Nucleic Acids Research, 1999, Vol. 27, No. 4 **919–929**

SURVEY AND SUMMARY

Yeast forward and reverse 'n'-hybrid systems

Marc Vidal* and Pierre Legrain[1]

MGH Cancer Center, Charlestown, MA 02129, USA and [1]Laboratoire du Métabolisme des ARNs, CNRS URA 1300, Institut Pasteur and Hybrigenics, 75724 Paris Cedex 15, France

Received July 27, 1998; Revised and Accepted November 4, 1998

ABSTRACT

Since its original description almost 10 years ago, the yeast two-hybrid system has been used extensively to identify protein–protein interactions from many different organisms. Simultaneously, a number of 'variations on a theme' based on the original concept have been described. In one set of variations, systems were developed to detect other macromolecular interactions: DNA–protein (one-hybrid), RNA–protein (RNA-based three-hybrid) and small molecule–protein interactions (ligand-based three-hybrid). These different versions are collectively referred to here as 'n-hybrid systems'. In another set of variations, the original configuration of the two-hybrid fusion proteins was modified to expand the range of possible protein–protein interactions that could be analyzed. For example, systems were developed to detect trimeric interactions, ligand–receptor interactions or interactions that require particular post-translational modifications. Finally, the original concept was turned upside down and 'reverse n-hybrid systems' were developed to identify mutations, peptides or small molecules that *dissociate* macromolecular interactions. These reagents can be used to validate, in the relevant biological systems, the potential interactions identified with the 'forward n-hybrid systems'. The powerful genetic selections of the forward and reverse n-hybrid systems are proving useful in proteomic projects aimed at generating macromolecular interaction maps.

INTRODUCTION

Protein–protein interactions are critical to most biological processes, extending from the formation of cellular macromolecular structures and enzymatic complexes to the regulation of signal transduction pathways. This statement is becoming somewhat superfluous since in the last 20 years large numbers of stable interactions were uncovered and shown to be biologically relevant, in practically every field of molecular biology (1,2).

Nevertheless, this statement is still worth mentioning in light of recently published observations that implicate stable protein–protein interactions in unexpected mechanisms. First, many enzymatic activities are mediated by complexes much larger than originally anticipated. For example, although the original purification procedure of RNA polymerase II led to a 12 subunit complex (3), recent co-immunoprecipitation experiments using antibodies raised against one of its subunits gave rise to a 55 subunit complex required for accurate transcriptional regulation activity (4). Similarly, nuclear pre-mRNA splicing requires a highly dynamic and organized molecular machine composed of five small nuclear RNAs and >50 polypeptides (reviewed in 5,6). Second, the transmission of regulatory signals, originally described as successive catalytic activities required for the amplification of weak inputs into cellular responses, now appears to be mediated by protein complexes structurally constrained around scaffolding proteins (7–9). Finally, the widely accepted concept of transient enzyme–substrate interactions is being challenged by the identification of enzymes which stably associate with their substrates. For example, the cyclin-dependent kinases (CDK) required for cell cycle regulation have been found stably associated with some of their cognate substrates (10–13). Thus, the identification of protein–protein interactions remains fascinating and very helpful in understanding biological phenomena.

The identification of potential protein–protein interactions leads to hypotheses that need to be tested in the relevant biological systems. For such functional analyses, the isolation of mutant proteins specifically altered in their ability to interact with a potential partner ('interaction-defective alleles') can be critical. In these experiments, the interaction-defective mutant proteins are compared with their wild-type counterpart in a functional assay. Interaction-defective mutant proteins are predicted to exhibit functional defects if their corresponding wild-type versions indeed function in relevant protein–protein interactions. This strategy is exemplified in the characterization of the association between the adenovirus E1A oncoprotein and the retinoblastoma gene product (pRB) (14). To demonstrate the physiological relevance of this potential interaction, interaction-defective alleles of E1A were generated and shown to affect E1A's ability to transform cells (15). In more recent developments of this strategy, compensatory mutations that restore the interaction were isolated in the potential interaction partner. These compensatory changes are expected to restore the function mediated by the interaction. For example, in the case of the yeast splicing factor composed of Prp9p, Prp11p and Prp21p, it had been demonstrated that *prp9* and *prp11* thermosensitive mutations disrupt the interaction with Prp21p. Subsequently, *prp21-91*, an extragenic suppressor mutation of the *prp9* thermosensitive phenotype, was shown to restore Prp9–Prp21 interaction (16,17).

Prior to the development of the yeast two-hybrid system, both the identification of physical protein–protein interactions and

*To whom correspondence should be addressed. Tel: +1 617 724 3404; Fax: +1 617 726 7808; Email: vidal@helix.mgh.harvard.edu

their subsequent functional characterization traditionally relied upon time and labor intensive biochemical approaches. For example, the molecular cloning of genes encoding interacting proteins identified biochemically is often difficult, as is the identification of interaction-defective alleles. Moreover, while protein–protein interactions are important, many biological processes rely upon other macromolecular interactions such as DNA–protein and RNA–protein interactions. To understand how *cis*-acting DNA sites are involved in transcription or replication control and to address how particular RNA molecules are involved in splicing, translation or development, one powerful approach is to identify the protein(s) that stably binds these nucleic acids.

As will be described in this review, the yeast one-, two- and three- (or 'n'-) hybrid systems in a forward or reverse configuration provide genetic solutions to bypass the problems inherent in biochemical approaches, for both the identification of potential interactions and their characterization.

HISTORICAL PERSPECTIVE

The yeast two-hybrid system presents three major advantages over alternative assays for gene identification. First, since it is based on a powerful genetic selection scheme performed with a convenient microorganism, it allows very high numbers of potential coding sequences to be assayed in a relatively simple experiment. Second, it relies on an assay performed *in vivo* and thus it is not limited by the artificial conditions of *in vitro* assays. Finally, since it is based on a physical binding assay, a wide variety of protein–protein interactions can be detected and characterized following one single commonly used protocol. In this section, we attempt to give some historical perspective on these three aspects.

Geneticists have long exploited the advantages of growth selection to identify relatively rare events. In these experiments, a few growing colonies detected among very high numbers of non-growing cells are indicative of an interesting and rather rare event, such as a mutation in a particular gene or the loss of a plasmid. In all cases, the genotype and/or the growth medium are manipulated to obtain a set of conditions under which the starting 'wild-type' strain is not able to grow. For example, yeast auxotrophic mutants affected in their ability to synthesize pyrimidine can be selected on the basis of their resistance to ureidosuccinic acid since wild-type strains are sensitive to this drug (18). Since mutations occurring at frequencies as low as 10^{-10} can be recovered (19), similar yeast selections applied to the identification of mammalian cDNAs can be very powerful (the frequency of relatively rare cDNAs is 10^{-6}–10^{-7} in most non-normalized libraries).

Almost exactly 20 years ago, it was demonstrated that functional complementation gene cloning experiments could be performed in yeast mutants with the goal of cloning yeast genes (20). Soon thereafter, it was also shown that human (or other species) cDNAs could also be identified by functional complementation of relevant yeast mutations (21,22). Since then many human genes have been cloned this way and, in most cases, the yeast complementation cloning succeeded where biochemical attempts had not. The reasons for this success were probably related to the fact that, unlike in biochemical experiments, the physiological conditions need not be adjusted in each case. Although powerful, complementation cloning is limited by the availability of the relevant yeast mutants. However, in some settings, functional assays were designed to clone and/or characterize human genes in yeast without the need for any particular mutant (23,24). The two-hybrid system represents the ultimate example of such a strategy since it was designed from the outset for the detection of essentially any protein–protein interaction, independent of the function of the corresponding proteins.

The basic concept of the two-hybrid system emerged from the analysis of transcription factors such as the archetypal yeast Gal4p. These transcription factors increase the rate of transcription of their target genes by binding to upstream activating DNA sequences (UAS) and thus 'activating' RNA polymerase II at the corresponding promoters. It was demonstrated that the DNA binding and the activating functions are located in physically separable domains of Gal4p (25). These two domains are referred to as the DNA-binding domain (DB) and the activation domain (AD), respectively. In the most extreme version of such structure–function experiments, a hybrid protein, consisting of the bacterial LexA DB fused to the Gal4p AD, was shown to activate, in yeast cells, the transcription of a bacterial reporter gene containing the LexA operator site in its promoter (26). In addition, the fusion between DB and AD was shown to exhibit an unexpected level of structural flexibility. For example, 1% of random bacterial DNA sequences were found to be capable of encoding a functional AD when fused to the Gal4p DB (DB–AD) (27). This observation suggested, rather surprisingly, that the structural constraints of transcription factors for correctly activating RNA polymerase II upon DNA binding can be extremely loose. Furthermore, it was shown that functional DB–AD fusions need not be covalent. For example, functional dissection of VP16, the *trans*-activator of herpes simplex virus, predicted that AD domains can be recruited by protein–protein interactions (28). This aspect was further exemplified by using Gal80p, a repressor protein which interacts with Gal4p and lacks a naturally occurring AD: an artificial Gal80p–AD fusion could restore, *in trans*, the ability of a Gal4p mutant lacking a functional AD to activate transcription (29).

