Yeast Two-Hybrid: State of the Art

Wim Van Criekinge[1*] and Rudi Beyaert[2]

[1] DEVGEN N.V. Technologiepark 9, 9052 Zwijnaarde-Gent and the [2]Department of Molecular Biology, Flanders Interuniversity Institute for Biotechnology and University of Ghent, B-9000 Ghent, Belgium.

[*] To whom correspondence should be addressed: Devgen N.V. Technologiepark 9, 9052 Zwijnaarde-Gent. Tel: 32-9-264-54-03; Fax: 32-9-264-54-18, E-Mail: wim.vancriekinge@devgen.com

ABSTRACT

Genome projects are approaching completion and are saturating sequence databases. This paper discusses the role of the two-hybrid system as a generator of hypotheses. Apart from this rather exhaustive, financially and labour intensive procedure, more refined functional studies can be undertaken. Indeed, by making hybrids of two-hybrid systems, customised approaches can be developed in order to attack specific function-related problems. For example, one could set-up a "differential" screen by combining a forward and a reverse approach in a three-hybrid set-up. Another very interesting project is the use of peptide libraries in two-hybrid approaches. This could enable the identification of peptides with very high specificity comparable to "real" antibodies. With the technology available, the only limitation is imagination.

INTRODUCTION

Protein-protein interactions

Protein-protein interactions are intrinsic to virtually every cellular process ranging from DNA replication, transcription, splicing and translation, to secretion, cell cycle control, intermediary metabolism, formation of cellular macrostructures and enzymatic complexes. The formation of large cellular structures such as the cytoskeleton, the nuclear scaffold, and the mitotic spindle result from complex interactions between proteins. Relatively smaller structures such as nuclear pores, centrosomes and kinetochores are beginning to be characterized and, in each case, protein-protein interactions seem to play a crucial role.

Apart from the evident structural requirements provided by a plethora of protein-protein interactions, there are a large number of transient protein-protein interactions that control and regulate a large number of cellular processes. All modifications of proteins involve such transient protein-protein interactions. Indeed kinases, phosphatases, glycosyl transferases, acyl transferases and proteases interact only transiently, i.e. for a limited period of time, with their protein substrates. Such protein-modifying enzymes encompass a large number of fundamental processes such as cell growth, the cell cycle,

metabolic pathways and signal transduction. Surprisingly, very large protein complexes also mediate many of these enzymatic activities. Transmission of regulatory signals from the external environment to relevant locations in the cell was originally thought to consist of successive catalytic activities that could amplify a weak signal into a significant cellular response. However, more recent experiments suggest that in many signal transduction pathways, the catalytic activities involved, such as protein kinases, may bind strongly to their protein substrates. In addition, structural proteins required for signal transmission have been suggested to act as scaffolds, bridging several proteins involved at consecutive steps in a signal transduction pathway (1). Thus, signal transduction pathways might be considered as large protein structures through which a signal is being transmitted. A striking example is the formation of the Death-Inducing Signaling Complex (DISC) at the Fas receptor. Only a few seconds after receptor triggering, a highly complex mixture of signal transducing molecules is recruited to the intracellular part of the receptor. This newly formed complex is capable of transmitting multiple specific signals provoking a highly regulated cellular response (2).

Alteration of protein-protein interactions is known to contribute to many diseases. Hence, the manipulation of protein-protein interactions that contribute to disease is a potential therapeutic strategy. Summarising, protein-protein interactions make up biological machines that are like intricate three-dimensional jigsaw puzzles, forming arrays of interlocking protein components that assemble and disassemble over time and in response to complex signals.

Tools for the study of protein-protein interactions

The study of protein-protein interactions can be conceptually divided into three major domains: identification, characterization and manipulation. In general, assemblies of proteins have been analyzed using two complementary approaches: the biochemical and the genetic. In the well-known analogy to understanding how car runs, biochemists disassemble the engine, transmission and body, characterize all the pieces and attempt to rebuild a working vehicle. Geneticists, by contrast, break single components, turn the key and try to determine what effect the single missing part has on the car's operation. This implies that genetic methods often require a specific phenotype before they can be carried out.

Traditionally, the tools available to analyze protein-protein interactions in multicellular organisms have been restricted to biochemical approaches. However, despite obvious advantages, biochemical approaches can be time-consuming. Biochemical methods that detect proteins that bind to another proteins generally result in the appearance of a band on a polyacrylamide gel. These methods are sometimes referred to as physical methods and include protein affinity chromatography, affinity blotting, immunoprecipitation and cross-linking (3).

Protein probing uses a labeled protein as a probe to screen an expression library in order to identify genes encoding interacting proteins. Since all combinations of protein-protein interactions are assayed, including those that might never occur *in vivo*, the possibility of identifying artifactual partners exists and is a typical disadvantage of most exhaustive screening procedures. A second drawback derives from the use of a bacterial host, where not all posttranslational modifications needed for the interaction might occur. A third disadvantage is that screening rather than selection is used as the means of detection, which inherently limits the number of plaques that can be assayed. The "Phage Display" method is

noteworthy in this matter. Here the methodology is based on the fact that an *Escherichia coli* filamentous phage can express a fusion protein on its surface.

A major advantage of the described method is the affinity purification step through "panning" cycles, that enables one to enrich every cycle 1000-fold for specific phages that contain an interacting protein. Disadvantages of phage display include the size limitation of protein sequence for polyvalent display and the requirement that proteins be secreted form *E. coli.* Moreover, all phage-encoded proteins are fusion proteins, which may limit the activity or accessibility for the binding of some proteins (3).

THE TWO-HYBRID SYSTEM

What is it?

Apart from the above-mentioned methods, a new technology has been developed during the past decade. This technique, entitled "two-hybrid" or "interaction trap", enables not only the identification of interacting partners but also the characterization of known interaction couples and even embodies the technological means to manipulate protein-protein interactions.

The modular properties of GAL4 and other transcription factors in general fostered this strategy. Indeed many eukaryotic transcription activators have at least two distinct functional domains, one that directs binding to a promoter DNA sequence and one that activates transcription (4,5). This fact was illustrated by exchanging DNA binding domains and activation domains from one transcription factor to the next while retaining its function. It was shown that the activation domain of yeast GAL4 could be fused to the DNA-binding domain of *E. coli* LexA to create a functional transcription activator in yeast (6).

