

2

Yeast Two-Hybrid Protein–Protein Interaction Networks

Daniel Auerbach and Igor Stagljar

ABSTRACT

The availability of complete genome sequences of numerous model organisms has initiated the development of new approaches in biological research to complement conventional biochemistry and genetics. Consequently, high-throughput methodologies also need to be applied in the emerging field of proteomics. Here, we discuss several methods that have been developed in the past years in order to characterize proteins and their functions on a large scale. We focus on the yeast two-hybrid system, which is the most widely used method to study protein–protein interactions and which has been used several times now to successfully map entire interaction networks on a large scale. We discuss small-scale pilot projects and how they have been upscaled to genome-wide screens, such as for the budding yeast *Saccharomyces cerevisiae*. We then compare the yeast two-hybrid system with several other screening methods that have been developed to investigate interactions between proteins in a high-throughput format, such as affinity purification methods coupled to mass spectrometry. Efficient adaptation of such methods to a high-throughput format, coupled with the increasing use of databases to compare interaction maps generated with different methods, will help in elucidating protein–protein interactions on a scale that would have been unthinkable just a few years ago.

DANIEL AUERBACH • Dualsystems Biotech Inc., Winterthurerstrasse 190, CH-8057 Zurich, Switzerland.

IGOR STAGLJAR • Institute of Veterinary Biochemistry and Molecular Biology, University of Zurich, Winterthurerstrasse 190, CH-8057 Zurich, Switzerland.

Proteomics and Protein–Protein Interactions: Biology, Chemistry, Bioinformatics, and Drug Design, edited by Waksman. Springer, New York, 2005.

1. INTRODUCTION

The availability of complete genome sequences of numerous model organisms has initiated the development of new approaches in biological research to complement conventional biochemistry and genetics. For example, only one third of all 6200 predicted yeast genes had been functionally characterized when the complete sequence of the yeast genome first became available (Goffeau et al., 1996). At present, only 3800 yeast genes have been characterized by genetic or biochemical means and there still remain approximately 1800 genes encoding proteins of unknown function (Kumar and Snyder, 2001). The same observation holds true for the human genome: approximately 80% of all predicted human genes have not been characterized to date (Aach et al., 2001). To answer this challenge, researchers have developed different high-throughput strategies to characterize unknown genes on a large scale.

To date, most interaction maps have been created by genetic screening in yeast, namely by using the yeast two-hybrid system (Fields and Song, 1989). The reasons for the success of the yeast two-hybrid system in large-scale screening projects are manifold: as an *in vivo* genetic screening system, it is easily scalable, no purification steps or optimizations with regard to binding or washing conditions are involved, and automatization using robotic platforms is very easy. On the other hand, false positives and false negatives remain a problem of the yeast two-hybrid system; consequently, large-scale interaction maps derived by such methods require stringent selection criteria to yield useful information. Below, we first discuss the protein–protein interaction maps from various organisms that have been created using the yeast two-hybrid system and then discuss the advantages and disadvantages of this method. Finally, we briefly describe what has been done in analyzing those interaction maps to date.

2. THE YEAST TWO-HYBRID SYSTEM

The yeast two-hybrid system originally created by Fields and Song is a genetic system wherein the interaction between two proteins of interest is detected via the reconstitution of a transcription factor and the subsequent activation of reporter genes under the control of this transcription factor (Fields and Song, 1989). As depicted in Figure 2.1A, a protein X is expressed as a fusion to a DNA binding domain (DBD). The DBD–X fusion is commonly termed the “bait.” Because of the affinity of the DBD for its operator sequences the bait is bound to a promoter element upstream of a reporter gene but does not activate it because it lacks an activation domain. A second protein Y is expressed as a fusion to an activation domain (AD) and is commonly termed the “prey.” The prey is capable of activating transcription but usually does not do so because it has no affinity for the promoter elements upstream of the reporter gene (Fig. 2.1B). If bait and prey are coexpressed and the two proteins X and Y interact, then a functional transcription factor is reconstituted at the promoter site

