because of the difficulty in definitively demonstrating that any particular interaction does not have a biological function. Because the false positive rates may be substantial, the maps from high throughput studies might be usefully regarded as the results from a first pass filter, which reduces the possible search space for functionally important interactions. Thus, the question becomes how to identify the more likely true positives. Several studies have confirmed the general principle that interactions detected in multiple screens and by different techniques or in different species are more likely to be true positives than those only found once or twice [17]. Due to the high rates of false negatives in high throughput screens, however, there has been very little overlap between different datasets, thus, limiting the opportunities for such experimental cross-validation. Alternatively, a variety of confidence scoring systems have been developed that calculate the likelihood of an interaction being a true positive, based on various parameters, including attributes of the proteins and the specific assays, whether the interaction was detected by other technologies or screens, and network topology [20,46–50]. However, thus far most of these scoring systems are specific to particular datasets or methodologies, and no universal system has yet been effective.

Another limitation of most protein interaction maps is that each node generally represents some generic version of a

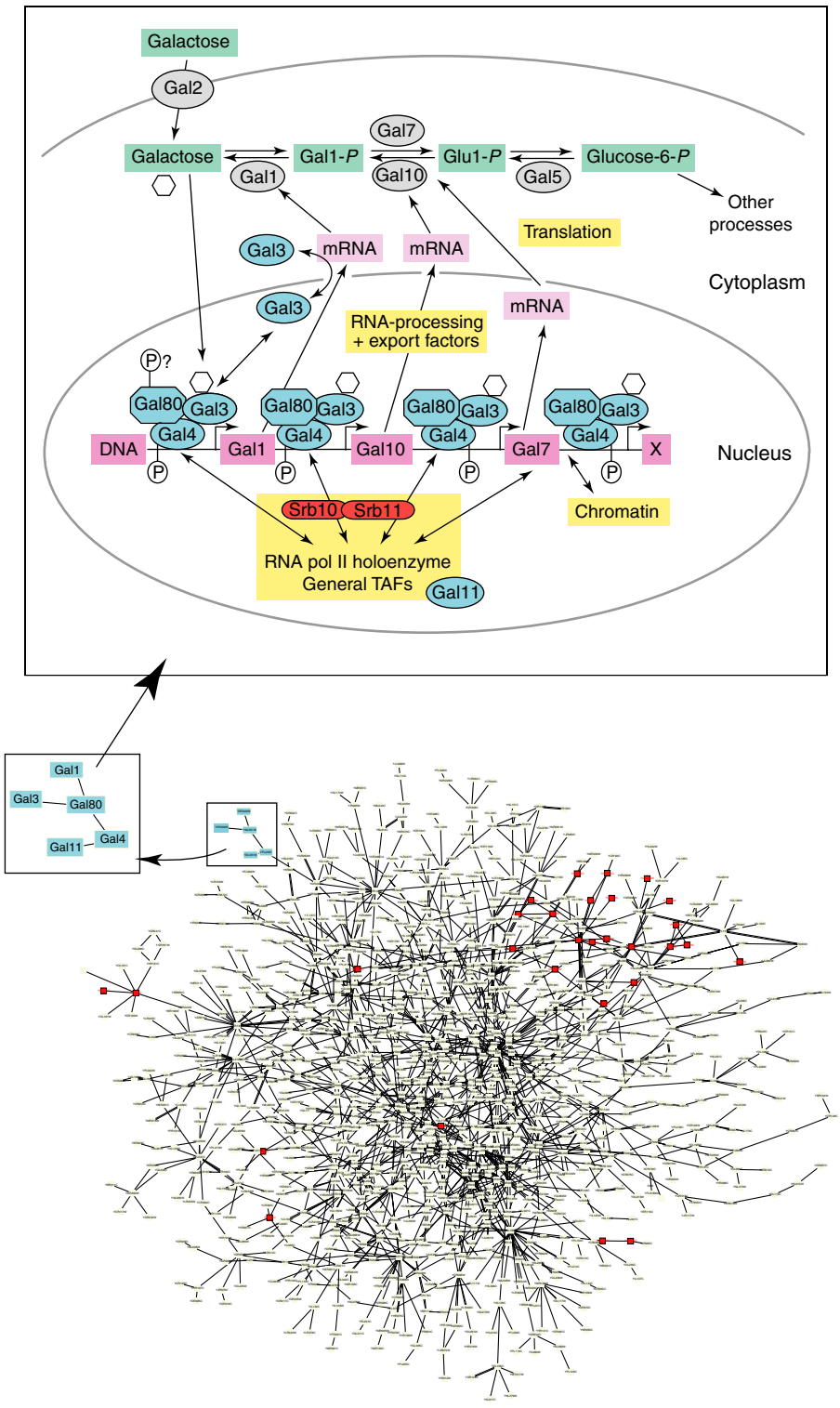

Fig. 2. Integrating protein networks and other biological information. (Bottom) The map of interactions among 1200 yeast proteins represents only the tip of the information iceberg. Highlighted in blue are proteins involved in galactose regulation, which in the map are found in a topological cluster (see text). A cluster of cytoskeletal proteins (highlighted in red) is also visible. (Top) Integrating the protein interaction network with the metabolic network (shown as green compounds and grey enzymes), the gene regulatory network (shown as pink genes and blue transcription factors), and the signaling network indicated by the Cyclin-dependent kinase Srb10 and its cyclin Srb11 (red). Yellow boxes indicate functional modules that involve additional protein interactions within complexes and with other proteins (not shown). Modified after Tucker et al. [12] and Rohde et al. [71].

cellular protein, without regard to the various splice variants or post-translationally modified forms that may exist. Isoforms could interact differently from the form that was actually used in the assay, which in many cases is unclear. This is particularly true for assays that use only one or a small number of the possible alternative transcripts from each gene. Thus, many

so-called "protein" interactions maps are actually gene or locus interaction maps, which tell us only that one or more of the proteins encoded by one locus is capable of interacting with one or more of the proteins encoded by another locus. Nevertheless, such maps have proven to be useful as starting points for additional studies, particularly if the caveats are borne in mind.

7. Using the interaction maps for systems biology

