Minireview

Interactome modeling

Marc Vidal

Center for Cancer Systems Biology and Department of Cancer Biology, Dana-Farber Cancer Institute, Harvard Medical School, 44 Binney Street, Boston, MA 02115, USA
Department of Genetics, Harvard Medical School, 44 Binney Street, Boston, MA 02115, USA

Received 14 February 2005; accepted 15 February 2005

Available online

Edited by Robert Russell and Giulio Superti-Furga

Abstract A long-term goal of the field of interactome modeling is to understand how global and local properties of complex macromolecular networks impact on observable biological properties, and how changes in such properties can lead to human diseases. The information available at this stage of development of the field provides strong evidence for the existence of such interesting global and local properties, but also demonstrates that many more datasets will be needed to provide accurate models with increasingly predictive capacity. This review focuses on an early attempt at mapping a multicellular interactome network and on the lessons learned from that attempt.

Keywords: ORFeome; Interactome; Network biology; Systems biology; Integrative omics; Reverse two-hybrid system; Interaction-defective alleles

1. Introduction

Macromolecular interactions such as protein–protein, protein–DNA and protein–RNA interactions are crucial for most biological processes. Tens of thousands of proteins and other macromolecules are expressed in a typical cell, mediating perhaps up to hundreds of thousands of physical interactions at any given moment, either to form molecular machines [1] or to participate in various regulatory processes [2]. In this context, the following questions are particularly intriguing. How are protein interactions organized at the scale of the whole cell? Could there be global and/or local principles that organize such complex networks of interactions? If so, how do we start tackling such topological features of macromolecular networks? And importantly, could it be that such organizational principles are disrupted in human diseases?

Since the beginnings of molecular biology, proteins have been studied mostly one or a few at-a-time using biochemistry and genetics. However, it is becoming increasingly clear that proteins perform their function together in complex networks, rather than in isolation. The notion of "interactome", defined as the complete list of physical interactions mediated by all proteins of an organism, reflects a drastic change in the way biologists have recently started asking fundamental questions. Indeed, biological questions are increasingly addressed in the framework of such complex molecular networks.

The completion of the first draft of the human genome has been compared to the discovery expeditions of unknown lands centuries ago. However, these discovery expeditions were able to only "glimpse" a fraction of the complexity of the unknown lands. Indeed, the one-gene/one-protein at-a-time approach of the last thirty years has provided some indication of function for only 5–10% of all predicted proteins so far. When we contemplate the draft of the human genome sequence and its resulting predicted proteome, we are thus facing a gigantic and daunting unknown territory, a sort of "terra incognita" of modern times. What were the opportunities, challenges and goals confronted a decade ago, as the genome and transcriptome sequencing projects were launched?

Below I present a view of early efforts at proteome-wide protein interaction mapping that is biased towards the technological development of high-throughput binary assays for multicellular organisms. As this work was unfolding, protein interaction maps gradually became available for simpler organisms from the mid- to late nineties [3–8].

2. Lessons from an early attempt

In the early 1990s, genome and transcriptome sequencing efforts were beginning to predict large numbers of completely uncharacterized proteins. Particularly, the sequence of the first chromosome of *Saccharomyces cerevisiae* [9] and the first few overlapping cosmids of the *Caenorhabditis elegans* genome [10] predicted many genes that had not been identified by any classical genetic screens. Likewise, the first systematic cDNA sequencing or "expressed sequence tags" (ESTs) projects revealed large numbers of unstudied gene products [11,12].

As a way to decipher the function of these "orphan" proteins and to determine how they work together in complex cellular networks, in the spring of 1993 I undertook the to development of a high-throughput (HT) system to map and characterize large numbers of physical protein–protein interactions, making sure that such a system would be applicable to the study of multicellular organisms, including humans.

At that time, it was clear that the yeast two-hybrid system (Y2H) [13] provided the only hope to ever generate such global protein–protein interaction maps. The challenge consisted in

E-mail address: marc_vidal@dfci.harvard.edu.

Abbreviations: Y2H, yeast two-hybrid system; R2H, reverse two-hybrid system; HT, high-throughput; DB, DNA binding domain; AD, activation domain; ORF, open reading frame; EST, expressed sequence tag; IST, interaction sequence tag; IDA, interaction-defective allele

0014-5793/$30.00 © 2005 Published by Elsevier B.V. on behalf of the Federation of European Biochemical Societies.
doi:10.1016/j.febslet.2005.02.030

setting up a system in which a cDNA library would be fused not only to the activation domain (AD-cDNA) of the Y2H, as originally proposed [13] and demonstrated [14,15], but also to the DNA binding domain (DB-cDNA) of the system. If somehow one could mix and match DB-cDNA and AD-cDNA libraries together, very large numbers of potentially interacting proteins could be identified using automated, high-throughput settings, somewhat in the spirit of the starting EST projects of the time.