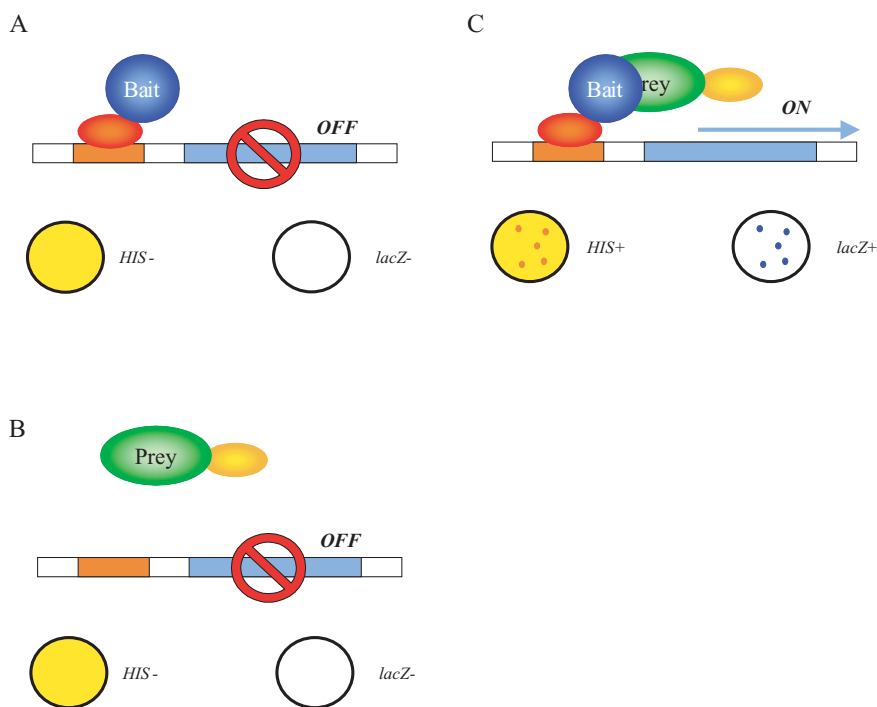


Figure 2.1. The yeast two-hybrid system. (A) A bait is expressed as a fusion to a DNA binding domain (DBD), for example, the *Saccharomyces cerevisiae* GAL4 transcription factor or the *Escherichia coli* LexA protein. The DBD–bait binds to the operator sequences present in the promoter region upstream of the reporter gene but does not activate its transcription since the DBD–bait does not contain an activation domain. (B) A prey is expressed as a fusion to an activation domain (AD), for example, from the GAL4 transcription factor or from the Herpes simplex virus protein VP16. The AD–prey fusion has the capability to activate transcription in yeast but because it is not actively targeted to the promoter it does not activate transcription of the reporter gene. (C) The interaction between bait and prey targets the AD–prey fusion protein to the promoter, thereby reconstituting an active transcription factor. The hybrid transcription factor is bound to the promoter upstream of the reporter gene and therefore activates transcription. The readout of the activated reporter gene is measured either as growth on selective medium (auxotrophic selection markers, such as *HIS3*, *URA3*, or *ADE2*) or in a color reaction (*lacZ*). Yeast expressing only the DBD–bait or the AD–prey on its own do not grow on selective medium (*HIS*[−]) and do not display blue staining in a color assay (*lacZ*[−]), whereas yeast harboring an interacting DBD–bait and AD–prey display growth (*HIS*⁺) and blue color (*lacZ*⁺).

upstream of the reporter gene. Consequently, transcription of the reporter gene is activated. Thus, in a yeast two-hybrid assay a protein–protein interaction is measured through the activation of one or several reporter genes in response to the assembly of a transcription factor by the said protein–protein interaction (Fig. 2.1C). In common yeast two-hybrid screening schemes the prey is usually replaced by a collection of unknown preys expressed from a cDNA or genomic library. Screening of entire libraries against a defined bait may then lead to the discovery of novel interaction

partners. For large-scale screenings, two approaches are commonly used: the library screening approach, in which multiple baits are screened against a library, and the matrix approach, in which an array of defined preys is substituted for the library.

3. LARGE-SCALE SCREENS USING THE LIBRARY APPROACH

The library approach is schematically shown in Figure 2.2A. A particular bait is expressed in a yeast reporter strain of the mating type a , whereas a collection of preys (the library) is transformed into a yeast reporter strain of the mating type α . The bait-bearing strain is then mated with the mixture of library strains, and clones expressing an interaction pair are isolated on selective media. To determine the identity of the interacting prey, the library plasmid encoding it has to be isolated from the yeast strain and amplified in *Escherichia coli*. The region encoding the prey is then sequenced.