A complete systems-level understanding of any biological process may require more input data than current technologies can offer. Intuitively, we imagine that we could model a system best only after knowing all of the molecules involved, their concentrations, how they fit together, the effect of each individual part on its neighbors, and dynamic parameters such as how concentrations, interactions, and mechanics change over time. But this seems unrealistic given the fact that the high throughput technologies for measuring many of these parameters are still on the drawing board, if they exist at all. Do we really need to know all of the details of a process to be able to develop a useful systems-wide understanding or to have a predictive model? Analysis of protein interaction maps has suggested that even sparse data can be used to derive initial, rudimentary models of biological networks.

Topological analyses, for example, initially of metabolic pathways and subsequently of protein interaction maps, began to reveal some common properties of biological networks [51,52]. These initial studies suggested the exciting possibility that cellular networks may be organized according to some general principles that could be understood without a detailed knowledge of all the constituent proteins and interactions [53,54]. Moreover, analysis of network topology can provide insights into protein and pathway function. For example, protein networks contain highly connected hub proteins, which have been shown to correlate with evolutionarily conserved proteins, and in yeast with proteins encoded by essential genes [51,55,56]. Thus, a protein's relative position in a network has implications for its function and importance. Analysis of topology also reveals clusters of highly interconnected proteins that correlate with conserved functional modules (Fig. 2), such as protein complexes or signaling pathways [57–59]. Thus, even the currently available noisy protein interaction maps can be used to explore the hierarchical organization of biological networks and to reveal interconnected modules that control specific biological processes. As these modules are defined and further elaborated, understanding them and their higher order organization will increasingly rely on advances in information technology.

8. Perspectives: iCell-TV

How can biologists access and integrate the deluge of proteomics data to help them understand biology? While this information should help drive the generation of hypotheses and hypothesis-testing research, we may be generating data faster than we are learning how to use it. Tools for accessing and analyzing molecular interaction data have just begun to emerge over the past few years. Several "visualization" tools and graphing programs, for example, allow users to construct a map of interactions [60–62]. These programs allow exploration and ad hoc analyses of interaction data but they rarely incorporate all of the useful available information about the molecules and interactions they represent (e.g., see Figs. 1 and 2). Moreover, they usually fail to capture the essential dynamic properties of biological networks. Animated cartoons, on the other hand, can provide at least a qualitative representation of the dynamics of a process (see, for example, Fig. 3). However, such oversimplification does not capture the details of the system or facilitate quantitative modeling. In a way, visualization of molecular networks is where word processing was in the early 1980s.