The "two-hybrid" technique exploits the fact that the DNA-binding domain of GAL4 is incapable of activating transcription unless physically, but not necessary covalently associated with an activating domain

Ma and Ptashne (7) demonstrated this principle for the first time. They showed that the GAL80 protein, normally a negative regulatory protein that interacts with GAL4, could be converted into a transcriptional activator by fusing it to an activation domain (AD). The activation by this fusion protein, GAL80-AD, was dependent on the presence of a GAL4 derivative bearing the GAL80 binding domain (C-terminal 30 amino acids) but lacking its own activation domain (7).

The actual use of those different functional modules of a transcription factor to study protein-protein interactions was first proposed by Fields and Song (8). They demonstrated the "proof-of-concept" by using SNF1 fused to the DNA-binding domain (DB) and SNF4 to an activation domain (AD). Only after expression of these two chimeras, and subsequent interaction of SNF1 and SNF4, did they reconstitute a functional transcription factor that is able to induce reporter gene expression. These initial experiments confirmed that a transcriptional read-out could be used as a tool to study interactions between proteins not involved in the transcription process.

In general, in any two-hybrid experiment a protein of interest is fused to a DNA-binding domain and transfected in a yeast host cell bearing a reporter gene controlling this DNA-binding domain. When this fusion protein cannot activate transcription on its own, it can be used as "bait" or as a "target" to screen a library of cDNA clones that are fused to an activation domain. The cDNA clones within the library that encode proteins capable of forming protein-protein interactions with the bait are identified by virtue of their ability to cause activation of the reporter gene. So the yeast two-hybrid system is devised to identify genes encoding proteins that are physically associated with a given protein *in vivo*. Since the emergence of the two-hybrid approach in 1989, a number of improvements have been incorporated that have increased its applicability (New developments).
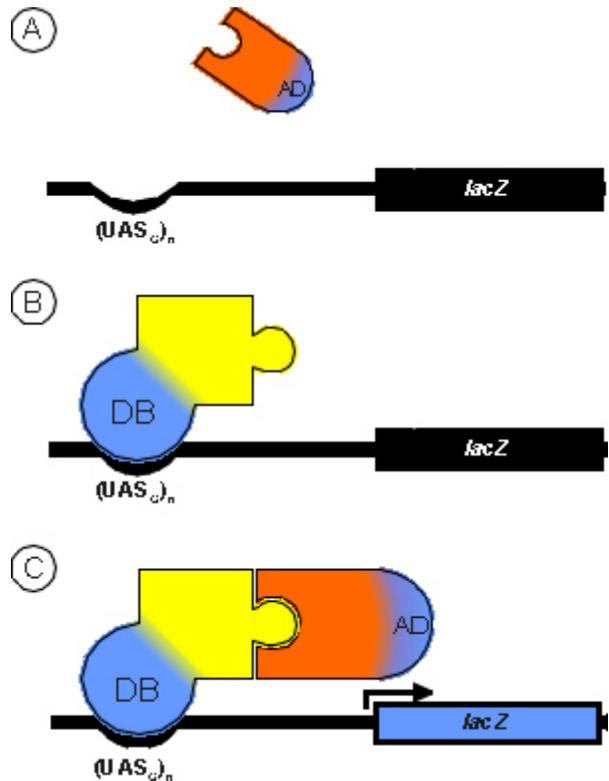
Why not use the two-hybrid system?

It should be noted, however, that the two-hybrid system does not provide a solution for all protein-protein problems. For different experimental reasons some proteins are not suited for this approach. The sceptics about the use of the two-hybrid system are furnished with a summary of the extensive list of disadvantages and drawbacks in the next paragraphs.

Fig. 1: Principle of the Two-hybrid system. (A), (B) Two chimeras, one containing the DNA-binding domain (DB: blue circle) and one that contains an activation domain (AD: half blue circle), are co-transfected into an appropriate host strain. **(C)** If the fusion partners (yellow and red) interact, the DB and AD are brought into proximity and can activate transcription of reporter genes (here LacZ).

Since the read-out of the two-hybrid system makes use of a transcription event, one of the most crucial initial experiments is to check whether your favorite protein (YFP) is able to initiate transcription. If this is the case, it might seriously handicap the successful use of this protein in any two-hybrid approach (see Auto-activation).

An obvious critique concerns the extensive use of chimeras. The use of artificially made fusion proteins always embodies a potential risk. The fusion might change the actual conformation of the bait and/or prey and consequently alter functionalities. This misconformation might result in a limited activity or in the inaccessibility of binding sites. However, the use of tagged proteins in general has been very successful in many biotechnological approaches. This success might rely on the fact that protein domains can fold rather independently, enabling the co-existence of different, even artificially introduced, modules in the same protein.

The best control to assay the correct conformation of your favorite protein is to clone a known positive interactor in the appropriate vector and "two-hybrid-assay" this interaction. This will only work if both

proteins are folded correctly. This kind of experiment is only conclusive for the domains involved in the "positive" interaction, which might differ from the ones involved in the new interactions that might be of interest. In summary, this category of drawbacks can be labeled "(mys)steric", related to folding and the three dimensional structure of the protein assayed. In this respect it is noteworthy that the reciprocal transfer of proteins, i.e. switching proteins from DNA-binding fusions to activating domain fusions, is not trivial, supporting the fact that (mys)steric constraints on folding are involved. This switching or "swapping" might provide an empirical way to escape the problem.

One of the most ambiguous disadvantages is that the two-hybrid system makes use of yeast, *S. cerevisiae*, as a host. This implicates, as mentioned above, that YFP must be able to fold correctly and exist as a stable protein inside the yeast cells. The use of yeast can also be seen as an advantage, since yeast is closer to higher eukaryotics than *in vitro* experiments or those systems based on bacterial hosts.

A major disadvantage of assaying protein-protein interaction in any heterologous system is that some interactions depend upon posttranslational modifications that do not, or inappropriately, occur in yeast. Such modifications are frequent and include the formation of disulfide bridges, glycosylation and most commonly phosphorylation. Some new two-hybrid systems, however, try to circumvent this inconvenience by co-expressing the enzyme responsible for the posttranslational modification.

Since the two-hybrid system needs the fusion proteins to be targeted to the yeast nucleus, it might be a disadvantage for extracellular proteins or proteins that contain strong(er) targeting signals.

When screening libraries, a good representation is crucial. In classical two-hybrid library preparations only one out of six fused cDNAs is in the correct frame, pushing the total number of independent clones to be screened to over a million, at the border of practical feasibility. Making directional libraries of a relevant tissue or cell type might be a solution. Another solution might be to go for less complex organisms like *C. elegans*.

Screening of libraries selects for optimized interactions. Many isolates may not represent full-length cDNA. Indeed, it has been shown that subdomains may interact better than full-length clones, probably reflecting domain function during folding of the protein. The best way to encompass this problem is probably to clone only full-length cDNAs in the correct open reading frame. Although extremely labor intensive, this approach was taken to establish the complete yeast protein linkage map (see Whole genome approaches using the two-hybrid system).