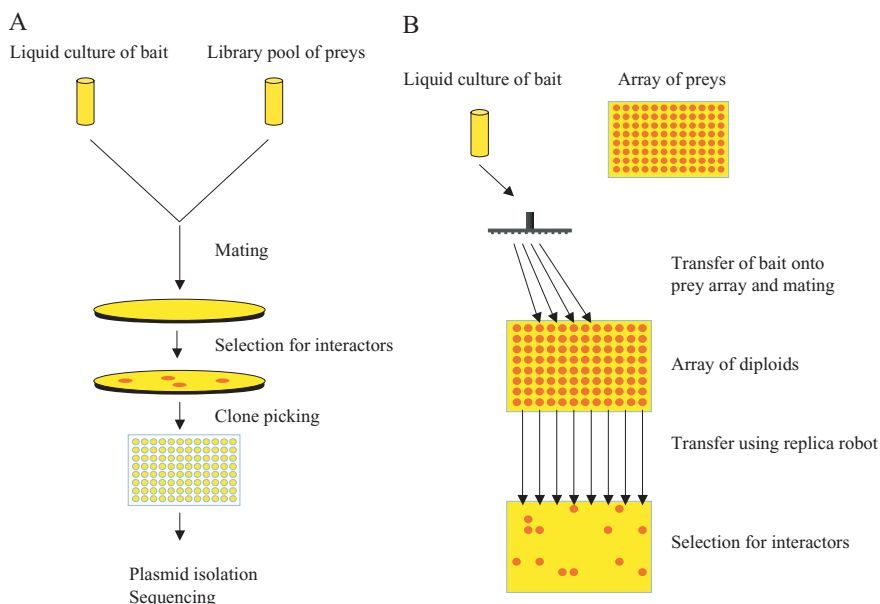


Figure 2.2. High-throughput approaches utilizing the yeast two-hybrid system. **(A)** The library screening approach. A yeast strain expressing a bait under investigation is mixed with a collection of yeast strains each expressing a random prey from a library. Incubation in rich medium allows the two strains to mate and diploids expressing bait and prey are selected. The diploids are then transferred to selective medium to isolate those clones containing interacting baits and preys (selection for interactors). Yeast clones that display growth on selective medium are picked up, transferred into multiwell plates, and processed for plasmid isolation and insert sequencing to identify the interacting prey. **(B)** The matrix or array approach. An array of preys is prepared by spotting yeast clones each expressing a known prey onto plates. The colonies on the array are then picked up by a robot and mated with a yeast strain expressing the bait under investigation. An exact replica of the array is transferred to a fresh plate to select for diploids expressing bait and prey and then to selective medium to select for interacting baits and preys. The identity of the prey in colonies that grow under selection is determined by its position within the array.

Although very powerful when used to investigate a single bait, the library approach is technically challenging when screening a large collection of baits: every yeast clone from the primary selection that may contain a valid protein–protein interaction has to be picked up, the library plasmid isolated, and the identity of the interacting protein established by sequencing of the cDNA insert. Even when using robotic equipment the screening of several thousands of baits against a library still presents a formidable challenge. Furthermore, false-positive interactors, for example, preys that activate the reporter genes without actually binding to the bait are common in library screens and are difficult to identify and eliminate. Several recent improvements in vectors and screening strains have led to a considerably lowered rate of false positives (James et al., 1996; Walhout et al., 2000b; Ito et al., 2001). For instance, when using only one auxotrophic selection marker, 10% to 70% of all clones in a screen may represent false positives. When using two auxotrophic selection markers, either in a simultaneous or in a sequential selection scheme, false positives are virtually eliminated from the screen (Fang and Macool, 2002).

4. LARGE-SCALE SCREENING USING THE MATRIX APPROACH

In the matrix approach a collection of defined preys is used instead of a random collection of open reading frames (ORFs) or ORF fragments. Each prey is separately introduced into yeast and the transformants are arrayed on plates using a robot (Figure 2.2B). A bait-bearing strain of the opposite mating type is then mated with every prey-bearing strain and the resulting diploid strains are replicated onto selective medium. If a particular diploid within the array grows on selection medium, its prey must interact with the bait under investigation. As opposed to the library screen, no plasmid isolation or sequencing is necessary since the position of the growing diploid on the array identifies the prey it expresses. In essence, a matrix screen consists of a series of defined interactions between a bait and a number of preys, rather than a screen of a bait against a collection of unknown preys. An advantage of the matrix screen is that repeated screenings will identify false-positive interactors present in the array. If a particular prey interacts with every bait being tested, chances are high that the interaction is either nonspecific or that the prey activates the reporter genes in the absence of a real protein–protein interactions. Such preys become apparent after several screens and can then be discarded.