To help us model biological processes, and to visualize and manipulate those models, we need programs to generate more dynamic and realistic representations of biological events and structures. We need what might be called "interactive Cell-TV" to visualize and manipulate models of cellular events and behavior. Importantly, iCell-TV must operate across several scales of time and space to allow biologists to navigate all available relevant information. Such a system, for example, might allow users to explore the changes in the molecular structure resulting from a post-translational modification, zoom out to witness the subsequent changes in network and pathway dynamics, and then change time scales to observe organelle movement or cell behavior. The number and complexity of the experiments that must be done to test hypotheses

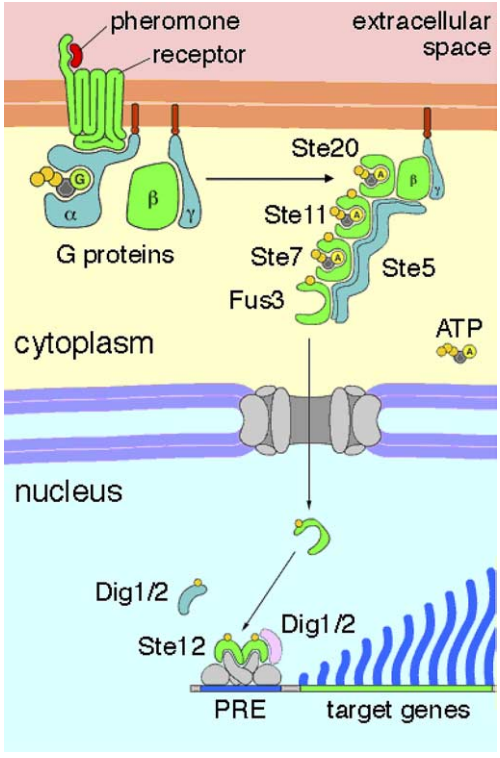

Fig. 3. Visualizing the dynamics of protein interaction and signaling networks. The pheromone signaling pathway in yeast is a highly dynamic process that involves numerous protein interactions, phosphorylation events, and small-molecule interactions involving ATP and GTP. Typical textbook (i.e., static) representations like this do not reflect the dynamics of this process. A more realistic representation is available through animation, as shown at http://www.bioveo.com/MAPK/MAPk.htm. Simplified from an animation by Tom Dallman, by permission of the author.

coming from network analyses are likely to be costly and inefficient. The next generation of biological information management systems must, therefore, allow us to do biology truly in silico. For this to be possible, they must enable the development and manipulation of quantitative models, which are often initially based on a qualitative understanding. However, it is often the case that about the time we understand a system well enough to be able to model it, it becomes too hard to understand in a qualitative sense. A system for navigating qualitative information based on quantitative data would give users the ability not only to understand the complexity of biological processes but also to manipulate those processes, to construct new models, and to test new hypotheses. Zoom in, change a K_d or a V_{max}, then zoom out and watch what happens to the system. This would be systems biology for the rest of us, and would open biological inquiry to a vast resource of creativity.

Acknowledgments: We thank Tom Dallman for permission to use his pheromone pathway animation.