Although the observations described above led to a clearer idea of 'how transcription factors work', the resulting potential practical applications remained unclear for a while, until a seminal concept was developed that extended the possibilities of the system beyond the study of transcription factors. It was demonstrated that protein–protein interactions unrelated to transcription factors can reconstitute a functional transcription factor by bringing DB and AD into close physical proximity (30). In this setting, the 'architectural blueprint' for the 'reconstitution' of a functional transcription factor can be summarized as follows: DB–X/AD–Y, where X and Y could be essentially any proteins from any organism. This rather unexpected twist opened the doors to a plethora of applications. For example, soon thereafter, this system involving two hybrid proteins was shown to be useful to identify, in complex AD–Y libraries, genes encoding proteins that potentially interact with DB–X (often referred to as the 'bait' in these experiments) (31–33).

The reporter gene used in these pioneer experiments was the bacterial *lacZ* gene. Subsequently, growth selection markers such as *LEU2* and *HIS3* were introduced to allow powerful growth selections to be used for the detection of protein–protein interactions (34–36). Finally, responding to the need for functional validation of newly identified protein–protein interactions by means of selecting for interaction-defective alleles, the original system was turned upside-down by introducing counterselectable markers (37). The expression of such markers can be lethal under certain conditions (reverse two-hybrid system), such that yeast growth selection can be applied to select mutations, proteins,

peptides or small molecules that dissociate particular interactions. The next two sections describe in greater detail the different versions of currently available forward and reverse two-hybrid systems.

THE FORWARD TWO-HYBRID SYSTEM

The two-hybrid system has rapidly become an attractive method because it allows the genetic selection of genes encoding potential interacting proteins without the need for protein purification. However, such a genetic screen only provides an indirect *in vivo* assay for proximity between two proteins. With this in mind, we describe the early versions of the two-hybrid system with emphasis on the strategies for the evaluation of screens and their limitations. We then detail critical parameters to control for in a two-hybrid screen and the most recent versions of the system that address these issues. We refer to Table 1 for a summary of the pioneer and most widely used versions of the system.

Table 1. Description of the two different versions of the two-hybrid system

DNA binding domain (DB)	Activation domain (AD)	DB-X	AD-Y library	Reporter genes	Reference
Gal4 (1)	Gal4 (2)	yeast Sir4p	yeast genomic DNA partially restricted	*lac Z*	31
none	VP16 (3)	human SRF	inducible HeLa random-primed cDNAs	*lac Z*	33
Gal4 (1)	Gal4 (2)	human Rb	human lymphocytes oligodT-primed cDNAs	*HIS3* *lac Z*	35
LexA (4)	VP16 (3)	human H-Ras	inducible mouse embryos random-primed cDNAs	*HIS3* *lac Z*	36
LexA (5)	E. coli B42 (6)	human Cdc2	inducible HeLa oligodT-primed cDNAs	*LEU2* *lac Z*	34
Gal4 (1)	Gal4 (2)	yeast proteins	sonicated yeast genomic DNA	*HIS3* *lac Z*	45
Gal4 (1)	Gal4 (2)	yeast proteins	yeast genomic DNA partially digested	*HIS3* *lac Z* *ADE2*	43
Gal4 (1) low expression	Gal4 (2) low expression	70 proteins of different origins	oligodT-primed cDNAs	*HIS3* *lac Z* *URA3*	37

(1) Gal4 amino acids 1–147.
(2) Gal4 amino acids 768–881.
(3) VP16 amino acids 410–490.
(4) LexA amino acids 1–211.
(5) LexA amino acids 1–202.
(6) Ma and Ptashne (27).

In the first two-hybrid screening experiment described, Sir4p, a yeast protein involved in transcriptional silencing, was the bait and the AD–Y library was generated using partially restricted yeast genomic DNA (31). The reporter gene consisted of *lacZ* under the control of a Gal4p-responsive promoter, allowing a screen for positive colonies on the basis of a colorimetric assay. Among >200 000 transformants screened, two Sir4p-interacting AD–Y fusion proteins were identified and they both corresponded to Sir4p itself. Thus, a potential homodimerization involved in Sir4p function was uncovered, although no novel protein had been identified. Nevertheless, this result was still very encouraging because it demonstrated that interacting fusion proteins could be identified from a large collection of AD–Y clones.

A two-hybrid screen performed with the yeast Snf1p protein as bait demonstrated that the method could indeed lead to the identification of interacting proteins with physiological relevance (38). Among the potential interactors identified, one was encoded by *SIP1*, a gene that exhibits genetic interactions with *SNF1*. For example, overexpression of *SIP1* suppresses the defects caused by reduced Snf1p kinase activity.

Simultaneously, several groups demonstrated that protein–protein interactions from other species than yeast, such as mammals, could also be identified in the yeast two-hybrid system (32,33). In these experiments, several dozen candidates were selected from pools of 10^6 transformants. However, only a few exhibited bait-dependent reporter gene expression. Hence, these experiments helped define the limits of the *lacZ*-based colony color assay.

Consequently, to address the question of the large numbers of transformants that need to be screened to adequately survey the complexity of mammalian cDNA libraries, genetic selections based on growth assays were designed to replace the *lacZ*-based screening strategy. This was achieved simultaneously by several groups which developed novel yeast strains containing *selectable* reporter genes. In most cases, after an initial selection of positive clones, a secondary screen was developed with a second and independent reporter gene such as *lacZ* to allow higher specificity (34–36). These different systems allowed the two-hybrid strategy to be used widely to analyze proteins from many model organisms and led to a large number of publications describing the successful use of the method to identify potential interacting proteins (e.g. 39).

However, limitations of the system soon became evident but were not always reported in the literature. These limitations included very large numbers of clones with no biological relevance (‘false positives’) or the lack of recovery of expected interactions (‘false negatives’). Since the screen relies upon the transcriptional activation of reporter genes, any mutational events leading to an increase in the rate of transcription might be misinterpreted as the signature of a DB–X/AD–Y interaction. Thus, as for any genetic selection, it is crucial to develop the proper criteria to evaluate both the *specificity* and the *sensitivity* of the assay. Unfortunately, these parameters are not always considered and, consequently, some experiments have been somewhat over-interpreted, leading to an abundance of false positives.

Among the most important criteria used to sort through the putative positive clones is the verification that the AD–Y plasmids were selected because they indeed encode fusion proteins and not fusions to irrelevant small peptides. In this regard, it should be emphasized that in non-directional AD–Y libraries, five out of six fusions lead to hybrid proteins involving polypeptides that do not correspond to those naturally expressed in the organism of interest. These fusions can result from the cloning of RNA-coding, non-coding, antisense or out-of-frame DNA sequences and usually give rise to short peptides fused to AD (‘out-of-frame’ peptides). However , it has been suggested that in some cases out-of-frame fusions might encode *bona fide* fusion proteins through a frame shifting event (45).

Finally, it should be kept in mind that a two-hybrid screen does not necessarily select for direct interactions. For example, peculiar DB–X fusions such as DB–lamin are capable of allowing activation by the AD–Y fusions independent of a direct contact between X and Y. These ‘non-touching’ pairs often mediate their effect in a promoter context-dependent manner (40). In addition, indirect interactions have been reported where an endogenous

yeast protein serves as a bridge. For example the HIV-encoded protein Rev was first shown to interact with the yeast nuclear protein Rip1, a member of the nucleoporin proteins (41). It has now been demonstrated that this interaction likely occurs via yeast Crm1p, a protein involved in nuclear export (42). Nevertheless, the selection of such indirect protein–protein interactions can still be considered as an indication of a potential functional link between these two proteins.

Most recently, modified two-hybrid strategies have been designed to increase the specificity in order to limit the above problems of false positives. Several strategies were chosen. First, the expression level of the two hybrid proteins was reduced by using yeast centromeric vectors, which are maintained at low copies in the cell, and truncated versions of promoters such *ADH1* (32,37). Second, multiple reporter genes were designed for which the corresponding promoters are unrelated, containing very different TATA boxes and initiation sites. Since many false positives of the two-hybrid method are promoter context-dependent, the specificity of the assay is increased by scoring the phenotypes conferred by three different reporter genes in the same cell (37,40,43). Third, mating strategies were introduced which allow screening with many different baits under reproducible conditions (44,45).

The quality of a two-hybrid screen can also be analyzed in terms of the number of expected interacting proteins that were *not* recovered. These are often referred to as 'false negatives'. In many cases, the reasons for the lack of detection in the two-hybrid system are unknown, but several possibilities can be imagined. For example, the folding and/or the stability of a DB–X or AD–Y fusion protein could affect its transcriptional activation properties. Alternatively, particular fusion proteins might be toxic and affect the viability of the corresponding transformed cells. These restrictions are exemplified by the fact that even though, in principle, the reconstitution of an interaction in the context of the two-hybrid system should be bidirectional, i.e. the DB–Y/AD–X combination would be expected to score as well as DB–X/AD–Y, it is actually often not the case. In addition, the gene encoding an interacting protein may simply not be represented in the library due a low complexity or a bias in the representation of that clone.