5. ADVANTAGES AND DISADVANTAGES OF THE TWO APPROACHES

Both the library screen and the matrix screen have advantages and disadvantages, and ultimately they complement each other. This is underscored by the fact that most large-scale screening projects have been carried out using both approaches (Flajolet et al., 2000; Uetz et al., 2000; Walhout et al., 2000b; Ito et al., 2001). A screen using a

cDNA or genomic library will find not only full-length interactors but also interacting subdomains of a given protein. This may be an advantage where a domain is capable of interacting with a bait but the entire protein is not, for instance, because other parts of the protein inhibit the interaction. It has been argued that library screens are more sensitive than matrix approaches (Ito et al., 2001). On the other hand, they often yield small protein fragments that interact nonspecifically with many unrelated baits, so-called false positives. In conventional yeast two-hybrid screens, in which only one protein is used as a bait, false positives are commonly eliminated by testing each prey against a set of bait proteins. Preys that interact nonspecifically with all baits tested are labeled as false positives and are discarded from the screen. However, when testing thousands of different baits in the context of a large-scale screening project, no such test can be carried out because the actual number of interactions to be assayed would be too high. For this reason, large-scale library screens probably yield a high number of interactions that do not occur in a physiological context and that must therefore be labeled as false positives. Mrowka et al. (2001) have compared results from the two genome-wide screens of *Saccharomyces cerevisiae* with published interaction data from the research community and have concluded that unless there is a severe bias in published interactions, the genome-wide datasets contain an estimated rate of false positives between 44% and 91%.

Many proteins are modular in nature and contain domains with specific functions, such as protein–protein interaction motifs. Often it is important to know which particular part of a prey interacts with the protein of interest. Since libraries contain random fragments of proteins, a library screen usually gives information on the minimal protein fragment necessary for interaction, provided a sufficiently large number of preys are rescued from the screen. A recent study has undertaken a genome-wide screen of the prokaryote *Helicobacter pylori*. A total of 261 baits were screened against a complex *H. pylori* library of small genomic fragments. More than 1200 interactions were identified, resulting in a protein interaction network that connected close to 50% of the entire proteome (Rain et al., 2001). The use of a small fragment library allowed the alignment of preys that were identified in each screen and the consequent identification of minimal interacting domains from these alignments. Such strategies may prove useful for the identification of as yet unknown interaction motifs in genome-wide screens. Matrix screens do not yield this type of information because they commonly use full-length proteins. Using computer algorithms in combination with high-throughput cloning strategies to create arrays containing mixtures of full-length proteins and protein fragments may be envisaged, however.

When compared to library screens the matrix approaches have an important advantage: false positives can be easily identified because every prey has a defined position within the array. Preys that interact nonspecifically with multiple baits can be identified easily and removed from the dataset. Furthermore, the matrix approach works well with sequenced genomes, in which every ORF has been predicted by computational methods and can be cloned easily using high-throughput methods (Uetz et al., 2000; Ito et al., 2001). For genomes of higher eukaryotes such as human or mouse, however, the situation is more difficult: genome annotation is less reliable,

and because of the presence of genomic introns, ORFs cannot simply be cloned using high-throughput genomic polymerase chain reaction (PCR). For this reason, library screens will probably remain the method of choice for constructing interaction maps of higher eukaryotes for several years to come.

Interestingly, screens using the library approach tend to yield a higher number of interactions than those carried out with the matrix approach. The reasons for this are probably manifold. One explanation is that in matrix screens many ORFs have to be discarded for reasons inherent to the proteins under investigation. For instance, interactions in the yeast two-hybrid system have to take place in the nucleus. Consequently, proteins that possess hydrophobic transmembrane domains will be unable to reach the nucleus. This problem is partially circumvented in library screens: libraries that express protein fragments have a greater chance of containing prey fragments that lack “difficult” protein regions, such as transmembrane domains. Consequently, they may detect interactions that are lost when using arrays of full-length proteins. On the other hand, library screens have the aforementioned problem of reproducibility. Whereas any interaction in a matrix screen can be repeated multiple times, library screens are essentially irreproducible. If the resulting interactions are not verified, for example, using multiple control baits, a dataset derived from a library screen may contain a very high number of false-positive interactors that would not occur in a physiological setting (Mrowka et al., 2001).

In the following, the large-scale studies that have been carried out to date are discussed in detail. They range from small-scale interaction maps, which draw protein interaction networks of protein complexes such as the proteasome, to truly genome-wide screens that manage to interconnect a large fraction of the genome of the organism under investigation. The results of each study are summarized in Table 2.1.