A first attempt of a DB-library/AD-library screen with phage lambda genomic DNA [16] demonstrated that, to obtain biologically interpretable protein interaction maps, a HT DB-cDNA/AD-cDNA random clone picking and sequencing system would require the following features. First, a cleaner and more stringent version of the Y2H was needed to reduce the rate of false-positives. To accomplish this, a version of Y2H with both more "physiological" (or else "reasonable") levels of expression of DB-X and AD-Y hybrid proteins and multiple Gal4-responsive reporter genes was developed [17,18], and tested using numerous bait proteins of particular biological interest [19–24]. Second, the possibility that DB-cDNA libraries could be used in the Y2H system needed to be demonstrated [25]. Third, a robust mating strategy had to be developed to mix large numbers, i.e. in the range of 10^5–10^6, of DB-cDNA- and AD-cDNA-containing yeast clones [18,26]. Fourth, a system was needed to eliminate DB-cDNA auto-activators expected to occur in high proportions [18,26,27]. Finally, it was clear that procedures would be needed to allow one, upon finding DB-X/AD-Y Y2H interactions, to isolate genetic reagents, such as interaction-defective alleles, or interaction-dissociating peptides or compounds, to study these interactions back in their natural in vivo environment [17,18,28–32].

After three years of technology development, the system was tested using two mouse cDNA libraries, one fused to DB and the other fused to AD. Thousands of yeast clones, each corresponding to a pair of "interaction sequence tags", or "ISTs", were recovered. At last, an IST database could now be obtained for the mouse, containing hopefully long lists of IST_{ij}, defined as DB-EST_i + AD-EST_j, where ESTs are "expressed sequence tags" and + represents an Y2H interaction. This was the good news.

The bad news, however, was that these initial ISTs overwhelmingly corresponded to a single interaction found in both Y2H orientations: DB-αGlobin + AD-βGlobin and DB-βGlobin + AD-αGlobin [26]. In retrospect, this was an expected result given the relatively high level of abundance of the transcripts of these two proteins in most cDNA libraries.

Beyond the first deception, two important lessons were learned from this experiment. First, to completely avoid such highly abundant ESTs in DB-cDNA and AD-cDNA libraries, it became clear that comprehensive "ORFeome" cloning projects needed to be launched, starting from completely sequenced and well-annotated genomes [33,34]. Second, even though a few additional Y2H interactions were found in the midst of hundreds of DB-αGlobin + AD-βGlobin clones, there was no way to derive their biological implications strictly from IST information. Thus it became clear that "interactome mapping" projects would have to be performed hand-in-hand with other functional genomic and proteomic approaches in order to obtain predictive models of interactome networks. In other words "interactome *modeling*" would require many additional large-scale approaches for biochemical and genetic characterization of the proteome [30,35,36].

3. ORFeome cloning

We selected *C. elegans* as model organism [30] to learn how complex interactome networks relate to metazoan development, because (i) its genome would turn out to be the first to be sequenced for a metazoan [37], and (ii) its cell lineage had been completely mapped [38]. Thus high quality models of the complete set of protein-encoding open reading frames, or "ORFeome", could be used to express and characterize most proteins using multiple approaches to start generating a "proteome atlas" [35] in the context of a nearly perfect "anatomy atlas".

We adapted the Gateway cloning technology [34,39] to attempt the cloning of all 19 000 predicted *C. elegans* ORFs [33,40–42], and more recently of ~10 000 human ORFs [43]. Gateway allows efficiency and adaptability in HT ORFeome cloning projects [44,45], by providing ways to directionally clone PCR products, obtained in our case from a worm cDNA library as template DNA, into a "Donor" vector. This generates a flexible resource to transfer the resulting cloned ORFs into many expression or "Destination" vectors in parallel [46].

4. Interactome mapping

C. elegans protein interaction maps were first attempted at the scale of individual biological processes, starting from all or most proteins known to be involved in these processes, such as vulval development, proteasome, germline, DNA damage response and Dauer formation [33,40,47–50]. The combined data from such "module-scale" interactome mapping attempts suggested among other things a higher level of interconnectivity between pathways than originally expected.