For this reason, the source of DNA for the library is a key parameter for the assessment of the data. In most screens described above, cDNAs were derived from random- or oligo(dT)-primed RNAs. It must be kept in mind that, in contrast to genomic libraries, the relative representation of each cDNA closely reflects the endogenous expression level of the corresponding gene. Thus, interesting interacting proteins might be under-represented, if their RNA is expressed at relatively low levels. One solution here is the use of normalized AD–Y libraries (M.Brasch and M.Vidal, in preparation). The process of normalizing cDNA libraries consists of reducing the representation of highly expressed cDNAs (46). In addition, the choice of a random-primed versus an oligo(dT)-primed cDNA library considerably modifies the nature of the screen. Discrete protein domains are more likely to be screened with random-primed libraries, while clones encoding nearly full-length proteins are enriched when oligo(dT)-primed libraries are used. However, in high complexity libraries, the occasional internal priming events of oligo(dT) primers to poly(A) sequences can still lead to optimal coverage of N-terminal domain-encoding sequences (M.Brasch, personal communication).

In contrast, the complexity of genomic libraries is directly correlated to the number of independent clones that compose the library and to the size of the genome. For organisms that are encoded by compact genomes, i.e. with small intergenic sequences and few introns, screening a genomic library instead of a cDNA library for two-hybrid experiments is advantageous. Indeed, such an exhaustive screen was recently published in which a highly complex yeast genomic library was screened with various yeast proteins as baits (45).

In summary, the two central issues to be considered are the specificity and the sensitivity of the screen. Usually, higher specificity is desirable because a large number of false positives can dangerously obscure the biologically relevant interactors. However, it should be emphasized that increasing the specificity usually leads to a reduction in sensitivity and thus a greater number of false negatives.

As a final comment, it is important to reiterate that the magnitude of the two-hybrid read-out cannot be correlated with the biological significance of the interaction nor with the affinity of the interaction between the two wild-type proteins in the relevant organism. This is because, in addition to the affinity, many other parameters can influence the two-hybrid read-out. These include the expression, stability, nuclear localization and three-dimensional structure of a fusion protein and the fact that discrete domains might interact more strongly than the corresponding full-length protein.

THE REVERSE TWO-HYBRID SYSTEM

Potential protein–protein interactions identified by the two-hybrid system merely represent hypotheses that need to be tested back in the relevant biological systems. Approaches conventionally used include co-immunoprecipitation of endogenous proteins, co-immunolocalization or gradient sedimentation. However, the most direct approach genetically correlates the potential physical interaction with a biological parameter: the physical interaction is dissociated and the consequences are analyzed in a functional assay. Logically, one would expect that if the newly detected interaction is critical for a function of interest, the dissociation of the interaction would impair that function.

Conceptually, protein–protein interactions can be dissociated by the use of *cis*-acting mutations in one partner (referred to here as interaction-defective alleles) or *trans*-acting molecules such as dissociating proteins, peptides or small molecules. For example, interaction-defective alleles can be compared with their wild-type counterparts for their ability to functionally complement a knockout in the corresponding gene or for their ability to function in an expression assay in the relevant cells. Alternatively, the corresponding proteins can be expressed and purified and subsequently compared with their wild-type counterpart in an *in vitro* biochemical assay. However, since they usually score as recessive mutations, the use of interaction-defective alleles is compromised in cases where the wild-type protein cannot be removed from the assay. This would be the case in model organisms when no knockout is available for a particular gene of interest or in biochemical assays when the corresponding wild-type protein cannot be immunodepleted from the tested fractions. In these cases, dominant *trans*-acting peptides or small molecules that specifically affect the ability of a particular protein pair to interact could be used.

Until recently, this genetic strategy to validate potential interactions had not been used widely due to the technical difficulties of identifying informative interaction-defective alleles or specific dissociating molecules. The main challenge for interaction-

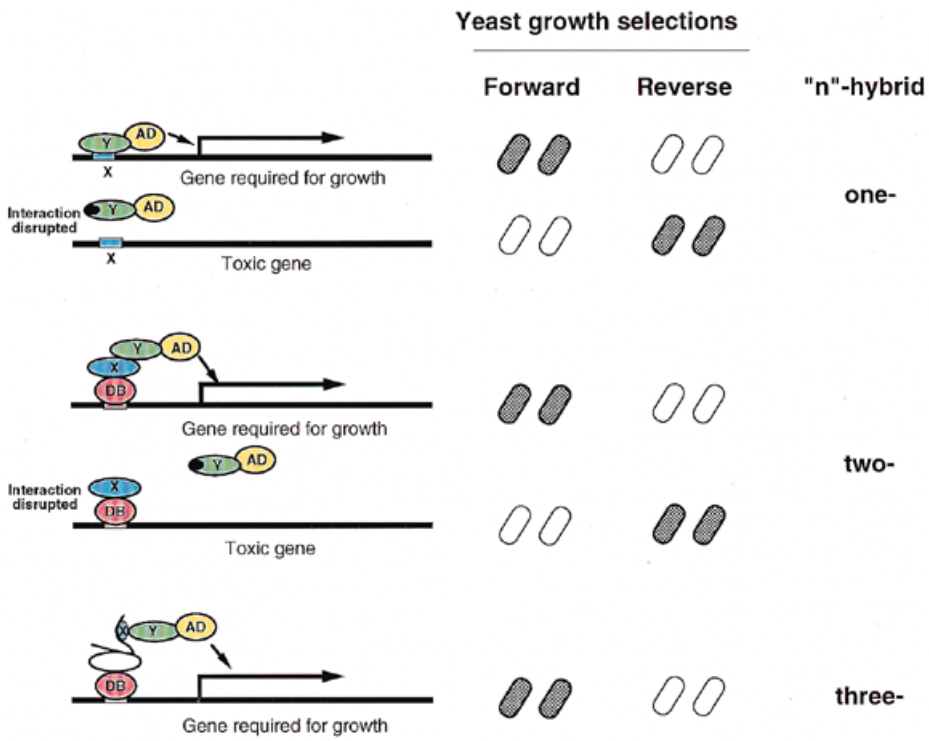

Figure 1. Forward and reverse n-hybrid systems. Four classes of macromolecular interactions can be detected using the original concept of reconstituting functional transcription factors. They include the following interactions: DNA–protein (one-hybrid system) (lines 1 and 2), protein–protein (two-hybrid) (lines 3 and 4), RNA–protein (RNA-based three-hybrid) and small molecule–protein (drug-based three-hybrid) (line 5). Potential interactions can be identified on the basis of the transcriptional activation of a 'gene required for growth' which confers a selective advantage. See the representation of two growing yeast patches (lines 1, 3 and 5) in the forward 'n'-hybrid configuration. It has also been demonstrated for the one- and two-hybrid systems that reagents that dissociate interactions can be identified by reversing the system. In this 'reverse' n-hybrid configuration, the reconstituted transcription factor activates the expression of a 'toxic gene' and thus it is the dissociation of the interaction that provides a selective advantage under those conditions. See the representation of two growing yeast patches (lines 2 and 4). X, DNA, protein, RNA or small molecule; Y, potential interacting protein; DB, DNA-binding domain; AD, activation domain (details in text).

defective alleles is the creation of subtle mutations that disrupt the interaction without grossly affecting the protein, i.e. in the context of a full-length, stable and correctly folded protein. Another problem is that for many novel interactors, no structural information is available and interaction-defective alleles cannot be rationally designed. To circumvent these two problems, large libraries of mutant alleles have to be generated randomly for each interaction partner with the goal of finding a few that exhibit the desired properties. Similarly, one might expect that complex libraries need to be screened to find a few specific dissociating peptides or small molecules.

The problem of screening very large libraries, of alleles or of molecules, should be overcome by using a genetic selection in which it is the *dissociation* of the interaction that provides a selective advantage. This situation is provided in the context of the reverse two-hybrid system. In this 'upside-down' version of the two-hybrid system, the wild-type DB–X/AD–Y interaction can be toxic or lethal for the yeast cells because a toxic marker is used as a reporter gene (negative selection) (37; Fig. 1). In this setting, DB–X/AD–Y dissociation confers a selective growth advantage which can conveniently identify both interaction-defective alleles and dissociating peptides or small molecules.

Below we will: (i) describe the different toxic reporter genes that can be used to provide negative selection; (ii) compare the multiple genetic strategies used to identify interaction-defective alleles in the context of stable full-length and correctly folded proteins; (iii) review the different aspects of the reverse two-hybrid that have already been validated.