References

[1] Phizicky, E., Bastiaens, P.I., Zhu, H., Snyder, M. and Fields, S. (2003) Protein analysis on a proteomic scale. Nature 422, 208–215.
[2] Fields, S. and Sternglanz, R. (1994) The two-hybrid system: an assay for protein–protein interactions. Trends Genet. 10, 286–292.
[3] Pandey, A. and Mann, M. (2000) Proteomics to study genes and genomes. Nature 405, 837–846.
[4] Rain, J.C., et al. (2001) The protein–protein interaction map of *Helicobacter pylori*. Nature 409, 211–215.
[5] Bartel, P.L., Roecklein, J.A., SenGupta, D. and Fields, S. (1996) A protein linkage map of *Escherichia coli* bacteriophage T7. Nat. Genet. 12, 72–77.
[6] Uetz, P., Rajagopala, S.V., Dong, Y.A. and Haas, J. (2004) From ORFeomes to protein interaction maps in viruses. Genome Res. 14, 2029–2033.
[7] Marsischky, G. and LaBaer, J. (2004) Many paths to many clones: a comparative look at high-throughput cloning methods. Genome Res. 14, 2020–2028.
[8] Bouwmeester, T., et al. (2004) A physical and functional map of the human TNF-alpha/NF-kappa B signal transduction pathway. Nat. Cell Biol. 6, 97–105.
[9] Brajenovic, M., Joberty, G., Kuster, B., Bouwmeester, T. and Drewes, G. (2004) Comprehensive proteomic analysis of human Par protein complexes reveals an interconnected protein network. J. Biol. Chem. 279, 12804–12811.
[10] Walhout, A.J., Boulton, S.J. and Vidal, M. (2000) Yeast two-hybrid systems and protein interaction mapping projects for yeast and worm. Yeast 17, 88–94.
[11] Grigoriev, A. (2003) On the number of protein–protein interactions in the yeast proteome. Nucleic Acids Res. 31, 4157–4161.
[12] Tucker, C.L., Gera, J.F. and Uetz, P. (2001) Towards an understanding of complex protein networks. Trends Cell Biol. 11, 102–106.
[13] Ito, T., Chiba, T., Ozawa, R., Yoshida, M., Hattori, M. and Sakaki, Y. (2001) A comprehensive two-hybrid analysis to explore the yeast protein interactome. Proc. Natl. Acad. Sci. USA 98, 4569–4574.
[14] Stanyon, C.A., et al. (2004) A *Drosophila* protein-interaction map centered on cell-cycle regulators. Genome Biol. 5, R96.
[15] Cornell, M., Paton, N.W. and Oliver, S.G. (2004) A critical and integrated view of the yeast interactome. Comp. Funct. Genom. 5, 382–402.
[16] Edwards, A.M., Kus, B., Jansen, R., Greenbaum, D., Greenblatt, J. and Gerstein, M. (2002) Bridging structural biology and genomics: assessing protein interaction data with known complexes. Trends Genet. 18, 529–536.
[17] von Mering, C., Krause, R., Snel, B., Cornell, M., Oliver, S.G., Fields, S. and Bork, P. (2002) Comparative assessment of large-scale data sets of protein–protein interactions. Nature 417, 399–403.
[18] Tong, A.H., et al. (2004) Global mapping of the yeast genetic interaction network. Science 303, 808–813.
[19] Tewari, M., et al. (2004) Systematic interactome mapping and genetic perturbation analysis of a *C. elegans* TGF-beta signaling network. Mol. Cell 13, 469–482.
[20] von Mering, C., et al. (2005) STRING: known and predicted protein–protein associations, integrated and transferred across organisms. Nucleic Acids Res. 33 (Database Issue), D433–D437.
[21] Jansen, R., Lan, N., Qian, J. and Gerstein, M. (2002) Integration of genomic datasets to predict protein complexes in yeast. J. Struct. Funct. Genom. 2, 71–81.