The properties of the *C. elegans* interactome network were then investigated at the scale of the whole proteome [51], focusing first on the subset of predicted worm proteins that have a clear ortholog in other multicellular organisms, but not in the yeast *S. cerevisiae*. From these screens, ~4000 Y2H interactions were identified, representing approximately 5–10% of the *C. elegans* interactome, a dataset referred to as WI5. WI5 is a useful resource to predict the function(s) of thousands of genes. Together with a *Drosophila* interactome mapping dataset [52], this work represented the first attempt to characterize a metazoan interactome.

We have recently demonstrated that no matter how primary Y2H or pull-down/mass-spectrometry screens [53,54] are conducted, the overall quality of an interactome dataset can be improved by systematically incorporating multiple data sets [55] or by retesting the "edges" of a network by different, orthogonal, secondary interaction assays [51]. The flexibility of the Gateway cloning system allows the transfer of thousands of ORFs at-a-time into different vectors that can then be used for secondary binary interaction assays. In our recent interactome map [51], we were able to show that 65% of the Y2H edges retested positive in a single co-affinity pull down assay performed in mammalian cells. In a recent experiment, we

tested 19 worm Y2H interactions, corresponding to orthologs of the human TGFβ pathway, out of which 17 retested positive after only one attempt at the co-affinity pull down assay, corresponding to ~90% retest [50].

While repeating directly either Y2H or pull-down/mass spectrometry assays can confirm initial observations and demonstrate reproducibility of a specific assay system [53], it goes without saying that testing the same interaction with two different assays is much more stringent than repeating the same procedure twice.

5. Comparative interactomics

A protein interaction map should be beneficial not only in the context of the species it was first intended for, but also for biologists involved in the study of other species. To test this idea we investigated to what extent the proteome-scale interactome map generated for yeast [5,7] could help making predictions of interactions for *C. elegans*. For each protein partner pair corresponding to a set of potential yeast interactions, we searched in silico for pairs of respective orthologs, or "interologs" [33], in *C. elegans* [56]. Approximately 20% of such worm potential interologs gave rise to detectable Y2H read-outs. Compared to experiments performed with randomly selected protein pairs tested in various two-hybrid settings, this number represents more than a 2000-fold increase.

We concluded that interologs are reasonable predictors of protein–protein interactions. Altogether, these observations suggested that protein interaction maps generated for a few model organisms might be useful to study, and may be design therapeutic strategies against, a large number of other organisms such as parasites and pathogens. Recent work further exemplified the potential power of such interolog searches [57].

6. Interactome modeling

Attempts have been described to model the function and dynamics of interactome networks by integrating various functional genomic approaches such as expression profiling and genome-wide phenotypic profiling generated by gene knockouts or RNA interference experiments. One should keep in mind that interactome mapping approaches have intrinsic caveats. For example, information is often missing because of the occurrence of false negatives, and information can be misleading because of the presence of false positives. Thus, data obtained from any single interactome mapping approach should be interpreted cautiously. In addition, data emerging from any single Y2H interaction can only indicate the possibility of related functions between two proteins, but does not constitute definitive proof.

It has been proposed that these limitations can be overcome by integrating data obtained from two or more distinct approaches [30,35,36]. For example, a Y2H interaction between two proteins whose genes are co-expressed under various experimental conditions and show overlapping loss-of-function phenotypes is more likely to be relevant in vivo than any interaction for which this additional information is not available. Recent investigations of the relationships between data sets obtained using distinct omic approaches demonstrated the use of such integrated approaches to model the interactome, thereby improving the analysis of biological systems [48,49,58–62].

7. Interaction-defective alleles and reverse two-hybrid system

To fully make use of interactome models it is important to develop HT strategies to validate potential protein–protein interactions back in the biologically relevant settings. A genetic strategy that can be used to validate potential interactions is to identify single amino-acid change that specifically affect one interaction while leaving all other known interactions intact. Such interaction-defective alleles (IDAs) can be tested for their ability to function either in vitro or in vivo [18]. Similarly *trans*-acting dissociators such as peptides or compounds could also be used. Correlation between loss-of-interaction and loss-of-function provides strong evidence of biological relevance for a potential interaction [33].

Integrated approaches are available for HT selection and manipulation of IDAs without the need for any structural information on the proteins involved [17,28,32]. Such approaches are based on a modified version of the reverse two-hybrid system (R2H). Of particular interest, the use of the Green Fluorescent Protein [63] as a C-terminal tag allows the recovery of single amino-acid substitutions that specifically prevent interaction rather than non-sense mutations that encode truncated proteins. In addition, the Gateway recombinational cloning technique can be used to rapidly transfer IDAs from the yeast assay into different expression vectors allowing subsequent characterization. This integrated version of the R2H is amenable to automation, which is important considering the large numbers of potential interactions already available.