Different yeast toxic markers that can be used for negative selection include *URA3* and *CYH2*. Traditionally, the *URA3* marker has been used most extensively because it allows both negative and positive growth selection, on medium containing 5-fluoroorotic acid (5-FOA) or on medium lacking uracil, respectively. The two-hybrid inducible *SPAL*::*URA3* reporter gene was designed by combining in the same promoter a strong *cis*-acting repressing sequence to maintain very low basal levels of expression and Gal4p-binding sites to allow Gal4p inducibility (47). A related reporter gene, based on an identical promoter conformation, is also available with LexA-binding sites (48). In cells containing these reporter genes, wild-type interactions confer a 5-FOA-sensitive phenotype. An alternative system uses the *CYH2* marker downstream of the *GAL1* promoter and in this case, wild-type interactions confer a cycloheximide-sensitive phenotype (49). Finally, in another strategy, the two-hybrid interaction activates expression of the Tet repressor, which in turn represses expression of the positive selectable marker *HIS3* (50). Under these conditions, wild-type interactions confer a histidine auxotrophic phenotype. Although many presumptions can be made on the respective advantages of these different selections, more experimental data will be needed for a thorough comparison.

Several genetic strategies have been implemented to select interaction-defective alleles in the context of a stable full-length

and properly folded protein. The first set of strategies is based on the selection for conditional mutations that retain some level of activity. For example, mutations that affect the interaction only weakly can be selected out from nonsense alleles (51). Additionally, thermosensitive mutations can also be selected that affect the interaction at the restrictive temperature while retaining wild-type activity at the permissive temperature (37,52). Finally, interaction-defective alleleles can be found that affect the interaction with one particular partner but not with another (51,53,54). In this particular case, simultaneous selections can even be performed in the same yeast cells if the first partner is fused to a DB that activates a reverse two-hybrid reporter gene while the second partner is fused to another DB that activates a forward two-hybrid reporter gene (55). Although conditional alleles can be very informative in some cases, it might be more convenient in other cases to use an exogenously added C-terminal marker protein to eliminate nonsense alleles. In this second set of strategies, the wild-type protein to be mutagenized is fused at its C-terminal end to an easily scorable protein. So far, both β-galactosidase (50) and green fluorescent protein (Y.Jacob, H.Endoh and M.Vidal, in preparation) have been used. Here, interaction-defective alleles are selected on the basis of negative selection combined with a β-galactosidase or fluorescence assay directly on the yeast colonies. In addition, fusions with a small 18 amino acid domain of the transcription factor E2F1, which is necessary and sufficient to interact with pRB, have also been used. Here, interaction-defective alleles are selected that retain the binding activity to a referential DB–pRB fusion protein (H.Endoh and M.Vidal, in preparation).

In the past 2 years, several aspects of the reverse two-hybrid system for interaction-defective alleles have been validated. First, it was shown that negative selection can be titrated to accommodate a large range of protein–protein interaction affinities (47). Second, well-characterized interaction-defective alleles that are found in human diseases were shown to rescue the negative growth phenotype of the reverse two-hybrid (47,50). Third, *de novo* selected interaction-defective alleles were described (47,50,53) and shown to be deficient for interaction in an *in vitro* binding assay (47,50). Finally, some of the interaction-defective alleles selected using the reverse two-hybrid system were shown to be functionally defective in several different biologically relevant assays. Human papillomavirus 16 E1 interaction-defective mutant proteins were shown to be defective in a replication assay in transfected cells (56). Similarly, yeast *ste5* interaction-defective alleles were shown to fail to complement the phenotypes of a *ste5* null mutation (53).

Finally, it should be added that such functional assays using interaction-defective alleles to validate potential interactions need to be substantiated by the proper controls. For example, in a few published cases, compensatory mutations were selected in the interacting partner by using a forward two-hybrid strategy and subsequently tested in a functional assay. This strategy implies that for relevant interactions, the restoration of the interaction can restore the function, at least partially (17,55,57,58).

Several aspects of the selection for *trans*-acting dissociating molecules have also been validated. On the one hand, a short adenovirus E1A peptide was shown to rescue the 5-FOA-sensitive phenotype in yeast cells expressing the DB–pRB/AD–E2F1 interaction (47). It is well known that, upon adenovirus infection, E1A mediates the dissociation of pRB from E2F1 (59). On the other hand, the small molecule FK506 was shown to rescue the 5-FOA-sensitive phenotype conferred by DB–FKBP12/AD–TGFβRIC using relatively simple and automatable plate assays (37,48,60). Although the interaction between FKBP12 and the TGF-β receptor IC has not been completely validated *in vivo*, it had been previously shown that FK506 affects its two-hybrid read-out (61). Last but not least, it was demonstrated in one case that small molecules can indeed be selected *de novo* from complex libraries using the reverse two-hybrid system (62). In this experiment, a novel compound, WAY141520, was identified in a collection of ~150 000 molecules as a dissociator of the interaction between two β3 and α1B subunits of N-type calcium channels. This compound was shown to inhibit N-type calcium channel activity in a specific manner (62).

ALTERNATIVE n-HYBRID SYSTEMS

Although protein–protein interactions form the basis of many biological processes, other macromolecular interactions such as DNA–protein and RNA–protein interactions are also critical. The original two-hybrid DB–X/AD–Y configuration was modified to accommodate the detection of such interactions.

The cloning of genes encoding DNA- or RNA-binding proteins has traditionally relied upon biochemical approaches. For example, the purification of a DNA-binding activity using DNA affinity chromatography can be followed by the cloning of a cDNA using a probe deduced from the sequence of the purified protein (63). Another approach involves the screening of an expression library with a radiolabeled DNA or RNA probe corresponding to the sequence of interest (64). Phage display can also be used to identify RNA-binding proteins *in vitro* (65). Although these methods have been successful in isolated cases, they are limited by the constraints of *in vitro* conditions. In addition, they are limited by the fact that each interaction tested requires a careful adjustment of the conditions. As argued for the identification and characterization of protein–protein interactions, the use of yeast selection can bypass these limitations.

The one-hybrid system used to study DNA-binding proteins is an extension, by simplification, of the two-hybrid concept (66; Fig. 1). In this configuration, the DB–X hybrid is eliminated and the the DNA Gal4p- or LexA-binding sites are replaced by a specific DNA sequence identified as an important binding site in the relevant biological system. The DNA-binding protein corresponding to this site can be identified and/or characterized as a fusion to AD, which extends the use of the system to proteins that are not necessarily transcriptional activators, such as proteins involved in transcriptional repression or DNA replication. Several versions of the one-hybrid system have been published (referenced in 66). They usually differ from each other by the choice of the selectable marker, such as *HIS3* or *lacZ*, and whether the marker is on a plasmid or integrated within the genome. The different steps of the protocols are as follows. The DNA-binding site is identified, located as precisely as possible and cloned into the promoter driving the reporter gene(s). This construct is introduced into yeast cells and tested for basal transcriptional activity since it is possible that an endogenously expressed yeast protein binds to the site or a neighboring sequence and activates the reporter gene. The subsequent steps are identical to two-hybrid selections and similar AD fusion libraries can be used (Fig. 1). An important control for the potential interactors selected from a one-hybrid experiment consists of verifying that the AD–Y protein loses its ability to activate the reporter gene when a mutant DNA sequence, known to be affected for its binding in the biological assay, replaces

the wild-type binding site used in the selection. So far the one-hybrid system has been used successfully to identify a few proteins, the yeast origin of replication complex Orc6p protein (67), the mammalian olfactory neuronal transcription factor Olf-1 (68), a mammalian silencer that restricts sodium channel expression in neurons (69) and a metal response binding factor (70).

The RNA-based three-hybrid system used to study RNA-binding proteins is a more complex extension of the two-hybrid concept (Fig. 1). In this configuration, both DB–X and AD–Y hybrid proteins are expressed and a third hybrid, an RNA molecule, bridges them (71). The RNA hybrid molecule consists of a fusion between a known RNA sequence (referred to here as RNA) which binds to a known RNA-binding protein (referred to here as RBP) and a novel RNA sequence to be characterized (referred to here as X by analogy with the two-hybrid configuration). In this setting RBP is fused to DB and reconstitution of the transcription factor relies upon two interactions, as indicated in the following 'blueprint': DB–RBP/RNA–X/AD–Y. Novel RNA-binding proteins Y can be identified and potentially characterized as fusions to AD. In one version of the RNA-based three-hybrid (72), DB–RBP consists of the LexA DB fused to the coat protein of bacteriophage MS2 and the hybrid RNA molecule consists of two MS2-binding sites (MS2RNA) linked to the RNA sequence X (72). In other words, DB–MS2, MS2RNA and AD are fixed while X, an RNA sequence, and Y, an RNA-binding protein, vary among experiments. The system was validated by reconstituting the well-characterized interaction between the RNA iron response element (IRE) and a protein to which it binds tightly (IRP1). Subsequently, RNA–protein interactions exhibiting a range of affinities between 10^{-7} and 10^{-11} M were shown to be detected in the assay (71). Most importantly, RNA-based three-hybrid selections have proven to be successful in identifying novel and biologically relevant interacting proteins. In these experiments, the RNA sequence of interest is identified in a biological system and the corresponding DNA sequence is cloned into a vector from which the proper hybrid RNA molecule can be expressed. This construct is introduced into yeast cells, along with DB–MS2, and tested for basal transcriptional activity since it is possible that the RNA–X molecule directly or indirectly activates reporter gene transcription in the absence of any exogeneous AD–Y hybrid protein. The subsequent steps are identical to the two-hybrid selections and, again, identical AD cDNA libraries can be used. An important control for the potential interactor selected from RNA-based three-hybrid experiments is verification that MS2RNA–X is required for the read-out. Thus proteins that bind DB–RBP directly or that are bridged with DB–RBP through an endogenous yeast RNA molecule or protein can be eliminated. In addition, when available, a functionally defective mutant version of X should be tested to verify specificity of the potential RNA-binding protein. So far, the RNA-based three-hybrid system has been used successfully to select a *Caenorhabditis elegans fem-3* mRNA binding factor which was shown genetically to be required for sexual fate in the hermaphrodite germline (73). In addition, a histone mRNA-binding protein, called SLBP, was identified using the system and shown to be required for processing of this mRNA (74).