Since only reporter gene activity is measured, it is impossible to exclude the possibility that a third protein Z is bridging the two interacting partners. Although this possibility is rather unlikely and might even be considered as "specific," it holds for many of the conventional biochemical techniques.

Some proteins might become toxic upon expression in yeast. A number of proteins, such as cyclins or homeobox gene products are indeed toxic when expressed and targeted into the yeast nucleus. Such genes might be counterselected during growth and result in problems. The use of an inducible promoter might circumvent the problem. Other proteins might proteolyse essential yeast proteins or proteins essential for the system like the DNA binding domain or the activation domain.

Since all combinations of protein-protein interactions are assayed, the possibility of identifying artifactual partners exists and is a typical disadvantage of all exhaustive screening procedures. Due to the so-called time/space constraints it is potentially possible that both proteins, although able to interact, are never in close proximity to each other within the cell. The two proteins could be expressed in different cell types, or even when found in the same cell they could be localized in distinct subcellular compartments. Moreover, interacting proteins can be expressed at different points during embryogenesis or during homeostasis (e.g. at different time points in the cell cycle). So once two interacting partners are identified, the biological relevance of this interaction remains to be determined.

Why use the two-hybrid system?

Apart from the above-mentioned drawbacks, the two-hybrid system has some clear advantages over classical biochemical and genetic approaches. First of all it embodies an *in vivo* technique using the yeast host cell as a live test tube. This yeast system brings the higher eukaryotic reality closer than most *in vitro* approaches or techniques based on bacterial expression. Appealing features of this system are the minimal requirements to initiate a screening. Only the cDNA, full-length or even partial of the gene of interest is needed, in contrast to sometimes-high quantities of purified proteins or good quality antibodies needed in classical biochemical approaches.

Weak and transient interactions, often the most interesting in signaling cascades, are more readily detected in two-hybrid since the genetic reporter gene strategy results in a significant amplification. It is useful to keep in mind that there is a trade-off between the identification of weak interactions and the number of false positives encountered when performing a screening procedure. Apart from the ability to screen libraries, the two-hybrid system also allows for the analysis of known interactions. This can be achieved by pinpointing crucial residues for interaction or by a functional characterization of the entire subdomain. By doing semi-quantitative experiments one can even interpret affinities from two-hybrid experiments. It was demonstrated that the strength of interaction as predicted by the two-hybrid approach generally correlates with that determined *in vitro*, permitting discrimination of high-, intermediate- and low-affinity interactions (34). In addition, binding affinities of peptides to retinoblastoma (Rb), as determined by surface plasmon resonance, correlated with results from the two-hybrid assay.

The two-hybrid system was predicted to be limited to the analysis of cytoplasmic proteins. Indeed extracellular proteins or protein domains are often N-glycosylated and contain disulfide bonds, both of which are not expected to occur in the yeast nucleus (9). However, several successes were reported with transmembrane receptors. Appropriate extracellular receptor-ligand interactions were demonstrated for the growth hormone,prolactin and growth hormone releasing receptors (10,11). Thus, receptors with whole extracellular critical ligand binding determinants can sometimes be evaluated by the two-hybrid system. But, it may be inappropriate for receptors with determinants in transmembrane domains that form intramembraneous ligand binding pockets (11).

One of the most appealing features of the yeast two-hybrid system is that the identification of an interacting protein implies that at the same time the corresponding gene is cloned. Two-hybrid screens

are sometimes referred to as functional screens, since interacting proteins might give a functional hint if at least one of the partners has a known functional commitment in a well understood signaling pathway. Trying to attribute function to an unknown target is often more difficult. Here, the identified partners need to be known or the problem will propagate. Although the outcome of a screening often results in many new hypotheses, they still need to be validated by other techniques. As a conclusion, there is enough reason to remain sceptic about two-hybrid screenings but the most convincing argument in favor of the two-hybrid is the number and speed in which many signaling cascades have been resolved in molecular detail.
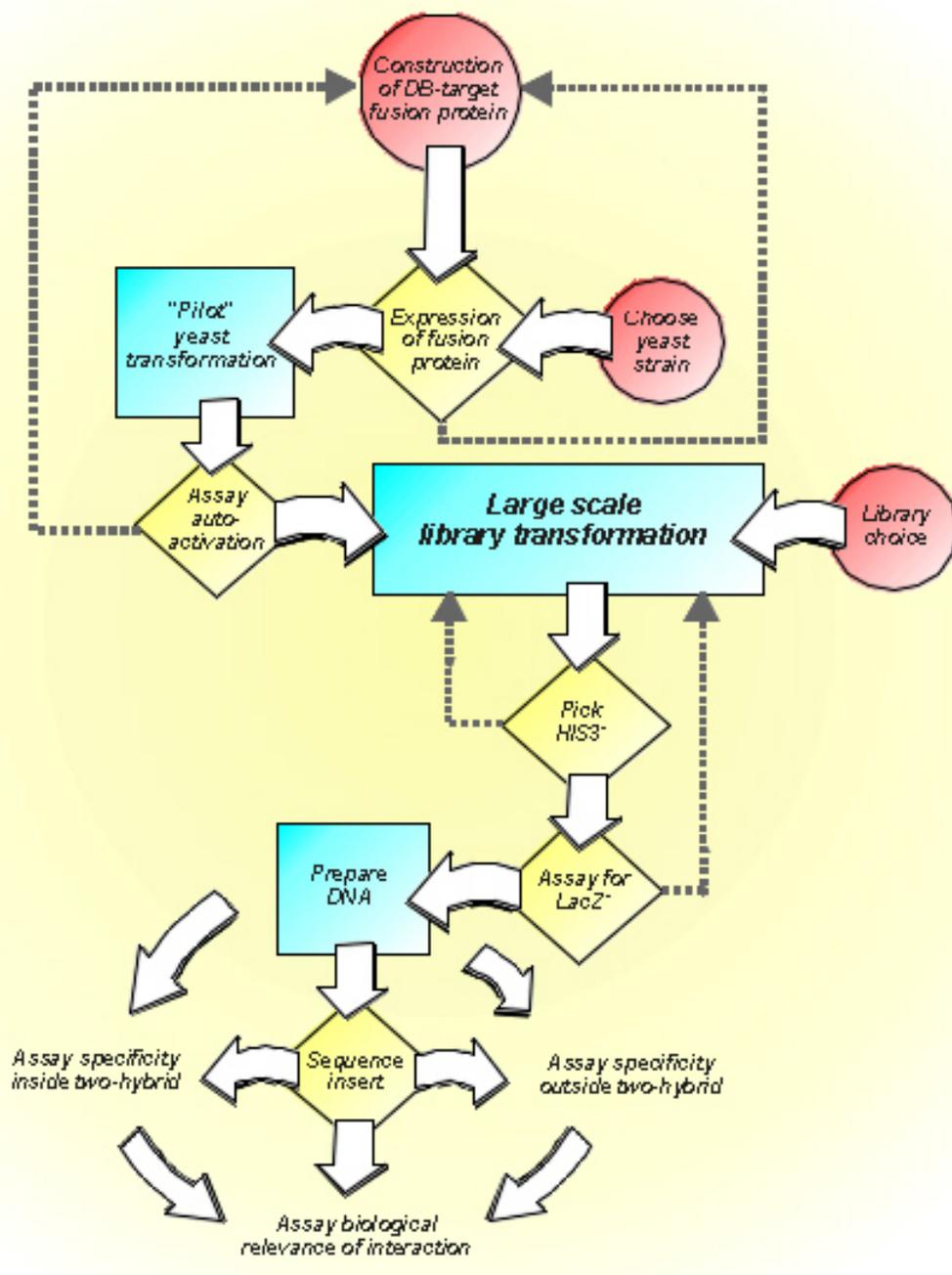
Fig. 2: General flowchart of a hypothetic two-hybrid screen. Pointed lines should be followed in case of negative results.