8. Conclusions

Altogether the work performed in both unicellular and multicellular organisms has shown that a systematic approach to the challenge of globally mapping interactome networks is possible and can be highly informative. At this stage, the technology should be sufficiently mature to start analyzing the human interactome network at the proteome-scale. Among future goals, we can now focus on studying the evolution of interactome networks, by comparing those of yeast, *C. elegans*, *Drosophila* and humans, and understanding cellular organizational principles of the human interactome. An interesting approach would be to study the global effects of viral proteomes upon infection into their host cells. Is it possible that evolution has shaped global strategies employed by viral proteomes to rewire the host's cellular networks and by doing so, forcing the host to reorganize its cellular activities?

Among the major challenges of the field of interactomics is the fact that the proteome is a dynamic entity. In terms of defining the human genome nucleotide sequence, the Human Genome Project was a finite enterprise. In contrast, it is harder to define the ultimate goal of a "human interactome project". Indeed, that there are as many "sub"-proteomes in the human body as there are cells and conditions. In other words, the proteome is constantly changing through time and space. Future

versions of interactome maps will have to take this dimension into account.

Acknowledgements: Many thanks to our colleagues and collaborators who made this work possible over the years, particularly to M. Walhout, J. Reboul, P. Vaglio, S. Boulton, E. Harlow, J. Boeke, M. Cusick and D. Hill. The work performed in this laboratory is funded by grants from NHGRI, NIGMS and NCI. Special thanks to M. Cusick and D. Hill for their help with this manuscript. Apologies to those whose work could not be included in this mini-review because of space limitations.