6. SMALL-SCALE PROTEIN INTERACTION MAPS

Initially, large-scale yeast two-hybrid screens were used to generate small interactions networks that focused on protein complexes or components of a defined signaling pathway. In 1994, Finley and Brent generated a protein interaction map of cyclin-dependent kinase interacting proteins in *Drosophila melanogaster*. They identified 19 interactions and demonstrated that each cyclin-dependent kinase interacting protein associates with a specific spectrum of Cdks (Finley and Brent, 1994). In another example, a large-scale screen was used to identify interactions between several yeast proteins involved in mRNA splicing. The investigators carried out library screens with 15 defined baits and identified 170 interactions (Fromont-Racine et al., 1997). The screen identified nine preys that encoded known pre-mRNA splicing factors, as well as five preys that were homologous to human splicing factors.

In 2000, Walhout et al. created the first partial protein interaction map for a multicellular organism, the nematode *Caenorhabditis elegans*. Focusing on proteins involved in vulval development, they screened a cDNA library with a total of 27 baits and identified 148 interactions, including 15 previously known interactions and 109

Table 2.1. Large-scale screening projects carried out using the yeast two-hybrid system

Organism	Predicted ORFs	Methods	AD hybrids	DBD hybrids	Interactions	References
Bacteriophage T7	55	Matrix screen	11 ORFs	34 ORFs	3	Bartel et al., 1996
		Library screen	Library	34 ORFs		
			11 ORFs	Library	22	
Vaccinia virus	266	Matrix screen	Library	Library		McCraith et al., 2000
		Matrix screen	266 ORFs	266 ORFs	37	
		Library screen	11 ORFs	10 ORFs	0	
Hepatitis C virus	~10	Library screen	Library	200 ORFs	15	Flajolet, et al., 2000
		Library screen	Library	261 ORFs	1280	
		Library screen	Library	27 ORFs	148	
<i>Helicobacter pylori</i>	1590	Matrix screen	29 ORFs	29 ORFs	11	Rain et al., 2001
		Matrix screen	30 ORFs	30 ORFs	17	
		Library screen	Library	30 ORFs	138	
<i>Caenorhabditis elegans</i>	19,099	Matrix screen	5 ORFs	9 ORFs	19	Davy et al., 2001
		Library screen	Library	15 ORFs	170	
		Library screen	159 ORFs	159 ORFs	183	
<i>Drosophila melanogaster</i>	13,600	Matrix screen	192 ORFs	5345 ORFs	281	Finley et al., 1994
		Library screen	5345 ORFs	5345 ORFs	692	
		Library screen	Library screen	Library screen		
<i>Saccharomyces cerevisiae</i>	6200	Matrix screen	159 ORFs	159 ORFs	183	Fromont-Racine et al., 1997
		Library screen	192 ORFs	5345 ORFs	281	
		Library screen	5345 ORFs	5345 ORFs	692	

interactions that had been predicted based on the *C. elegans* genome sequence (Walhout et al., 2000b). Another group has focused on proteins of the 26S proteasome in *C. elegans*. A matrix screen using 30 baits and preys yielded 17 interactions, whereas library screens with the 30 baits resulted in 138 interactions (Davy et al., 2001).

7. COMPREHENSIVE PROTEIN INTERACTION MAPS

Not surprisingly, the first genome-wide interaction maps were created using bacteriophages and viruses as model systems. Their small, well-characterized genomes make a comprehensive screening study relatively straightforward. The first genome-wide yeast two-hybrid study was carried using bacteriophage T7 as a model system, which has a total of 55 predicted ORFs (Dunn and Studier, 1983). Using a combination of matrix and library screening approaches, random bait and prey fragments were screened against each other and 25 interactions were identified in this way (Bartel et al., 1996).

In the subsequent study, McCraith et al. applied the matrix approach to examine interactions between 266 ORFs of the vaccinia virus as bait and prey fusions (McCraith et al., 2000), resulting in a total of 70,000 combinations of DBD and AD fusions. This study identified 37 protein–protein interactions, including 28 that had previously been identified. The same matrix approach has also been applied to build a protein interaction map of 200 ORFs of the hepatitis C virus (Flajolet et al., 2000). Interestingly, no interactions were identified using this approach, most probably because of incorrect folding or mistargeting of full-length DBD and AD fusions. To circumvent this problem, the authors applied the exhaustive library screening approach in which the above-mentioned 200 DBD–ORF fusions were screened against a random genomic library. This approach yielded 15 interactions that included both previously known and novel interacting pairs.