MATERIALS AND METHODS

Introduction

Actual screening involves many choices, since several comparable two-hybrid systems are available. Many of the most encountered problems and potential pitfalls are reviewed in this section. In most cases a typical two-hybrid experience goes through a sequence as outlined in Fig. 2.

Construction of the "target" or "bait"

When performing a two-hybrid screen, the first decision to be made is the choice of the most appropriate vector system. A large number of different DNA-binding domain (DB) and transcription activation domain (AD) containing vectors have been used successfully. The most extensively used vectors are GAL4 based, probably because they were the first commercially available. Alternative systems make use of the DB of the bacterial LexA protein and the AD of VP16 or the so-called B42AD. Both systems, GAL4 and LexA, have advantages and drawbacks which make the choice more difficult. In the following paragraphs the major differences between the two systems will be pointed out together with the major pitfalls in a typical screening procedure. Since the LexA and the GAL4-based two-hybrid systems have different properties, it is not unreasonable to imagine that some interactions might be detected differently in both systems. Trying both will increase the chance of success.

The promoter regulates the expression level of the target protein

The 1500-bp full length ADH1 promoter, that normally drives the expression of the metabolic enzyme alcohol dehydrogenase 1, leads to high-level expression of sequences under its control. This promoter is used in the pAS2(-1) and pLexA plasmids that are used to clone the target-fusion. It is also present in pGAD-GH that can be used to clone the cDNA library. Expression from this promoter is maximal during logarithmic growth of the yeast cells and becomes repressed in late log phase by ethanol accumulation in the medium. This explains why Western blot analysis of the expressed fusion protein gives the best results when logarithmic growing yeast is used. Although the ADH1 is generally considered to be a strong constitutive promoter, expression is actually repressed as much as 10-fold on non-fermentable carbon sources (12). In contrast to this full-length promoter, several cloning vectors, including pGBT9 (DB) and pGAD424 (AD) contain a truncated 410-bp ADH1 promoter. Expression from this promoter leads to low or very low levels of fusion protein expression that are hardly detectable on a Western blot (13-15). The choice of expression plasmid might be influenced by the nature of the target. If the target is expected to interfere with the endogeneous yeast metabolism, a lower expression might be beneficial. However if the expression of the fusion-protein needs to be assayed, the higher expression level is more convenient. Legrain *et al.* (15) used known interaction partners to study the influence of expression levels of the fusion proteins on their detection sensitivity in two-hybrid assays. Combinations of weak expressing plasmids revealed that detection of these interactions was roughly 50-100 times less sensitive. Therefore, an interaction assayed by use of these plasmids may sometimes escape detection. Important, however, is to notice that this observation was strongly dependent upon the protein-couples used. Some protein combinations were unaffected in sensitivity whereas others dropped only 10-fold. Another important but less documented parameter is the relative position of the various "cassettes" in the different vectors. A simple correlation between sensitivity and length of the

promoter fragment used is excluded since the use of other plasmids containing the same proteins under the same truncated promoter did not result in a loss of sensitivity (15).

Total expression level does not only depend on the promoter strength and the carbon source used, but also on the copy number of the plasmid. In most commonly used two-hybrid plasmids, the origin of replication is the so-called 2µm origin. The 2-µm circle plasmids are maintained stable and at high copy numbers (50-100 copies per cell) in yeast, and function solely for their own replication (16). It was shown that 2-µm DNA replication is similar to chromosomal DNA replication. 2-µm DNA is the only known example of a multiple-copy extrachromosomal DNA in which every molecule replicates in each cell cycle (17). It should be noted that in the context of a reverse hybrid system (see Reverse Hybrid System), the expression levels of the hybrid proteins should be maintained as low as possible, since every "background" interaction, which might occur more often at high protein expression levels, will kill the yeast. Therefore, the vectors used in these reverse hybrid systems make use of low-copy, so-called centromeric, expression plasmids. The difference in expression levels between high- and low-copy plasmids can be as much as twenty- or thirty fold. Also when using toxic proteins, the use of these centromeric vectors in forward screens can be taken into consideration.

Features of target/bait vectors

The pAS vectors also encode a hemaglutinin (HA), YPYDVPDYA, epitope tag in frame with the GAL4DB (18). This allows the protein to be visualized with commercially available anti-HA antibodies. There has been some controversy about the introduction of the HA tag since this results in a weak auto-activation when the empty vector is transformed and reporter gene activity is measured. However, when a protein is fused in frame, in general, the fusion protein looses the auto-activation properties. In principle, this could be used as an "in-frame control" but it has never been advertised in this way. It might however explain why in the new version of the pAS2, named pAS2-1, the HA tag is removed. The same HA is now introduced in pACT2, which is the successor of pGAD424. Other characteristics that might influence the choice of a vector are from a more practical nature. In all vector systems care must be taken to maintain the proper reading frame when creating the two-hybrid proteins. Compatibility of the multiple cloning site (MCS) in the DB and AD containing vector can be beneficial. Cloning in both vectors permits both verification of the correct length on Western blot and enables the establishment of a panel to score new targets in simple co-transformation or mating experiments.