Small organic molecules can be useful in elucidating biological mechanisms if they specifically affect functions of interest. Thus, it is often important to identify the protein targets of interesting small molecules. Traditionally, the identification of small molecule-binding proteins has relied upon biochemical methods including *in vitro* affinity chromatography and radiolabeled ligand binding reactions. However, as for the RNA-based three-hybrid system, a yeast genetic strategy has been developed that relies upon the presence of a third hybrid molecule bridging DB–X and AD–Y (Fig. 1). In this case, the third hybrid organic molecule is synthesized *in vitro* and consists of a well-known ligand (referred to here as ligand) that binds to a known small molecule-binding protein (referred to here as LBP) and is covalently attached to a small molecule of interest (referred to here as X by analogy with the two-hybrid configuration) (75). Such hybrid small molecules were shown to be useful for the inducible homodimerization of membrane receptors or activity of transcription factors (76). In the three-hybrid setting, LBP is fused to DB and reconstitution of the transcription factor relies upon two interactions indicated as follows: DB–LBP/ligand–X/AD–Y. Novel proteins Y that bind the small molecule X can be identified and/or characterized as fusions to AD. In one version of the small molecule-based three-hybrid system (75), DB–LBP corresponds to the LexA DB fused to the rat glucocorticoid receptor (GR) and the hybrid molecule is dexamethasone (Dex) covalently attached to a small molecule X. In other words, DB–GR, Dex and AD are fixed while X, a small molecule, and Y, a small molecule-binding protein, vary among experiments. It should be noted that Dex has previously been shown to penetrate yeast cells and activate GR in the nucleus (77). The system was validated by reconstituting the well-characterized interaction between FK506 and one of its binding proteins using micromolar concentrations of the Dex–FK506 hybrid molecule. In addition, small molecule-based three-hybrid selection was successful in identifying FKBP-encoding cDNAs from a Jurkat cell cDNA library. In these cloning experiments, the dependence upon the presence of the Dex–X molecule was tested by incubating the potential positives on medium containing an excess of X and the relevant positives corresponded to those for which the read-out was eliminated.

As for protein–protein interactions, there is a need for interaction-defective alleles corresponding to DNA–, RNA– and small molecule–protein interactions as well. Conceptually, the reasoning developed above on the negative selections possible with the reverse two-hybrid system should be applicable. We refer to these strategies as reverse n-hybrid systems. Just as it is possible to impose positive selections for AD–Y cDNAs that mediate different n-hybrid interactions, it should also be possible to select for interaction-defective alleles of AD–Y using the negative selections described above. This concept has been demonstrated in a reverse one-hybrid system configuration (47; Fig. 1). After reconstitution of p53 binding to its DNA recognition sequence incorporated within the promoter expressing the *URA3* reporter gene, interaction-defective mutant alleles were selected from a randomly generated library of mutant alleles. Remarkably, many alleles recovered from this selection corresponded to mutations found in patients (78). It is also conceivable that compensatory mutations in the corresponding DNA or RNA binding sites, or derivatives of the small molecules, could be recovered that restore the interaction.

INTRINSIC LIMITATIONS AND VARIATIONS OF THE NUCLEAR TWO-HYBRID

The conventional two-hybrid strategy suffers intrinsic limitations because of its reliance upon transcriptional activation. A major limitation is that neither the bait protein nor the potential interacting protein should be able to activate transcription on their own. This can be a problem with transcriptional activators which naturally contain

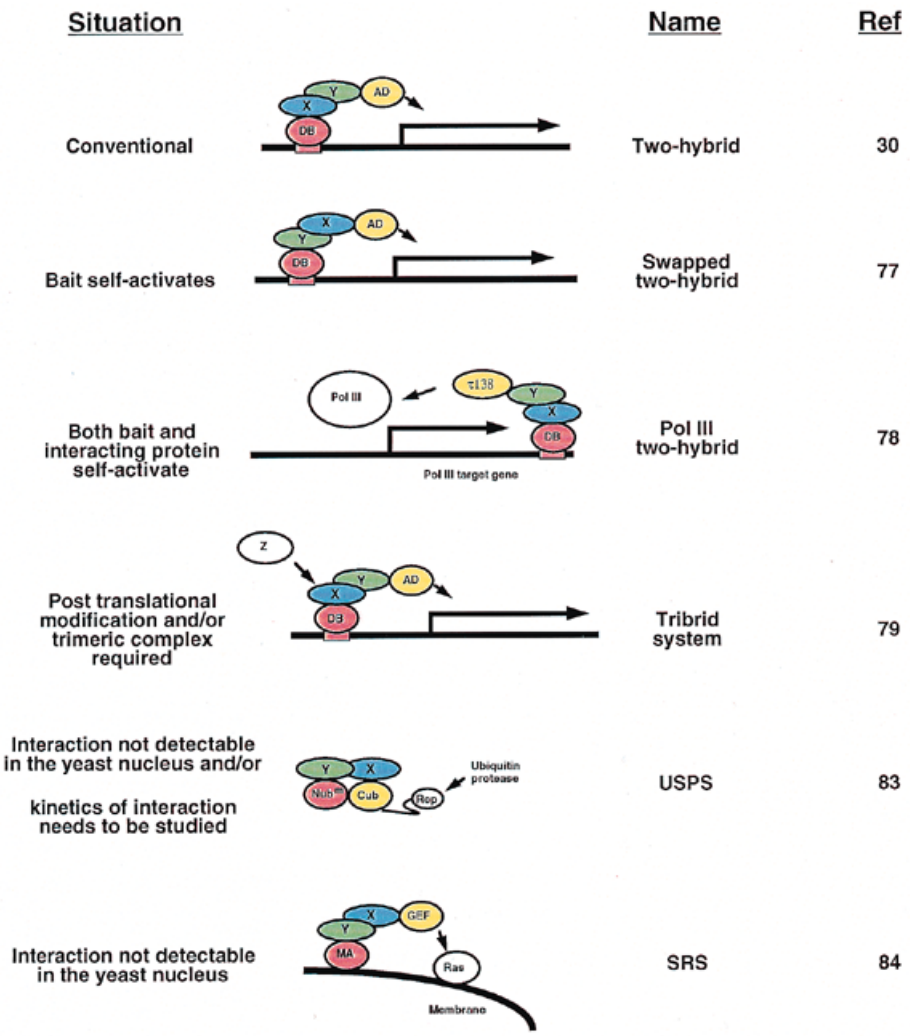

Figure 2. Variations on the two-hybrid theme. The conventional description of the two-hybrid system (line 1) has been extended to different configurations of reconstituted transcription factors (lines 2–4). Furthermore, it was demonstrated that molecules other than transcription factors can be reconstituted using two hybrid proteins as long as two interacting moieties can bring them into close proximity (lines 5 and 6). X, protein of interest; Y, potential interacting protein; DB, DNA-binding domain; AD, activation domain; PolIII, RNA polymerase III; τ138, activator of PolIII; Z, third molecule that modifies or stabilizes X–Y interaction; Nub^{m}, mutant form of N-terminal domain of ubiquitin; Cub, C-terminal domain of ubiquitin; Rep, reporter protein such as DHFR; MA, membrane-anchoring domain of SRC; GEF, GTPase exchange factor. The names and the references for the different systems are described in the text.

domains that activate the reporter genes when fused to the DB. In addition, many proteins other than transcription factors are also found to activate transcription when artificially fused to DB (27). Both classes of proteins are referred to here as 'self-activators'.

When dealing with self-activators, several approaches can be considered. First, the expression level of the DB–X fusion can be decreased by using centromeric vectors and/or weaker promoters (above). Second, when using *HIS3* as a reporter gene, the 3-aminotriazole (3AT) concentration can be increased to elevate the growth threshold of the host strain. The drug 3AT acts as a competitive inhibitor of the *HIS3*-encoded enzyme (35). Under these conditions, it is expected that, even though DB–X activates *HIS3* transcription to some extent, the DB–X/AD–Y interaction leads to more *HIS3* expression to overcome the growth threshold imposed by 3AT (79). However, these two strategies are not always successful, especially for strong self-activators.