References

[1] Alberts, B. (1998) The cell as a collection of protein machines: Preparing the next generation of molecular biologists. Cell 92, 291–294.
[2] Pawson, T. and Nash, P. (2003) Assembly of cell regulatory systems through protein interaction domains. Science 300, 445–452.
[3] 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.
[4] Fromont-Racine, M., Rain, J.C. and Legrain, P. (1997) Toward a functional analysis of the yeast genome through exhaustive two-hybrid screens. Nat. Genet. 16, 277–282.
[5] Uetz, P., et al. (2000) A comprehensive analysis of protein–protein interactions in *Saccharomyces cerevisiae*. Nature 403, 623–627.
[6] Flajolet, M., Rotondo, G., Daviet, L., Bergametti, F., Inchauspe, G., Tiollais, P., Transy, C. and Legrain, P. (2000) A genomic approach of the hepatitis C virus generates a protein interaction map. Gene 242, 369–379.
[7] 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.
[8] Rain, J.C., et al. (2001) The protein–protein interaction map of *Helicobacter pylori*. Nature 409, 211–215.
[9] Oliver, S.G., et al. (1992) The complete DNA sequence of yeast chromosome III. Nature 357, 38–46.
[10] Sulston, J., et al. (1992) The *C. elegans* genome sequencing project: A beginning. Nature 356, 37–41.
[11] McCombie, W.R., et al. (1992) Caenorhabditis elegans expressed sequence tags identify gene families and potential disease gene homologues. Nat. Genet. 1, 124–131.
[12] Adams, M.D., et al. (1992) Sequence identification of 2375 human brain genes. Nature 355, 632–634.
[13] Fields, S. and Song, O. (1989) A novel genetic system to detect protein–protein interactions. Nature 340, 245–246.
[14] Chevray, P.M. and Nathans, D. (1992) Protein interaction cloning in yeast: Identification of mammalian proteins that react with the leucine zipper of Jun. Proc. Natl. Acad. Sci. USA 89, 5789–5793.
[15] Dalton, S. and Treisman, R. (1992) Characterization of SAP-1, a protein recruited by serum response factor to the c-fos serum response element. Cell 68, 597–612.
[16] Vidal, M., Boeke, J.D., and Harlow, E. (1999) Reverse Two-Hybrid System. US Patent 5955280.
[17] Vidal, M., Brachmann, R.K., Fattaey, A., Harlow, E. and Boeke, J.D. (1996) Reverse two-hybrid and one-hybrid systems to detect dissociation of protein–protein and DNA-protein interactions. Proc. Natl. Acad. Sci. USA 93, 10315–10320.
[18] Vidal, M. (1997) The reverse two-hybrid system in: The Yeast Two-hybrid System (Bartels, P. and Fields, S., Eds.), pp. 109–147, Oxford University Press, New York.
[19] Sardet, C., Vidal, M., Cobrinik, D., Geng, Y., Onufryk, C., Chen, A. and Weinberg, R.A. (1995) E2F-4 and E2F-5, two members of the E2F family, are expressed in the early phases of the cell cycle. Proc. Natl. Acad. Sci. USA 92, 2403–2407.
[20] Kim, S.S., Chen, Y.M., O'Leary, E., Witzgall, R., Vidal, M. and Bonventre, J.V. (1996) A novel member of the RING finger family, KRIP-1, associates with the KRAB-A transcriptional repressor domain of zinc finger proteins. Proc. Natl. Acad. Sci. USA 93, 15299–15304.
[21] Hu, G., Zhang, S., Vidal, M., Baer, J.L., Xu, T. and Fearon, E.R. (1997) Mammalian homologs of seven in absentia regulate DCC via the ubiquitin-proteasome pathway. Genes Dev. 11, 2701–2714.
[22] Ronco, L.V., Karpova, A.Y., Vidal, M. and Howley, P.M. (1998) Human papillomavirus 16 E6 oncoprotein binds to interferon regulatory factor-3 and inhibits its transcriptional activity. Genes Dev. 12, 2061–2072.
[23] Cenciarelli, C., Chiaur, D.S., Guardavaccaro, D., Parks, W., Vidal, M. and Pagano, M. (1999) Identification of a family of human F-box proteins. Curr. Biol. 9, 1177–1179.
[24] Zukerberg, L.R., et al. (2000) Cables links Cdk5 and c-Abl and facilitates Cdk5 tyrosine phosphorylation, kinase upregulation, and neurite outgrowth. Neuron 26, 633–646.
[25] Du, W., Vidal, M., Xie, J.E. and Dyson, N. (1996) RBF, a novel RB-related gene that regulates E2F activity and interacts with cyclin E in Drosophila. Genes Dev. 10, 1206–1218.
[26] Walhout, A.J. and Vidal, M. (1999) A genetic strategy to eliminate self-activator baits prior to high- throughput yeast two-hybrid screens. Genome Res. 9, 1128–1134.
[27] Vidalain, P.O., Boxem, M., Ge, H., Li, S. and Vidal, M. (2004) Increasing specificity in high-throughput yeast two-hybrid experiments. Methods 32, 363–370.
[28] Vidal, M., Braun, P., Chen, E., Boeke, J.D. and Harlow, E. (1996) Genetic characterization of a mammalian protein–protein interaction domain by using a yeast reverse two-hybrid system. Proc. Natl. Acad. Sci. USA 93, 10321–10326.
[29] Yasugi, T., Vidal, M., Sakai, H., Howley, P.M. and Benson, J.D. (1997) Two classes of human papillomavirus type 16 E1 mutants suggest pleiotropic conformational constraints affecting E1 multimerization, E2 interaction, and interaction with cellular proteins. J. Virol. 71, 5942–5951.
[30] Walhout, M., Endoh, H., Thierry-Mieg, N., Wong, W. and Vidal, M. (1998) A model of elegance. Am. J. Hum. Genet. 63, 955–961.
[31] Vidal, M. and Endoh, H. (1999) Prospects for drug screening using the reverse two-hybrid system. Trends Biotechnol. 17, 374–381.
[32] Endoh, H., Vincent, S., Jacob, Y., Real, E., Walhout, A.J. and Vidal, M. (2002) Integrated version of reverse two-hybrid system for the postproteomic era. Methods Enzymol. 350, 525–545.
[33] 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.
[34] Walhout, A.J., Temple, G.F., Brasch, M.A., Hartley, J.L., Lorson, M.A., van den Heuvel, S. and Vidal, M. (2000) GATEWAY recombinational cloning: Application to the cloning of large numbers of open reading frames or ORFeomes. Methods Enzymol. 328, 575–592.
[35] Vidal, M. (2001) A biological atlas of functional maps. Cell 104, 333–339.
[36] Ge, H., Walhout, A.J. and Vidal, M. (2003) Integrating 'omic' information: A bridge between genomics and systems biology. Trends Genet. 19, 551–560.
[37] The C. elegans Sequencing Consortium (1998) Genome sequence of the nematode C. elegans: A platform for investigating biology. Science 282, 2012–2018.
[38] Sulston, J.E., Schierenberg, E., White, J.G. and Thomson, J.N. (1983) The embryonic cell lineage of the nematode *Caenorhabditis elegans*. Dev. Biol. 100, 64–119.
[39] Hartley, J.L., Temple, G.F. and Brasch, M.A. (2000) DNA cloning using in vitro site-specific recombination. Genome Res. 10, 1788–1795.
[40] Reboul, J., et al. (2003) *C. elegans* ORFeome version 1.1: Experimental verification of the genome annotation and resource for proteome-scale protein expression. Nat. Genet. 34, 35–41.
[41] Reboul, J., et al. (2001) Open-reading-frame sequence tags (OSTs) support the existence of at least 17,300 genes in *C. elegans*. Nat. Genet. 27, 332–336.
[42] Lamesch, P., et al. (2004) *C. elegans* ORFeome version 3.1: Increasing the coverage of ORFeome resources with improved gene predictions. Genome Res. 14, 2064–2069.