Another "add-on" found in some of the commercially available vectors is the CYH2 gene. CYH2 encodes the L29 protein of the yeast ribosome. Cycloheximide (Chx), a drug which blocks polypeptide elongation during translation, prevents the growth of cells that contain the wild-type CYH2 gene. Chx resistance results from a single amino acid change in the CYH2 protein (19). When a wild type CYH2 containing plasmid is present in a mutant yeast strain resistant to Chx, it confers Chx sensitivity. Adding Chx to the medium results in a selection for yeast that has lost the CYH2 containing plasmid. This technique might be useful to select for only the library insert containing plasmid after screening (see Separating the two-hybrids). A potential danger is that due to leakage on the promoter of this CYH2 gene the yeast growth is retarded.

Most two-hybrid strains have a lesion in either *URA3, LEU2*, *HIS3*, *TRP1* and/or *ADE2* which allows selection for yeast cells that were transformed with plasmids that carry the corresponding gene by growth in the absence of the appropriate amino acids. The most widely used markers in yeast are genes encoding amino acid biosynthetic enzymes. However, in a recently introduced target vector (pHybLex/Zeo) zeocin is used as selection marker. Parent *et al.* (17) have compiled an extensive list of yeast cloning vectors. In addition, a list of the most commonly used plasmids in two-hybrid and their distinguishable features is listed in Table 1.

Table 1: Overview of the most commonly used two-hybrid vectors. The last column describes the origin of the promoter and the accession number in EMBL (AC)

Name	Selection-marker	Functional domain	promoter, AC
GAL4-based			
pMA424	*HIS3*	GAL4DB	Original vector, 12 kb
pGBT9	*TRP1*	GAL4DB	ADH1(truncated) AC: U07646
pAS1	*TRP1*	GAL4DB+HA	ADH1(full length), CYH2
pAS2	*TRP1*	GAL4DB+HA	ADH1(full length) CYH2 AC: U30496
pAS2-1	*TRP1*	GAL4DB	ADH1(full length) CYH2 AC:U30497
pGAD2F	*LEU2*	GAL4AD	Original vector, 13 kb
pGAD424	*LEU2*	GAL4AD	ADH1(truncated) AC:U07647
pGAD10	*LEU2*	GAL4AD	ADH1(truncated) AC:U13188
pGAD-GL	*LEU2*	GAL4AD	ADH1(truncated)
pGAD-GH	*LEU2*	GAL4AD	ADH1(full length)
pGAD1318	*LEU2*	GAL4AD	ADH1(full length)
pSE1107	*LEU2*	GAL4AD	
pSD-10	*URA3*	VP16AD	
pACT1	*LEU2*	GAL4AD	
pACT2	*LEU2*	GAL4AD+HA	ADH1(truncated), medium expression AC:U29899
LexA-based			
pBTM116	*TRP1*	LexA	ADH1(truncated)
pLexA	*HIS3*	LexA	ADH1(full length) = pEG202
pB42AD	*TRP1*	B42+ SV40 NLS + HA	GAL1 (full length), inducible promoter

			= pJG4-5
pHybLex/Zeo	*Zeocin*	LexA	ADH1(truncated)
pYESTrp	*TRP1*	V5 epitope + SV40 NLS + B42	GAL1 (full length), inducible promoter
pGilda	*HIS3*	LexA	GAL1 (full length), inducible promoter centromeric vector

Sequencing and Western Blot analysis

Before proceeding with actual screening it is advisable to check the chimeric cDNA by sequencing and to verify the actual expression of the fusion protein inside yeast. The latter can be achieved by "classical" SDS-PAGE and subsequent Western blot analysis as outlined in Protocol 1. Both anti-GAL4DB and anti-LexA antibodies are commercially available (Clontech #5399-1). This is preferably performed on the yeast strain used for screening the library to be sure that the target is properly expressed (20).

Protocol 1 SDS-PAGE and Western blot analysis to check for full-length expression of the target fusion protein in the GAL4 system

After transformation of the plasmids encoding GAL4DB fusion proteins into *Saccharomyces cerevisiae* strain HF7c and incubation on plates lacking the appropriate amino acids (e.g.Trp), single colonies are inoculated into 15 ml synthetic medium lacking the same amino acid(s). At an optical density (A_{600}) around 0.7, the culture is centrifuged at 2500 rpm for 5 min, the pellet is washed in distilled water and boiled for 3 min in 200 µl of 2 x Laemmli loading buffer. 50 µl of this sample is separated by 10% SDS-polyacrylamide gel electrophoresis and blotted onto a nitrocellulose membrane (Schleicher & Schuell, Dassel, F.R.G.). Detection of the expressed fusion proteins can be performed with polyclonal anti-yeast GAL4DB antibody and peroxidase-conjugated anti-rabbit antibody using ECL (Amersham Life Science, Amersham, U.K.).

Auto-activation

Since the two-hybrid system is based on reconstitution of a functional transcription factor, checking the auto-activation capacity of the target is crucial for the overall feasibility. Initiation of transcription, due to some latent activating activity is present in approximately 5% of all proteins and even more in randomly generated fragments (like in libraries). The use of a library fused to the DNA binding domain would result in a large amount of false positives illustrating that auto-activation can cause problems. Since DNA-binding to the upstream activating sequence (UAS_G) is more stringent, it provides a rational for making the libraries in the vector containing the AD.

A single transformation and reporter gene assay tells you whether your favorite protein (YFP) is able to induce reporter gene activity. What if YFP is auto activating? If it is only weak, e.g. only a minor background on the *HIS3* and not on *LacZ*, one could try to increase the amount of AT in the plates (see Pilot transformation in the GAL4 system). In the LexA system using less sensible reporter hosts or LacZ expressing plasmids containing variable DNA binding regions can genetically control the sensitivity. This ability to tune genetically the sensitivity of the reporter might be one of greatest advantages of the LexA system (see Table 4). If this tuning does not help, more radical approaches need to be applied. It is often

possible to delete a small region of a protein that activates transcription. Removal of this activation function while retaining other properties of the protein might enable the use of the target to screen libraries. Alternatively, instead of using a C-terminal fusion to the DB, an N-terminal fusion can be tried. Indeed normal two-hybrid methods detect interactions between two proteins fused at the C-termini of the DB and AD, respectively. This implicates that the N-terminus of none of these proteins is available for interaction. It was demonstrated, using constructs with reverted polarity, that such constructs give a specific interaction signal that is dramatically increased. Such constructs might lead to the identification of partners missed during classical two-hybrid screens and might at least in some cases provide a solution to the auto-activation problem (21).