Finally, a swapped two-hybrid system can be used to deal with self-activator baits (80; Fig. 2). In this approach, the bait protein was fused to the activation domain (AD–X) and a DB–Y library was screened. In the initial selection, many positive clones were selected (5000 for 2×10^6 transformants), most of them representing DB–Y that can self-activate. These self-activators could be eliminated by selecting for loss of the AD–X bait plasmid using *CYH2* (above; 35) and subsequently testing the ability of the DB–Y hybrid protein to activate *HIS3* expression. Among the initial 5000 positive colonies, 46 expressed AD–X bait-dependent interactors and among these a few have been shown to be functionally relevant (80).

In some extreme cases, including the study of protein complexes involved in RNA polymerase II-mediated transcription, one might imagine that, in addition to a self-activating bait, the expected interacting protein(s) could also act as a self-activator. In this case, the swapped two-hybrid configuration described above would eliminate the desired interacting protein. To bypass these problems, an RNA polymerase III (PolIII)-based two-hybrid system was designed (81; Fig. 2). Here the endogenous *SNR6*

gene, which encodes the essential U6 snRNA, is replaced by a mutated allele that contains Gal4p-binding sites in its 3′ regulatory region to constitute the reporter gene of the assay. The bait protein X is fused to the Gal4p DB (DB–X) and the potential interacting proteins Y are fused to the PolIII activating protein τ138 (τ138–Y). In this particular configuration, AD–X/τ138–Y interaction reconstitutes a functional PolIII-activating transcription factor. This two-hybrid version was shown to be capable of detecting a well-characterized interaction (81). However, it remains to be demonstrated that a screening procedure will be possible in such a PolIII-driven transcriptional system.

Another potential limitation of the two-hybrid involves the failure of yeast cells to undergo various post-translational modifications required for particular interactions in higher organisms, in particular, glycoproteins or lipoproteins for which the non-peptidyl part of the molecule is suspected to be involved in the interaction. To our knowledge, it has not been reported that such modifications naturally found in cytoplasmic proteins can indeed be reproduced in yeast to generate active two-hybrid fusions. However, other modifications have been reported. In one case, it has been shown that a tyrosine phosphorylation-dependent interaction could be detected when the corresponding kinase was co-expressed in yeast cells (82; Fig. 2). Several plasmids have been designed to allow conditional expression of such a 'third' *trans*-acting partner. In addition to mediating post-translational modifications of one of the two interacting proteins, this third protein can directly contribute to the formation of a trimeric complex with greater stability (83). These different variations that involve third partners as native proteins, in the absence of any fused domains, are referred to as 'tribrid systems'.

Finally, it has been shown that the two-hybrid system can be used as a tool to select for peptides that can interact with a protein of interest (84). This strategy has been applied with a combinatorial library of constrained 20mer peptides displayed by the active site loop of bacterial thioredoxin (85). If this protein–peptide two-hybrid system turns out to be applicable for the general selection of ligands that affect the activity of proteins, it might open an avenue toward the design of powerful research tools for *in vivo* functional studies in organisms for which genetic strategies are not readily available.

TWO-HYBRID SYSTEMS BASED ON DIFFERENT CELLULAR PROCESSES

One of the main limitations of the two-hybrid system is that both the DB–X and the AD–Y fusions need to be transported to and properly folded in the nucleus. Thus, most current two-hybrid vectors encode DB and AD with a nuclear localization signal to target the fusion proteins to the nucleus. However, proper localization can represent a major difficulty, especially when dealing with membrane-anchored proteins. Variations of the system that have been designed for those proteins that are not active in the yeast nucleus are described in this section.

One such system relies on the properties of the ubiquitin protein (86; Fig. 2). It is based on the observation that covalent addition of ubiquitin polypeptides can recruit particular proteases for the specific degradation of target proteins. When ubiquitin is experimentally cleaved into C-terminal (Cub) and N-terminal (Nub) domains, neither domain can mediate ubiquitin function, while simultaneous expression of both domains reconstitutes active ubiquitin *in trans*. However, a particular allele of the N-terminal domain (Nub^m) was found to be unable to mediate this functional interaction and, under these conditions, the two domains fail to reconstitute active ubiquitin. The 'ubiquitin-based split protein sensor' (USPS) takes advantage of these observations to detect exogenous protein–protein interactions (86). In this configuration, the protein of interest is fused to Cub (Cub–X) and the potential interacting proteins are fused to Nub^m (Nub^m–Y). Thus the Cub–X/Nub^m–Y interaction reconstitutes a functional ubiquitin which can target specific reporter proteins. In the original description of the system, the reporter protein used was human dihydrofolate reductase (DHFR) expressed as a fusion to Cub–X and cleavage of that fusion was used as an indication for X/Y interaction. This two-hybrid version was proven capable of reconstituting the well-characterized homodimerization of the yeast Gcn4p leucine zipper domain. However, this strategy has not yet been adapted for the selection of novel interacting proteins. It is possible to imagine a reporter protein that would be functionally inactive when fused to Cub–X. Under these conditions, Cub–X/Nub^m–Y interaction would provide a selective advantage by releasing the reporter protein from its covalent link to Cub–X (86).

A recent modification to bypass the reconstitution of a transcription factor takes advantage of a cell proliferation signaling pathway (87; Fig. 2). The 'Sos recruitment system' (SRS) uses as a signal recruitment to the yeast plasma membrane of the human guanyl nucleotide exchange factor (GEF) hSos, which can functionally substitute for the essential yeast Ras-GEF Cdc25p. In a yeast strain containing a thermosensitive allele of *CDC25*, growth at non-permissive temperature occurs only if hSos can be recruited efficiently to the membrane and this requires a myristylation signal, such as the one contained in the Src protein kinase. Thus, in the SRS system the bait protein of interest is fused to the GEF domain of hSos (GEF–X) and the interacting proteins are fused to the Src membrane-anchoring domain (MA) (MA–Y). Under these conditions, GEF–X/MA–Y interaction allows recruitment of hSos to the membrane and rescue of the yeast *cdc25* mutation. The SRS system has already been used successfully in a screen for c-Jun interacting proteins (87). In that experiment, expression of the MA–Y hybrids was controlled by a galactose-inducible promoter, thus allowing the convenient elimination of chromosome-encoded *cdc25-2* suppressors. However, one of the limitations of the system is that Cdc25p function could be rescued by overexpression of mammalian Ras proteins and other MA–Y fusion proteins, independent of any interaction with c-Jun. This class of proteins could be a serious limitation of this otherwise efficient alternative two-hybrid strategy. A possible solution to the problem is based on the observation that overexpression of GTPase activating protein (GAP) can suppress the bypass of Cdc25 function by Ras (88). However, it remains to be shown whether this improvement significantly reduces the number of false positives obtained in screening procedures.

TWO-HYBRID SYSTEMS IN PROTEOMICS

The recent release of the complete genomic sequence of several organisms has introduced the need for large-scale projects that address the function of the predicted proteins. The generation of protein interaction maps corresponding to complete genomes, or sets of expressed genes in particular tissues, would represent a reasonable strategy to add valuable predictive functional information to crude sequence alignments (89,90). As a genetic system, the

two-hybrid assay has been proposed as the optimal method for this enormously challenging task. Indeed, recent publications suggest that the method should allow such protein interaction maps to be generated in a reasonable period of time.

A first attempt to analyze protein–protein interactions in a systematic manner was proposed for *Drosophila* cell cycle regulators (91). It had been previously demonstrated that the plasmids encoding the DB–X and AD–Y fusions can be introduced into the same yeast cell by a mating procedure (44). In this method, one plasmid is introduced into haploid cells of one mating type and the other plasmid is introduced into cells of the opposite mating type. Since yeast mating is very efficient, numerous combinations of DB–X/AD–Y can be assayed simultaneously in diploid cells. Using a similar mating technique, a two-dimensional matrix was generated for cyclin-dependent kinases in which potentially interacting proteins were detected in diploid cells. This approach is nevertheless restricted by several considerations. First, only known or predicted proteins could be tested and this is an obvious key problem in genome-wide projects. Second, the matrix approach necessitates the use of a single set of growth conditions, which precludes the possibility of using an adapted selective pressure for every specific interaction. For example, the 3AT concentration cannot be adjusted to account for weak self-activation of certain DB–X baits. Third, the use of full-length proteins for both the DB–X bait fusions and the AD–Y interacting proteins might prevent the identification of several interactions due to various intrinsic problems such as toxicity, folding and degradation.