Similarly to bacteriophages and viruses, bacterial genomes, because of their small number of protein coding genes, also represent ideal model organisms for generation of protein interaction maps. However, only one systematic yeast two-hybrid approach has so far been undertaken in bacteria to analyze protein interactions at a global level—the one of the human gastric bacterial pathogen *H. pylori*, whose genome encodes 1590 predicted ORFs. Using the matrix approach and exhaustive genomic library screening on 261 ORFs fused to DBD, Rain et al. identified a total of 1280 interactions, resulting in a protein interaction map covering much of the *H. pylori* proteome (Rain et al., 2001). This study indicated that building a protein interaction map of a pathogenic bacterium may represent a powerful new tool for understanding the molecular mechanisms of infection and drug resistance and for developing novel innovative therapies.

The most comprehensive yeast two-hybrid screenings to date focus on the yeast *S. cerevisiae*. In 2000, two groups completed the comprehensive yeast two-hybrid mappings on all 6000 yeast ORFs as baits using both matrix and library screening approaches (Ito et al., 2000, 2001; Uetz et al., 2000). In an effort led by Ito et al.,

large-scale matrix screen using 159 ORFs cloned as DBD and AD fusions were performed and resulted in identification of 175 interactions, of which 163 had not been reported previously (Ito et al., 2000). Recently, the same group completed their exhaustive yeast two-hybrid screenings on yeast ORFs and identified 841 interactions in total (Ito et al., 2001). The second comprehensive study performed by the Fields group utilized both array and library screenings (Uetz et al., 2000). Using the array method, 192 ORFs were created as DBD fusions and then mated with the 6000 ORFs of yeast fused to the AD. Only 20% of all interactions were found in both screens, resulting in 281 protein pairs. For the exhaustive library screen, a library was made by pooling the 5345 AD-fused ORFs. These were then mated separately to the same 5345 ORFs fused to the DBD, yielding a total of 692 protein–protein interactions.

8. COMPARISON OF INTERACTION MAPS CREATED BY DIFFERENT SCREENING METHODS

When comparing the datasets of Ito et al. and Uetz et al. it is interesting to note that despite the fact that both groups used the same 6000 ORFs in their experiments, only 20% of all interactions in the two datasets actually overlap (Ito et al., 2001). The reasons for this small overlap are difficult to explain. A significant factor may have been the use of different experimental systems: bait and prey plasmids used in the two studies differ with regard to copy number within the cell, selection markers, and promoters driving expression of bait and prey proteins. The use of PCR to amplify the yeast ORFs may have introduced mutations that abolish interactions and, most importantly, the stringency of selection may have been different, eliminating interactions seen by one group from the other group's dataset (see discussion in Ito et al., 2001). The small overlap can be taken to mean that even when using exhaustive library screens that potentially cover all interactions in a genome, the subset of protein–protein interactions that can be identified using the yeast two-hybrid system is far from representative.

How do large-scale interaction maps therefore compare with the interaction data gathered by the research community in the past decades? A recent publication has compared the data from several large-scale studies on yeast with data available from public protein interaction databases (von Mering et al., 2002). They also compared yeast two-hybrid screenings with data from several other high-throughput methods, such as affinity purification coupled to mass spectrometry (Gavin et al., 2002; Ho et al., 2002). Their findings highlight again the problem of small overlap between different interaction datasets: despite the fact that high-throughput methods in yeast have generated some 80,000 interactions to date, only a small fraction (2400 interactions) are supported by more than one method. Thus, every method probably has a bias toward certain protein–protein interactions and may fail to detect others. As already mentioned, any screen based on the yeast two-hybrid system will have difficulties in detecting interactions between integral membrane proteins or membrane-associated proteins since these proteins are unable to reach the nucleus. Consequently,

membrane proteins are underrepresented in yeast two-hybrid datasets (von Mering et al., 2002). On the other hand, methods based on copurification of protein complexes often miss signaling pathways and transport pathways. In this respect it is also important to note that the yeast two-hybrid system is a method for detecting binary interactions, because under normal circumstances only two proteins are assayed against each other. In contrast, complex purification schemes aim at isolating native protein complexes; for example, they will identify several interaction partners of a given bait. As the two large-scale yeast two-hybrid studies were carried out on yeast proteins, this problem may have been partially circumvented because endogenous yeast proteins may have formed complexes with bait and prey proteins and may thus have acted as bridging partners. Therefore, the Uetz and Ito datasets may also contain annotated interactions between two proteins that may in fact have been mediated by a third bridging partner.