[22] Jansen, R., et al. (2003) A Bayesian networks approach for predicting protein–protein interactions from genomic data. Science 302, 449–453.
[23] Bergmann, S., Ihmels, J. and Barkai, N. (2004) Similarities and differences in genome-wide expression data of six organisms. PLoS Biol. 2, E9.
[24] Stuart, J.M., Segal, E., Koller, D. and Kim, S.K. (2003) A gene-coexpression network for global discovery of conserved genetic modules. Science 302, 249–255.
[25] Pellegrini, M., Marcotte, E.M., Thompson, M.J., Eisenberg, D. and Yeates, T.O. (1999) Assigning protein functions by comparative genome analysis: protein phylogenetic profiles. Proc. Natl. Acad. Sci. USA 96, 4285–4288.
[26] Date, S.V. and Marcotte, E.M. (2003) Discovery of uncharacterized cellular systems by genome-wide analysis of functional linkages. Nat. Biotechnol. 21, 1055–1062.
[27] Bowers, P.M., Cokus, S.J., Eisenberg, D. and Yeates, T.O. (2004) Use of logic relationships to decipher protein network organization. Science 306, 2246–2249.
[28] Wong, S.L., et al. (2004) Combining biological networks to predict genetic interactions. Proc. Natl. Acad. Sci. USA 101, 15682–15687.
[29] Marcotte, E.M., Pellegrini, M., Thompson, M.J., Yeates, T.O. and Eisenberg, D. (1999) A combined algorithm for genome-wide prediction of protein function. Nature 402, 83–86.
[30] Aloy, P., et al. (2004) Structure-based assembly of protein complexes in yeast. Science 303, 2026–2029.
[31] Aloy, P. and Russell, R.B. (2002) Interrogating protein interaction networks through structural biology. Proc. Natl. Acad. Sci. USA 99, 5896–5901.
[32] Lu, L., Arakaki, A.K., Lu, H. and Skolnick, J. (2003) Multimeric threading-based prediction of protein–protein interactions on a genomic scale: application to the *Saccharomyces cerevisiae* proteome. Genome Res. 13, 1146–1154.
[33] Zhang, L.V., Wong, S.L., King, O.D. and Roth, F.P. (2004) Predicting co-complexed protein pairs using genomic and proteomic data integration. BMC Bioinform. 5, 38.
[34] Enright, A.J., Iliopoulos, I., Kyrpides, N.C. and Ouzounis, C.A. (1999) Protein interaction maps for complete genomes based on gene fusion events. Nature 402, 86–90.
[35] Reiss, D.J. and Schwikowski, B. (2004) Predicting protein–peptide interactions via a network-based motif sampler. Bioinformatics 20 (Suppl 1), I274–I282.
[36] Sharon, R., et al. (2005) Conserved patterns of protein interaction in multiple species. Proc. Natl. Acad. Sci. USA 102, 1974–1979.
[37] Walhout, A.J., Sordella, R., Lu, X., Hartley, J.L., Temple, G.F., Brasch, M.A., Thierry-Mieg, N. and Vidal, M. (2000) Protein interaction mapping in *C. elegans* using proteins involved in vulval development. Science 287, 116–122.
[38] Matthews, L.R., Vaglio, P., Reboul, J., Ge, H., Davis, B.P., Garrels, J., Vincent, S. and Vidal, M. (2001) Identification of potential interaction networks using sequence-based searches for conserved protein–protein interactions or "interologs". Genome Res. 11, 2120–2126.
[39] Yu, H., et al. (2004) Annotation transfer between genomes: protein–protein interologs and protein–DNA regulogs. Genome Res. 14, 1107–1118.
[40] Lehner, B. and Fraser, A.G. (2004) A first-draft human protein-interaction map. Genome Biol. 5, R63.