Another systematic approach was carried out for a small viral genome (92). In this experiment, a large number of combinations between the proteins encoded by the *Escherichia coli* bacteriophage T7 were tested and 25 interactions were identified. The approach consisted of mating yeast cells containing libraries of randomly generated fragments of T7 phage DNA cloned into the DB and AD fusion-encoding plasmids. Non-self-activating bait colonies were mated, 10 at a time, with 10^5–10^7 yeast colonies expressing the AD–Y library, leading to hundreds of positive combinations defining 19 interactions. Subsequently, more specific screens with defined baits led to the identification of six additional interactions. This analysis of a 55 protein-encoding genome led to the discovery of many novel interactions between different proteins as well as between different domains of the same protein, thus defining intramolecular interaction domains. The potential interactions identified have yet to be correlated with biological significance. However, this study demonstrated that it is possible to systematically and efficiently identify many of the possible connections between proteins encoded by a given genome.

Recently, an efficient mating strategy coupled with the generation of a large genomic library was used to perform exhaustive screens of the yeast *Saccharomyces cerevisiae* genome with proteins involved in RNA splicing (45). By using this strategy it was possible to classify the potential interacting proteins in sets of various heuristic values. The most likely candidates for interaction were subsequently used as baits and, within a reasonable period of time, potential interaction networks were built 'around' proteins of well-known functions. Subsequently, dedicated functional assays were applied to confirm the suspected function of the various interacting proteins (93,94). Finally, a directed approach has been suggested to generate a yeast protein interaction map by systematically cloning the complete set of predicted ORFs into both the DB and AD vectors of the two-hybrid system (95). Similar mating strategies could also be used here to generate the large number of combinations required to test every single pair of predicted proteins (95).

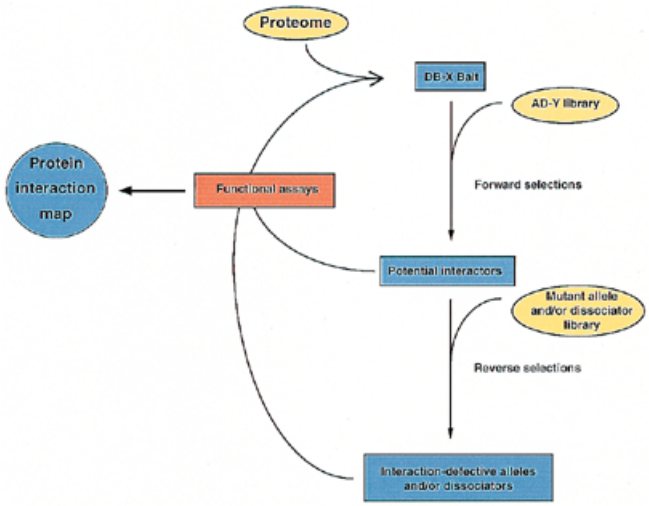

Figure 3. Forward and reverse two-hybrid systems in proteomics. Proteomes predicted by large-scale genome sequence analysis and cDNA sequencing lead to the choice of proteins of interest (DB–X baits). Forward two-hybrid selections with DB–X baits lead to potential interactors. These interactions can be validated using interaction-defective alleles and/or specific dissociators in functional assays. The combination of such functional assays and two-hybrid screens could lead to comprehensive protein interaction maps. Adapted from Walhout *et al*. (97).

It is conceivable that in the near future, similar strategies will be applied to protein interaction maps for higher eukaryotes (96,97). Many interactions will be reported and it will be critical to define parameters that allow a careful comparison of results within and between organisms. Two-hybrid screens will have to be performed under comparable standardized procedures. The combination of advanced forward and reverse n-hybrid strategies, in their present forms as well as using future improvements, might continue to deliver functional information on a more rational basis (Fig. 3).

ACKNOWLEDGEMENTS

We thank our colleagues and friends Mike Brasch, Rong Jiang, Linda Ko, Stan Fields and Marian Walhout for careful reading of the manuscript. We apologize to the authors of the papers not cited because of space limitations. The work in our respective laboratories is funded by NHGRI grant 1 R01 HG01715-01 A1 to M.V. and by EU grant Biotech 95-0009 to P.L.

REFERENCES

1 Lane,D.P. and Crawford,L.V. (1979) *Nature*, **278**, 261–263.
2 Linzer,D.I. and Levine,A.J. (1979) *Cell*, **17**, 43–52.
3 Sawadogo,M. and Sentenac,A. (1990) *Annu. Rev. Biochem.*, **59**, 711–754.
4 Koleske,A.J. and Young,R.A. (1994) *Nature*, **368**, 466–469.
5 Kramer,A. (1996) *Annu. Rev. Biochem.*, **65**, 367–409.
6 Guthrie,C. (1995) *Harvey Lect.*, **90**, 59–80.
7 Choi,K.Y., Satterberg,B., Lyons,D.M. and Elion,E.A. (1994) *Cell*, **78**, 499–512.
8 Printen,J.A. and Sprague,G.F. (1994) *Genetics*, **138**, 609–619.
9 Tsunoda,S., Sierralta,J., Sun,Y., Bodner,R., Suzuki,E., Becker,A., Socolich,M. and Zuker,C.S. (1997) *Nature*, **388**, 243–249.
10 Dowdy,S.F., Hinds,P.W., Louie,K., Reed,S.I., Arnold,A. and Weinberg,R.A. (1993) *Cell*, **73**, 499–511.