In theory, another solution could be considered. If auto-activation is strong like in the case of transcription factors, it could be possible to apply a reverse approach. By using a toxic reporter gene and an auto-activating target under an inducible promoter, all yeast cells that are not repressed by an interaction, will die upon induction of the promoter. All yeast cells surviving such a screening procedure should contain a plasmid encoding an interaction partner of the auto-activating bait.

Nuclear localization: the repression assay

Another requirement, apart from correct folding of the fusion protein and impossibility to initiate autonomously transcription, is the ability of the fusion to be localized to the yeast nucleus. Indeed all transcription events are dedicated nuclear processes. Specific binding of GAL4 to the $(UAS)_G$ is conferred by a zinc-cluster motif located within the N-terminal 64 residues (5,22,23). GAL4 binds as a dimer to DNA, which is mediated at least in part by an α-helical region within residues 65-94 which form a coiled-coil interaction (23-24). GAL4DB also has its own nuclear localization signal (NLS) within its N-terminal 74 residues (25-27). A positive control can be used to check the nuclear localization in the GAL4 system. Unless, in the case of auto-activation, where apart from the problems described above, the nuclear localization problem is trivial.

Since LexA is a bacterial protein it contains no NLS. Therefore, the NLS of SV40 large T is fused in frame ensuring nuclear localization. The nuclear localization of LexA fusions can be assayed by a so-called "Repression" or "Blocking assay". The repression assay is based on the observation that LexA and non-activating LexA fusions can repress transcription of a yeast reporter gene that has LexA operators positioned between the TATA and upstream activating sequence (UAS) (28). *LacZ* expression is induced by galactose and is detectable in the presence of glucose. A transcriptional inert LexA fusion that binds to the operator located between the UAS_G and the TATA box is able to repress or block the *LacZ* expression, a clear indication that the LexA fusion is properly located to the yeast nucleus. The reporter plasmid pJK101 is used to perform this repression analysis. Since repression is never complete, a positive and negative control is needed to calibrate the repression capability of the protein assayed.

One could argue that nuclear localization of the target or prey as such is not strictly required. Indeed, if interaction occurs in the cytoplasm, both interacting proteins, of which only one contains a NLS, might be targeted to the nucleus. However, this might only be true for strong interactions that are not always the most interesting.

Positive control

The best way to avoid most problems concerning bait construction is to clone a known interactor in the opposite vector and to assay the interaction in a two-hybrid system. If this results in specific reporter gene activation, your fusion protein is correctly folded and properly targeted to the nucleus. However, if this test fails, finding out why, might be cumbersome.

Library choice

It is always a good idea to start with a library prepared from a tissue in which the target protein is known to be biologically relevant. To screen a mammalian cDNA library until saturation, more than 5-10 x 10^6 yeast transformants need to be screened. Both oligo-dT and random primed libraries are used. Important quality parameters are the number of independent clones before and after amplification, the number of clones that contain an insert and the mean insert length. Having multiple cloning sites available in the library plasmid is beneficial to subclone in later stages the insert for additional interaction controls outside the two-hybrid system. The most commonly used plasmids for library construction are listed in Table 1.

Important is also the relative strength of the activation domains, indicating their ability to initiate transcription. Both VP16 and the AD of GAL4 are known to be strong activators making the system more sensitive. This might be needed for the detection of weak interactions but results inevitable in higher backgrounds. Therefore the use of B42AD, a random fragment that was isolated for its intermediate transactivation capability, might be beneficial in setting up a screening.

For the library plasmid there is a major difference in promoter between the LexA and GAL4 system. While in the most commonly used plasmids of the GAL4 system, fusion proteins are weakly and constitutively expressed, they are cloned behind a stronger but inducible promoter in the LexA system. In the latest version of the library plasmid used in the GAL4 system, pACT2, a truncated weakly expressing version of ADH1 is used. However, since this promoter is adjacent to a section of pBR322 that acts as a transcriptional enhancer in yeast, a medium expression level is obtained.

Inducible expression has the advantage that there is less opportunity for AD fusions proteins to have a toxic effect on the yeast host, which could also result in the elimination of this protein from the pool of potentially interacting proteins. However, the experimental protocol is longer in the case of induction since transformation efficiency drops dramatically if the transformations are directly selected on all auxotrophic markers and on a carbon source that induces expression of the library fusion protein. Therefore, one typically selects for all plasmids before induction. Although this procedure is longer, it might be beneficial since only part of the original transformation mix needs to be induced, enabling means to perform screening in shifts and having back-ups in case of contamination. It is possible to combine vectors as long as they are compatible with each other with regards to the selectable markers and the UAS before the reporters (see Table 2) (29).

Table 2: Comparison of the LexA and GAL4 Yeast Two-Hybrid Systems

	DB vector selection marker	AD vector selection marker	Chromosomal reporter gene(s)	Plasmid reporter gene
GAL4 Systems	*TRP1*	*LEU2*	*HIS3, LacZ*	(none)
LexA Systems	*HIS3*	*TRP1*	*LEU2*	*LacZ*

Choose a yeast strain as a host

Reporter gene(s) and resulting sensitivity

The upstream activating regions and TATA regions are the basic building blocks of yeast promoters. The initiation of gene transcription in yeast, as in other organisms, is achieved by several molecular mechanisms working in concert. All yeast genes are preceded by a region containing a TATA box. Many genes are also associated with cis-acting transcription elements and sequences to which transcription factors and other trans-acting regulatory proteins bind and affect transcription levels. The term promoter usually refers to both the TATA box and these associated cis-regulatory elements. Gene regulation in yeast involves cis regulatory elements that are relatively closely associated with the TATA box. The most common type of cis-acting transcription elements in yeast are upstream activating sequences (UAS). UAS sequences are recognized by specific transcriptional activators that enhance transcription. The enhancing function of yeast UASs is generally independent of the orientation but it is sensitive to distance effects, if moved more than a few hundred base pairs from the TATA region. There may be multiple copies of a UAS upstream of a yeast coding region. In yeast, the genes required for galactose metabolism are controlled by two regulatory proteins, GAL4 and GAL80, as well as by the carbon source in the medium (30). When galactose is present, the GAL4 protein binds to the GAL-responsive elements within the UAS of at least 20 known galactose-responsive genes (including GAL1). In the absence of galactose, GAL80 binds to GAL4 and this interaction blocks transcriptional activation. Furthermore, in the presence of glucose, transcription of galactose genes is immediately repressed. The 17-mer consensus sequence, referred to as UAS_G, functions in an additive fashion. Indeed multiple sites lead to higher transcription levels than a single site (31). This explains why the number of UAS_G located before the reporter determines the sensitivity of the interactions that can be assayed.