9. THE CHALLENGES AHEAD

The future of proteomics is the definition of the exact function of every protein in a cell, and how this function may change in different cellular conditions, with different modification states of a protein, and with different interacting partners. For more than 10 years, the yeast two-hybrid system and its variations have played an important role in the study of physical protein–protein interactions. The recent application of the yeast two-hybrid system to large-scale screenings has culminated in the construction of several protein interaction maps that manage to connect the majority of proteins encoded by the organism under investigation. These screens facilitate the understanding of gene function in several ways. First, they provide insight into the possible functional roles of previously unknown genes by linking them to already characterized proteins. Second, they help to assign additional, novel functions to many previously characterized proteins. Third, they identify novel interactions between proteins that have previously been assigned to common biological processes based on circumstantial evidence such as transcriptional coregulation or subcellular colocalization.

Recently, alternative approaches for the identification of protein–protein interactions on a genome-wide scale have been developed, which are based on the characterization of protein complexes using mass spectrometry (Gavin et al., 2002; Ho et al., 2002). As opposed to the yeast two-hybrid system, which largely detects binary interactions, this method relies on the selective purification of entire protein complexes from the cell, followed by separation of its subunits and their identification by mass spectrometry. The methodology has been applied to characterize multiprotein complexes systematically on a large-scale in yeast and has identified hundreds of novel protein–protein interactions and protein complexes (Gavin et al., 2002; Ho et al., 2002). The future progress of interactive proteomics will involve refinement of such approaches, as well as the integration of data sets derived from as many different methods as possible. Ultimately, protein interaction networks that have been

constructed from different datasets will hopefully show little bias with regard to protein classes or functions, will represent the entire proteome under investigation, and will contain only interactions that have been proven by several methods, for example, they are likely to represent actual interactions that occur under physiological conditions within a cell. Once such representative maps have been constructed it will also become feasible to address the issue of regulation: by comparing representative protein interaction maps from the same cell type or organism under different growth conditions it may become possible to dissect cellular reactions in response to changing environmental conditions at the level of protein–protein interactions.

REFERENCES

- Aach, J., Bulyk, M.L., Church, G.M., Comander, J., Derti, A., Shendure, J. (2001). Computational comparison of two draft sequences of the human genome. *Nature* 409:856–859.
- Bartel, P.L., Roecklein, J.A., SenGupta, D., Fields, S. (1996). A protein linkage map of *Escherichia coli* bacteriophage T7. *Nat. Genet.* 12:72–77.
- Davy, A., Bello, P., Thierry-Mieg, N., Vaglio, P., Hitti, J., Doucette-Stamm, L., Thierry-Mieg, D., Reboul, J., Boulton, S., Walhout, A.J., Coux, O., Vidal, M. (2001). A protein–protein interaction map of the *Caenorhabditis elegans* 26S proteasome. *EMBO Rep.* 2:821–828.
- Fang, Y. and Macool, D.J. (2002). Development of a high-throughput yeast two-hybrid screening system to study protein–protein interactions in plants. *Mol. Genet. Genomics* 267:142–153.
- Dunn, J.J. and Studier, F.W. (1983). Complete nucleotide sequence of bacteriophage T7 DNA and the locations of T7 genetic elements. *J. Mol. Biol.* 166:477–535.
- Fields, S. and Song, O. (1989). A novel genetic system to detect protein–protein interactions. *Nature* 340:245–246.
- Finley, R.L., Jr. and Brent, R. (1994). Interaction mating reveals binary and ternary connections between *Drosophila* cell cycle regulators. *Proc. Natl. Acad. Sci. USA* 91:12980–12984.
- Flajolet, M., Rotondo, G., Daviet, L., Bergametti, F., Inchauspe, G., Tiollais, P., Transy, C., Legrain, P. (2000). A genomic approach of the hepatitis C virus generates a protein interaction map. *Gene* 242:369–379.
- Fromont-Racine, M., Rain, J.C., Legrain, P. (1997). Toward a functional analysis of the yeast genome through exhaustive two- hybrid screens. *Nat. Genet.* 16:277–282.
- Gavin, A.C., Bosche, M., Krause, R., Grandi, P., Marzioch, M., Bauer, A., Schultz, J., Rick, J.M., Michon, A.M., Cruciat, C.M., Remor, M., Hofert, C., Schelder, M., Brajenovic, M., Ruffner, H., Merino, A., Klein, K., Hudak, M., Dickson, D., Rudi, T., Gnau, V., Bauch, A., Bastuck, S., Huhse, B., Leutwein, C., Heurtier, M.A., Copley, R.R., Edelmann, A., Querfurth, E., Rybin, V., Drewes, G., Raida, M., Bouwmeester, T., Bork, P., Seraphin, B., Kuster, B., Neubauer, G., and Superti-Furga, G. (2002). Functional organization of the yeast proteome by systematic analysis of protein complexes. *Nature* 415:141–147.
- Goffeau, A., Barrell, B., Bussey, H., Davis, R.W., Dujon, B., Feldmann, H., Galibert, F., Hoheisel, J.D., Jacq, C., Johnston, M., Louis, E.J., Mewes, H.W., Murakami, Y., Philippsen, P., Tettelin, H., Oliver, S.G. (1996). Life with 6000 genes. *Science* 274:563–567.
- Ho, Y., Gruhler, A., Heilbut, A., Bader, G.D., Moore, L., Adams, S.L., Millar, A., Taylor, P., Bennett, K., Boutilier, K., Yang, L., Wolting, C., Donaldson, I., Schandorff, S., Shewnarane, J., Vo, M., Taggart, J., Goudreaault, M., Muskut, B., Alfarano, C., Dewar, D., Lin, Z., Michalickova, K., Willems, A.R., Sassi, H., Nielsen, P.A., Rasmussen, K.J., Andersen, J.R., Johansen, L.E., Hansen, L.H., Jespersen, H., Podtelejnikov, A., Nielsen, E., Crawford, J., Poulsen, V., Sorensen, B.D., Matthiesen, J., Hendrickson, R.C., Gleeson, F., Pawson, T., Moran, M.F., Durocher, D., Mann, M., Hogue, C.W., Figeys, D., and Tyers, M. (2002). Systematic identification of protein complexes in *Saccharomyces cerevisiae* by mass spectrometry. *Nature* 415:180–183.