11 Ewen,M.E., Sluss,H.K., Sherr,C.J., Matsushime,H., Kato,J. and Livingston,D.M. (1993) *Cell*, **73**, 487–497.
12 Zhu,L., Harlow,E. and Dynlacht,B.D. (1995) *Genes Dev.*, **9**, 1740–1752.
13 Krek,W., Xu,G. and Livingston,D.M. (1995) *Cell*, **83**, 1149–1158.
14 Whyte,P., Buchkovich,K.J., Horowitz,J.M., Friend,S.H., Raybuck,M., Weinberg,R.A. and Harlow,E. (1988) *Nature*, **334**, 124–129.
15 Whyte,P., Williamson,N.M. and Harlow,E. (1989) *Cell*, **56**, 67–75.
16 Legrain,P. and Chapon,C. (1993) *Science*, **262**, 108–110.
17 Legrain,P., Chapon,C. and Galisson,F. (1993) *Genes Dev.*, **7**, 1390–1399.
18 Bach,M.L. and Lacroute,F. (1972) *Mol. Gen. Genet.*, **115**, 126–130.
19 Silar,P., Butler,G. and Thiele,D.J. (1991) *Mol. Cell. Biol.*, **11**, 1232–1238.
20 Nasmyth,K.A. and Reed,S.I. (1980) *Proc. Natl Acad. Sci. USA*, **77**, 2119–2123.
21 Lee,M.G. and Nurse,P. (1987) *Nature*, **327**, 31–35.
22 Elledge,S.J. and Spottswood,M.R. (1991) *EMBO J.*, **10**, 2653–2659.
23 Klein,R.D., Gu,Q., Goddard,A. and Rosenthal,A. (1996) *Proc. Natl Acad. Sci. USA*, **93**, 7108–7113.
24 Ishioka,C., Frebourg,T., Yan,Y.-X., Vidal,M., Friend,S.H., Schmidt,S. and Iggo,R. (1993) *Nature Genet.*, **5**, 124–129.
25 Keegan,L., Gill,G. and Ptashne,M. (1986) *Science*, **231**, 699–704.
26 Brent,R. and Ptashne,M. (1985) *Cell*, **43**, 729–736.
27 Ma,J. and Ptashne,M. (1987) *Cell*, **51**, 113–119.
28 Triezenberg,S.J., Kingsbury,R.C. and McKnight,S.L. (1988) *Genes Dev.*, **2**, 718–729.
29 Ma,J. and Ptashne,M. (1988) *Cell*, **55**, 443–446.
30 Fields,S. and Song,O. (1989) *Nature*, **340**, 245–246.
31 Chien,C.-T., Bartel,P.L., Sternglanz,R. and Fields,S. (1991) *Proc. Natl Acad. Sci. USA*, **88**, 9578–9582.
32 Chevray,P. and Nathans,D. (1992) *Proc. Natl Acad. Sci. USA*, **89**, 5789–5793.
33 Dalton,S. and Treisman,R. (1992) *Cell*, **68**, 597–612.
34 Gyuris,J., Golemis,E., Chertkov,H. and Brent,R. (1993) *Cell*, **75**, 791–803.
35 Durfee,T., Becherer,K., Chen,P.-L., Yeh,S.-H., Yang,Y., Kilburn,A.E., Lee,W.-H. and Elledge,S.J. (1993) *Genes Dev.*, **7**, 555–569.
36 Vojtek,A.B., Hollenberg,S. and Cooper,J. (1993) *Cell*, **74**, 205–214.
37 Vidal,M. (1997) In Bartel,P.L. and Fields,S. (eds), *The Yeast Two-Hybrid System*. Oxford University Press, New York, NY, pp. 109–147.
38 Yang,X., Hubbard,E.J.A. and Carlson,M. (1992) *Science*, **257**, 680–682.
39 Golemis,E.A. and Khazak,V. (1997) *Methods Mol. Biol.*, **63**, 197–218.
40 Bartel,P., Chien,C.T., Sternglanz,R. and Fields,S. (1993) *BioTechniques*, **14**, 920–924.
41 Stutz,F., Neville,M. and Rosbash,M. (1995) *Cell*, **82**, 495–506.
42 Neville,M., Stutz,F., Lee,L., Davis,L.I. and Rosbash,M. (1997) *Curr. Biol.*, **7**, 767–775.
43 James,P., Halladay,J. and Craig,E.A. (1996) *Genetics*, **144**, 1425–1436.
44 Bendixen,C., Gangloff,S. and Rothstein,R. (1994) *Nucleic Acids Res.*, **22**, 1778–1779.
45 Fromont-Racine,M., Rain,J.C. and Legrain,P. (1997) *Nature Genet.*, **16**, 277–282.
46 Soares,M.B., Bonaldo,M.F., Jelene,P., Su,L., Lawton,L. and Efstratiadis,A. (1994) *Proc. Natl Acad. Sci. USA*, **91**, 9228–9232.
47 Vidal,M., Brachmann,R., Fattaey,A., Harlow,E. and Boeke,J.D. (1996) *Proc. Natl Acad. Sci. USA*, **93**, 10315–10320.
48 Huang,J. and Schreiber,S.L. (1997) *Proc. Natl Acad. Sci. USA*, **94**, 13396–13401.
49 Leanna,C.A. and Hannink,M. (1996) *Nucleic Acids Res.*, **24**, 3341–3347.
50 Shih,H.M., Goldman,P.S., DeMaggio,A.J., Hollenberg,S.M., Goodman,R.H. and Hoekstra,M.F. (1996) *Proc. Natl Acad. Sci. USA*, **93**, 13896–13901.
51 Vidal,M., Braun,P., Chen,E., Boeke,J.D. and Harlow,E. (1996) *Proc. Natl Acad. Sci. USA*, **93**, 10321–10326.
52 Cayrol,C., Cabrolier,G. and Ducommun,B. (1997) *Nucleic Acids Res.*, **25**, 3743–3744.
53 Inouye,C., Dhillon,N., Durfee,T., Zambryski,P.C. and Thorner,J. (1997) *Genetics*, **147**, 479–492.
54 Jiang,R. and Carlson,M. (1996) *Genes Dev.*, **10**, 3105–3115.
55 White,M.A. (1996) *Proc. Natl Acad. Sci. USA*, **93**, 10001–10003.
56 Yasugi,T., Vidal,M., Sakai,H., Howley,P.M. and Benson,D. (1997) *J. Virol.*, **71**, 5942–5951.
57 Amati,B., Brooks,M.W., Levy,N., Littlewood,T.D., Evan,G.I. and Land,H. (1993) *Cell*, **72**, 233–245.
58 White,M.A., Nicolette,C., Minden,A., Polverino,A., Van Aelst,L., Karin,M. and Wigler,M. (1995) *Cell*, **80**, 533–541.
59 Fattaey,A.R., Harlow,E. and Helin,K. (1993) *Mol. Cell. Biol.*, **13**, 7267–7277.
60 Vidal,M. and Endoh,H. (1998) *Trends Biotechnol.*, submitted for publication.
61 Wang,T., Donahoe,P.K. and Zervos,A.S. (1994) *Science*, **265**, 674–676.
62 Young,K., Lin,S., Sun,L., Lee,E., Modi,M., Hellings,S., Husbands,M., Ozenberger,B. and Franco,R. (1998) *Nature Biotechnol.*, **16**, 946–950.
63 Kadonaga,J.T. and Tjian,R. (1986) *Proc. Natl Acad. Sci. USA*, **83**, 5889–5893.
64 Singh,H., LeBowitz,J.H., Baldwin,A.S.,Jr and Sharp,P.A. (1988) *Cell*, **52**, 415–423.
65 Laird-Offringa,I.A. and Belasco,J.G. (1995) *Proc. Natl Acad. Sci. USA*, **92**, 11859–11863.
66 Chong,J.A. and Mandel,G. (1997) In Bartel,P.L. and Fields,S. (eds), *The Yeast Two-Hybrid System*. Oxford University Press, New York, NY, pp. 289–297.
67 Li,J.J. and Herskowitz,I. (1993) *Science*, **262**, 1870–1874.
68 Wang,M.M. and Reed,R.R. (1993) *Nature*, **364**, 121–126.
69 Chong,J.A., Tapia-Ramirez,J., Kim,S., Toledo-Aral,J.J., Zheng,Y., Boutros,M.C., Altshuller,Y.M., Frohman,M.A., Kraner,S.D. and Mandel,G. (1995) *Cell*, **80**, 949–957.
70 Inouye,C., Remondelli,P., Karin,M. and Elledge,S. (1994) *DNA Cell Biol.*, **13**, 731–742.
71 Zhang,B., Kraemer,B., Sengupta,D., Fields,S. and Wickens,M. (1997) In Bartel,P.L. and Fields,S. (eds), *The Yeast Two-Hybrid System*. Oxford University Press, New York, NY, pp. 298–315.
72 SenGupta,D.J., Zhang,B., Kraemer,B., Pochart,P., Fields,S. and Wickens,M. (1996) *Proc. Natl Acad. Sci. USA*, **93**, 8496–8501.
73 Zhang,B., Gallegos,M., Puoti,A., Durkin,E., Fields,S., Kimble,J. and Wickens M.P. (1998) *Nature*, **390**, 477–484.
74 Wang,Z.F., Whitfield,M.L., Ingledue,T.C., Dominski,Z. and Marzluff,W.F. (1996) *Genes Dev.*, **10**, 3028–3040.
75 Licitra,E.J. and Liu,J.O. (1996) *Proc. Natl Acad. Sci. USA*, **93**, 12817–12821.
76 Crabtree,G.R. and Schreiber,S.L. (1996) *Trends Biochem. Sci.*, **21**, 418–422.
77 Godowski,P.J., Picard,D. and Yamamoto,K.R. (1988) *Science*, **241**, 812–816.
78 Brachmann,R., Vidal,M. and Boeke,J. (1996) *Proc. Natl Acad. Sci. USA*, , **93**, 4091–4095.
79 Yasugi,T., Benson,J.D., Sakai,H., Vidal,M. and Howley,P.M. (1997) *J. Virol.*, **71**, 891–899.
80 Du,W., Vidal,M., Xie,J.-E. and Dyson,N. (1996) *Genes Dev.*, **10**, 1206–1218.
81 Marsolier,M.C., Prioleau,M.N. and Sentenac,A. (1997) *J. Mol. Biol.*, **268**, 243–249.
82 Osborne,M.A., Dalton,S. and Kochan,J.P. (1995) *BioTechnology*, **13**, 1474.
83 Tirode,F., Malaguti,C., Romero,F., Attar,R., Camonis,J. and Egly,J.M. (1997) *J. Biol. Chem.*, **272**, 22995–22999.
84 Yang,M., Wu,Z. and Fields,S. (1995) *Nucleic Acids Res.*, **23**, 1152–1156.
85 Colas,P., Cohen,B., Jessen,T., Grishina,I., McCoy,J. and Brent,R. (1996) *Nature*, **380**, 548–550.
86 Johnsson,N. and Varshavsky,A. (1994) *Proc. Natl Acad. Sci. USA*, **91**, 10340–10344.
87 Aronheim,A., Zandi,E., Hennemann,H., Elledge,S.J. and Karin,M. (1997) *Mol. Cell. Biol.*, **17**, 3094–3102.
88 Aronheim,A. (1997) *Nucleic Acids Res.*, **25**, 3373–3374.
89 Lander,E.S. (1996) *Science*, **274**, 536–539.
90 Fields,S. (1997) *Nature Genet.*, **15**, 325–327.
91 Finley,R.L.,Jr and Brent,R. (1994) *Proc. Natl Acad. Sci. USA*, **91**, 12980–12984.
92 Bartel,P.L., Roecklein,J.A., SenGupta,D. and Fields,S. (1996) *Nature Genet.*, **12**, 72–77.
93 Siomi,M.C., Fromont,M., Rain,J.-C., Wan,L., Wang,F., Legrain,P. and Dreyfuss,D. (1998) *Mol. Cell. Biol.*, **18**, 4141–4148.
94 Rain,J.-C., Rafi,Z., Rhani,Z., Legrain,P. and Krämer,A. (1998) *RNA*, **4**, 551–565.
95 Hudson,J.R.,Jr, Dawson,E.P., Rushing,K.L., Jackson,C.H., Lockshon,D., Conover,D., Lanciault,C., Harris,J.R., Simmons,S.J., Rothstein,R. and Fields,S. (1997) *Genome Res.*, **7**, 1169–1173.
96 Regalado,A. (1998) *Technol. Rev.*, **101**, 51–53.
97 Walhout,M., Endoh,H., Thierry-Mieg,N., Wong,W. and Vidal,M. (1998) *Am. J. Hum. Genet.*, **63**, 955–961.