To avoid interference by endogeneous GAL4 and GAL80, proteins the yeast host strains used in the GAL4 based two-hybrid system must carry deletions of the GAL4 and GAL80 genes. Due to the deletion of these two genes, the yeast cells grow more slowly as compared to yeast containing the wild type version of these genes. The use of bacterial LexA circumvents this disadvantage.

Reporter genes can be integrated into the genome or reside on a plasmid. The inconvenience of having another plasmid containing the reporter gene, and the need for an additional auxotrophic marker is compensated by several advantages. Indeed, one of the major advantages of the LexA system, where the *LacZ* reporter gene is present on a high copy-number plasmid, is that weak signals are more efficiently amplified than in the GAL4 system, which makes it possible to assay β-Galactosidase activity directly on

the selection plate by including X-GAL in the medium. This avoids tedious replica and/or filter lift assays.

Reporter strains in LexA and GAL4

In the GAL4-based MATCHMAKER two-hybrid system, either the intact GAL1 UAS, which contains four GAL4-binding sites, or an artificially constructed UAS consisting of three copies of the 17-mer consensus binding sequence is used.

Table 3: Survey of the most commonly used yeast strains, used reporter genes and the constitution (origin) of their promoter in the GAL4 based two-hybrid system.

	Reporter genes	UAS regulated by + origin of UAS	Expression	
			uninduced	induced
H7Fc	*LacZ,*	GAL4, 3 x $UAS_{G\,17\text{-mer}}$	-	low
	HIS3	GAL4, GAL1 (= 4 x $UAS_{G\,17\text{-mer}}$)	-	high
YRG-2	*LacZ,*	GAL4, 3 x $UAS_{G\,17\text{-mer}}$	-	low
	HIS3	GAL4, GAL1 (= 4 x $UAS_{G\,17\text{-mer}}$)	-	high
SFY526	*LacZ*	GAL4, GAL1 (= 4 x $UAS_{G\,17\text{-mer}}$)	-	high
Y187	*LacZ*	GAL4, GAL1 (= 4 x $UAS_{G\,17\text{-mer}}$)	-	high
Y190	*LacZ,*	GAL4, GAL1 (= 4 x $UAS_{G\,17\text{-mer}}$)	-	high
	HIS3	GAL4, GAL1 (= 4 x $UAS_{G\,17\text{-mer}}$)	low	high
CG-1945	*LacZ,*	GAL4, 3 x $UAS_{G\,17\text{-mer}}$	-	low
	HIS3	GAL4, GAL1 (= 4 x $UAS_{G\,17\text{-mer}}$)	very low	high
L40	*HIS3,*	4 x LexA op		
	LacZ	8 x LexA op		

The relative sensitivity of both reporter genes used in the GAL4 based approach, respectively *LacZ* and *HIS3*, depends entirely on the constitution of the promoter. In the GAL4 system the *LacZ* is more stringent, meaning less likely to give false positives, than the *HIS3* gene. This was elegantly demonstrated in a recently described reverse three-hybrid approach. Here the expression of Raf from a third promoter could titrate out the Ras-DB fusion, keeping it from interacting with Raf-AD fusion. This reverse experiment resulted in the expected *LacZ* negative phenotype but not in a *HIS3* negative phenotype, reflecting the higher sensitivity of the histidine reporter gene as compared to the β-galactosidase reporter gene in HF7c (32).

In LexA-based two-hybrid systems, the DB is provided by the entire prokaryotic LexA protein. LexA normally functions as a repressor of SOS genes in *E. coli* by binding LexA operator sequences that are an integral part of the promoter (33). When used in the yeast two-hybrid system, the LexA protein does not act as a repressor because the LexA operators are integrated upstream of the minimal promoter and coding region of the *LEU2* reporter gene. Expression of the latter in EGY48 (Erica Golemis Yeast) is

under the control of six copies of the LexA operator (op) sequence and a minimal *LEU2* promoter. In the *LacZ* reporter plasmids, the *LacZ* reporter expression is under the control of 1-8 copies of the LexA operator and the minimal GAL1 promoter (34).

As all of the GAL1 UAS sequences have been removed from the *LacZ* reporter plasmids. This promoter is not regulated by glucose or galactose (35).

Table 4: Tuning sensitivity in the LexA system by using different host strains and reporter plasmids

EGY reporter strains listed by decreasing sensitivity

EGY 48		colE1 LexA operators = 6 LexA binding sites + TATA + *LEU2*
EGY 194		2 colE1 LexA operators = 4 LexA binding sites + TATA+ *LEU2*
EGY 188		1 colE1 LexA operators = 2 LexA binding sites + TATA + *LEU2*
EGY 40		contains no *LEU2* reporter gene

LacZ **reporter plasmids listed by decreasing sensitivity**

pSH18-34		4 colE1 operators, 8 LexA binding sites + TATA + GAL1-*LacZ*
pJK103		1 colE1 operators, 2 LexA binding sites + TATA + GAL1-*LacZ*
pRB1840		1 recA LexA operator + TATA + GAL1-*LacZ*

Finally there are many "hybrid two-hybrid systems" that combine GAL4 and LexA based systems. The reporter strain L40 permits the use of the same reporter genes as in the GAL4 system but under control of LexA operators. This permits the use of the plethora of libraries commercially available for the GAL4 system in a LexA related system.

Why use multiple reporter genes

The most significant improvement that accounts largely for the success of the two-hybrid system has been the conversion from a color-based assay to a nutritional selection via inclusion of the *HIS3* gene as a transcription reporter in addition to *LacZ*. This allows much larger libraries to be efficiently screened on far fewer plates. Apart from a higher plating density, the use of multiple reporter genes has many additional advantages. The use of two reporter genes under the control of a similar promoter but at distinct places of the genome can eliminate many false positives. Continuous histidine selection throughout subsequent procedures eliminates irrelevant library plasmids that may have been cotransformed with the actual plasmid of interest that initiated reporter gene activity. Moreover, plasmids that are toxic to the yeast can undergo rearrangements or mutations that eliminate expression of the activation domain hybrid. Since loss of the activation domain results in loss of reporter genes, these yeast cells do not survive under auxotrophic selection. The β-galactosidase activity of transformants under *HIS3* selection is generally enhanced, probably because the transformants are forced to produce the hybrid proteins in quantities that enable the *HIS3* gene to be expressed at levels that permit growth.