- Ito, T., Tashiro, K., Muta, S., Ozawa, R., Chiba, T., Nishizawa, M., Yamamoto, K. Kuhara, S., Sakaki, Y. (2000). Toward a protein–protein interaction map of the budding yeast: A comprehensive system to examine two-hybrid interactions in all possible combinations between the yeast proteins. *Proc. Natl. Acad. Sci. USA* 97:1143–1147.
- Ito, T., Chiba, T., Ozawa, R., Yoshida, M., Hattori, M., Sakaki, Y. (2001). A comprehensive two-hybrid analysis to explore the yeast protein interactome. *Proc. Natl. Acad. Sci. USA* 98:4569–4574.
- James, P., Halladay, J., Craig, E. A. (1996). Genomic libraries and a host strain designed for highly efficient two-hybrid selection in yeast. *Genetics* 144:1425–1436.
- Kumar, A. and Snyder, M. (2001) Emerging technologies in yeast genomics. *Nat. Rev. Genet.* 2:302–312.
- McCrath, S., Holtzman, T., Moss, B., Fields, S. (2000). Genome-wide analysis of vaccinia virus protein–protein interactions. *Proc. Natl. Acad. Sci. USA* 97:4879–4884.
- Mrowka, R., Patzak, A. (2001). Is there a bias in proteome research? *Genome Res.* 11:1971–1973.
- Rain, J.C., Selig, L., De Reuse, H., Battaglia, V., Reverdy, C., Simon, S., Lenzen, G., Petel, F., Wojcik, J., Schachter, V., Chemama, Y., Labigne, A., Legrain, P. (2001). The protein–protein interaction map of *Helicobacter pylori*. *Nature* 409:211–215.
- Uetz, P., Giot, L., Cagney, G., Mansfield, T.A., Judson, R.S., Knight, J.R., Lockshon, D., Narayan, V., Srinivasan, M., Pochart, P., Qureshi-Emili, A., Li, Y., Godwin, B., Conover, D., Kalbfleisch, T., Vijayadamar, G., Yang, M., Johnston, M., Fields, S., Rothberg, J.M. (2000). A comprehensive analysis of protein–protein interactions in *Saccharomyces cerevisiae*. *Nature* 403:623–627.
- von Mering, C., Krause, R., et al. (2002). Comparative assessment of large-scale data sets of protein–protein interactions. *Nature* 417:399–403.
- Walhout, A.J., Boulton, S.J., Vidal, M. (2000a). Yeast two-hybrid systems and protein interaction mapping projects for yeast and worm. *Yeast* 17:88–94.
- Walhout, A.J., Sordella, R., Lu, X., Hartley, J.L., Temple, G.F., Brasch, M.A., Thierry-Mieg, N., Vidal, M. (2000b). Protein interaction mapping in *C. elegans* using proteins involved in vulval development. *Science* 287:116–122.

<http://www.springer.com/978-0-387-24531-7>

Proteomics and Protein-Protein Interactions

Biology, Chemistry, Bioinformatics, and Drug Design

Waksman, G. (Ed.)

2005, VIII, 323 p., Hardcover

ISBN: 978-0-387-24531-7