[41] Lee, I., Date, S.V., Adai, A.T. and Marcotte, E.M. (2004) A probabilistic functional network of yeast genes. Science 306, 1555–1558.
[42] Alfarano, C., et al. (2005) The biomolecular interaction network database and related tools 2005 update. Nucleic Acids Res. 33 (Database Issue), D418–D424.
[43] Salwinski, L., Miller, C.S., Smith, A.J., Pettit, F.K., Bowie, J.U. and Eisenberg, D. (2004) The database of interacting proteins: 2004 update. Nucleic Acids Res. 32 (Database issue), D449–D451.
[44] Hermjakob, H., et al. (2004) IntAct: an open source molecular interaction database. Nucleic Acids Res. 32 (Database issue), D452–D455.
[45] Hermjakob, H., et al. (2004) The HUPO PSI's molecular interaction format – a community standard for the representation of protein interaction data. Nat. Biotechnol. 22, 177–183.
[46] Saito, R., Suzuki, H. and Hayashizaki, Y. (2003) Construction of reliable protein–protein interaction networks with a new interaction generality measure. Bioinformatics 19, 756–763.
[47] Goldberg, D.S. and Roth, F.P. (2003) Assessing experimentally derived interactions in a small world. Proc. Natl. Acad. Sci. USA 100, 4372–4376.
[48] Deane, C.M., Salwinski, L., Xenarios, I. and Eisenberg, D. (2002) Protein interactions: two methods for assessment of the reliability of high throughput observations. Mol. Cell Proteomics 1, 349–356.
[49] Giot, L., et al. (2003) A protein interaction map of *Drosophila melanogaster*. Science 302, 1727–1736.
[50] Bader, J.S., Chaudhuri, A., Rothberg, J.M. and Chant, J. (2004) Gaining confidence in high-throughput protein interaction networks. Nat. Biotechnol. 22, 78–85.
[51] Jeong, H., Mason, S.P., Barabasi, A.L. and Oltvai, Z.N. (2001) Lethality and centrality in protein networks. Nature 411, 41–42.
[52] Jeong, H., Tombor, B., Albert, R., Oltvai, Z.N. and Barabasi, A.L. (2000) The large-scale organization of metabolic networks. Nature 407, 651–654.
[53] Strogatz, S.H. (2001) Exploring complex networks. Nature 410, 268–276.
[54] Barabasi, A.L. and Oltvai, Z.N. (2004) Network biology: understanding the cell's functional organization. Nat. Rev. Genet. 5, 101–113.
[55] Said, M.R., Begley, T.J., Oppenheim, A.V., Lauffenburger, D.A. and Samson, L.D. (2004) Global network analysis of phenotypic effects: protein networks and toxicity modulation in *Saccharomyces cerevisiae*. Proc. Natl. Acad. Sci. USA 101, 18006–18011.
[56] Han, J.D., et al. (2004) Evidence for dynamically organized modularity in the yeast protein–protein interaction network. Nature 430, 88–93.
[57] Poyatos, J.F. and Hurst, L.D. (2004) How biologically relevant are interaction-based modules in protein networks. Genome Biol. 5, R93.
[58] Spirin, V. and Mirny, L.A. (2003) Protein complexes and functional modules in molecular networks. Proc. Natl. Acad. Sci. USA 100, 12123–12128.
[59] von Mering, C., Zdobnov, E.M., Tsoka, S., Ciccarelli, F.D., Pereira-Leal, J.B., Ouzounis, C.A. and Bork, P. (2003) Genome evolution reveals biochemical networks and functional modules. Proc. Natl. Acad. Sci. USA 100, 15428–15433.
[60] Shannon, P., et al. (2003) Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res. 13, 2498–2504.
[61] Breitkreutz, B.J., Stark, C. and Tyers, M. (2003) Osprey: a network visualization system. Genome Biol. 4, R22.
[62] Mrowka, R. (2001) A Java applet for visualizing protein–protein interaction. Bioinformatics 17, 669–671.
[63] Uetz, P., et al. (2000) A comprehensive analysis of protein–protein interactions in *Saccharomyces cerevisiae*. Nature 403, 623–627.
[64] Drees, B.L., et al. (2001) A protein interaction map for cell polarity development. J. Cell Biol. 154, 549–571.
[65] Fromont-Racine, M., et al. (2000) Genome-wide protein interaction screens reveal functional networks involving Sm-like proteins. Yeast 17, 95–110.
[66] Gavin, A.C., et al. (2002) Functional organization of the yeast proteome by systematic analysis of protein complexes. Nature 415, 141–147.
[67] Ho, Y., et al. (2002) Systematic identification of protein complexes in *Saccharomyces cerevisiae* by mass spectrometry. Nature 415, 180–183.
[68] Li, S., et al. (2004) A map of the interactome network of the metazoan *C. elegans*. Science 303, 540–543.
[69] Measday, V., Moore, L., Retnakaran, R., Lee, J., Donoviel, M., Neiman, A.M. and Andrews, B. (1997) A family of cyclin-like proteins that interact with the Pho85 cyclin-dependent kinase. Mol. Cell. Biol. 17, 1212–1223.
[70] Carroll, A.S. and O'Shea, E.K. (2002) Pho85 and signaling environmental conditions. Trends Biochem. Sci. 27, 87–93.
[71] Rohde, J.R., Trinh, J. and Sadowski, I. (2000) Multiple signals regulate GAL transcription in yeast. Mol. Cell. Biol. 20, 3880–3886.