[43] Rual, J.F., et al. (2004) Human ORFeome version 1.1: A platform for reverse proteomics. Genome Res. 14, 2128–2135.
[44] Rual, J.F., Hill, D.E. and Vidal, M. (2004) ORFeome projects: Gateway between genomics and omics. Curr. Opin. Chem. Biol. 8, 20–25.
[45] Brasch, M.A., Hartley, J.L. and Vidal, M. (2004) ORFeome cloning and systems biology: Standardized mass production of the parts from the parts-list. Genome Res. 14, 2001–2009.
[46] Rual, J.F., et al. (2004) Toward improving *Caenorhabditis elegans* phenome mapping with an ORFeome-based RNAi library. Genome Res. 14, 2162–2168.
[47] Davy, A., et al. (2001) A protein–protein interaction map of the *Caenorhabditis elegans* 26S proteasome. EMBO Rep. 2, 821–828.
[48] Walhout, A.J., et al. (2002) Integrating interactome, phenome, and transcriptome mapping data for the *C. elegans* germline. Curr. Biol. 12, 1952–1958.
[49] Boulton, S.J., Gartner, A., Reboul, J., Vaglio, P., Dyson, N., Hill, D.E. and Vidal, M. (2002) Combined functional genomic maps of the *C. elegans* DNA damage response. Science 295, 127–131.
[50] 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.
[51] Li, S., et al. (2004) A map of the interactome network of the metazoan *C. elegans*. Science 303, 540–543.
[52] Giot, L., et al. (2003) A protein interaction map of *Drosophila melanogaster*. Science 302, 1727–1736.
[53] Gavin, A.C., et al. (2002) Functional organization of the yeast proteome by systematic analysis of protein complexes. Nature 415, 141–147.
[54] Ho, Y., et al. (2002) Systematic identification of protein complexes in *Saccharomyces cerevisiae* by mass spectrometry. Nature 415, 180–183.
[55] Han, J.D., et al. (2004) Evidence for dynamically organized modularity in the yeast protein–protein interaction network. Nature 430, 88–93.
[56] 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.
[57] Sharan, R., et al. (2005) From the Cover: Conserved patterns of protein interaction in multiple species. Proc. Natl. Acad. Sci. USA 102, 1974–1979.
[58] Ge, H., Liu, Z., Church, G.M. and Vidal, M. (2001) Correlation between transcriptome and interactome mapping data from *Saccharomyces cerevisiae*. Nat. Genet. 29, 482–486.
[59] Grigoriev, A. (2001) A relationship between gene expression and protein interactions on the proteome scale: Analysis of the bacteriophage T7 and the yeast *Saccharomyces cerevisiae*. Nucleic Acids Res. 29, 3513–3519.
[60] Kemmeren, P., van Berkum, N.L., Vilo, J., Bijma, T., Donders, R., Brazma, A. and Holstege, F.C. (2002) Protein interaction verification and functional annotation by integrated analysis of genome-scale data. Mol. Cell 9, 1133–1143.
[61] Jansen, R., Greenbaum, D. and Gerstein, M. (2002) Relating whole-genome expression data with protein–protein interactions. Genome Res. 12, 37–46.
[62] Begley, T.J., Rosenbach, A.S., Ideker, T. and Samson, L.D. (2002) Damage recovery pathways in *Saccharomyces cerevisiae* revealed by genomic phenotyping and interactome mapping. Mol. Cancer Res. 1, 103–112.
[63] Chalfie, M., Tu, Y., Euskirchen, G., Ward, W.W. and Prasher, D.C. (1994) Green fluorescent protein as a marker for gene expression. Science 263, 802–805.