Pilot transformation in the GAL4 system

Since *HIS3* has a leaky expression in many yeast strains, it is recommended to inhibit the basal expression of this gene with 3-amino-triazole (AT), a known inhibitor of the *HIS3* gene product. Therefore, before starting a large-scale transformation procedure it could be very informative to perform a pilot transformation that enables titration of the optimal amount of AT needed. Using too much AT will result in a loss of weak interactions, whereas the use of no or too small amounts AT results in

unworkable numbers of false positives. The exact amount is largely dependent on the yeast strain and upon the target used. The reason why Y190 requires more AT to bring down the background expression of *HIS3* as compared to HF7c is related to the origin of the TATA sequence. In most cases the TATA box of GAL1 is used in front of *HIS3* (and the one of CYC1 in front of *LacZ*) except in Y190 where the TATA box of *HIS3* is used.

Typically a couple of strong, weak and intermediate interacting proteins are transformed together with the target and a dilution of library cDNA. These are plated on a logarithmic gradient of AT (0,5,10,25 and 50 mM of AT). This generally results in growth as shown in Figure 3. Here the optimal AT would be, enough to suppress background and not too much to be able to detect the weak interactions.
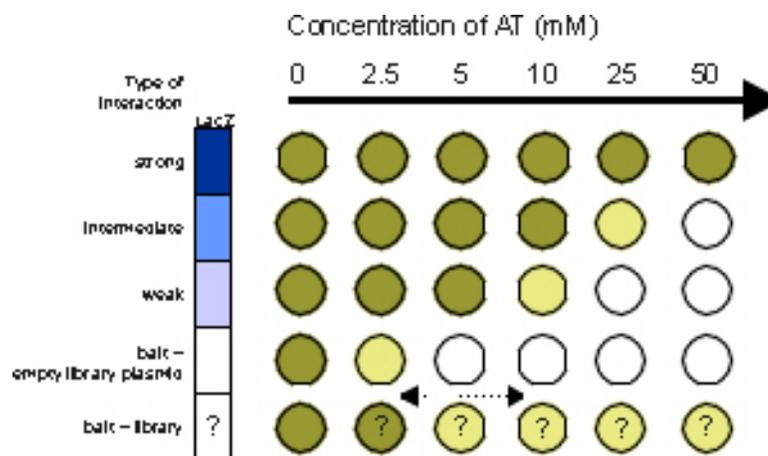
Figure 3: Determination of optimal AT concentration in a pilot transformation using strong, weak and intermediate interactors as positive and negative controls. Cells are plated out on a AT gradient. The optimal AT concentration (dotted double arrow) is this which elimates false positives (e.g. formed in case of bait + empty library plasmid but which still allows to detect weak interactions).

Library Transformation

Transformation efficiency

Transformation efficiency is often the bottle neck when performing two-hybrid experiments. The basic idea is to saturate the screening to obtain as many transformants as possible, and to avoid transformation of a single yeast cell with multiple library plasmids, which can seriously complicate subsequent analysis, the lowest amount of DNA needs to be used. A maximum number of transformants with a minimal amount of DNA requires a high transformation efficiency.

There are several commonly used methods to introduce plasmid DNA into yeast. These include the spheroblast method, electroporation and the lithium acetate method (LiAc) (36-38). Although the spheroblast method results in high transformation efficiencies, it is too much time-consuming for large-scale experiments like library screenings. In general, electroporation offers the highest transformation efficiency, and is especially useful when small quantities of DNA are transformed. Unlike the lithium acetate (LiAc) based method, electroporation saturates at low DNA levels, restricting its use in a two-hybrid approach. Taken together, all two-hybrid protocols make use of the LiAc method, which is a quick method to introduce foreign DNA into the yeast in a large scale experiment.

In the LiAc transformation method, yeast competent cells are prepared and resuspended in a LiAc solution together with the plasmid DNA to be transformed and an excess of carrier DNA. Polyethylene

glycol (PEG) with the appropriate amount of LiAc is then added to the mixture of DNA and yeast, and the mixture is incubated at 30°C. Before or after the incubation, DMSO is added and the cells are heat shocked, which allows the DNA to enter the cells. The cells are then plated on the appropriate medium to select for transformants containing the introduced plasmids (Protocol 2).

Many versions of this LiAc based protocol are available, probably illustrating that many of the different steps can be subject to change and are not strictly required for the final outcome. Like every transformation procedure, the exact mechanism of DNA uptake is not known, which makes all protocols highly empirical.

Of crucial importance in every strategy are the yeast cells used. Since growth on a master plate does not give any selection criteria for "transformability", there is always a risk of selecting for yeast cells that are more difficult to make competent for transformation. Best practice is to keep a stock of (competent) yeast cells known to be good for transformation. Even by using yeast cells known to have worked in previous experiments, it is advisable to perform several transformations in parallel. A second parameter found to be important is the growth phase. With respect to this, logarithmic growing cells were shown to be the best. Maximal transformation efficiency in each of the procedures depends on finding the optimal cell density. The cell density can be "controlled" by making appropriate dilutions and determining the optical density at 600 nm at different times. However, the best method is to look through the microscope and count the percentage budding yeast cells and calibrate this percentage to the number of transformants obtained. Important to notice is that the total volume of cultures is equally important. Too many yeast cells in a tranformation mix will significantly drop the efficiency and will give difficulties to spread on the plates. Too dense plating might result in higher background since more slowly growing yeast cells can grow on the corpses of others. Also crucial in two-hybrid experiments is the total amount of DNA used. The use of too much DNA in the transformation-mix might be fatal, not only because most protocols will saturate at a given amount of DNA, but it will increase the chance of multiple library plasmids to enter the same host cell. This will make subsequent analysis very hard. The quality of plasmid DNA is said to be less crucial since impurities resulting from the preparation method might function as carrier and consequently increase transformation efficiency.

Another critical parameter is the quality of carrier DNA, which must be of high molecular weight and completely denatured. Denaturation can be achieved by repeating the cycle of boiling and chilling immediately before use. Achieving the proper DNA size distribution upon sonication is probably the most difficult. Instead of running agarose gels to check the size distributions, a more direct way is to use several batches in a transformation experiment and retain the one that gives the highest transformation efficiency regardless of its size distribution. Competence of DNA uptake at the level of the cell wall is induced by treatment with both lithium ions and PEG (3350-4000). Some people claim that the quality of a PEG stock buffer decreases over time and might result in low transformation efficiency.

To ensure that all plasmids are present and the target is correctly expressed, a small aliquot of the competent yeast cells can be used to transform a positive control. One of best resources on transformation of yeast can be found on the webpage of the Gietz Lab (39).

Protocol 2 Overview of a typical library screening protocol

1. Start with a master plate of yeast transformed with the target DNA binding domain-fusion protein (don't keep master plates longer than 2 weeks and store them at 4°C)
2. Inoculate four 20-50 ml cultures of a single large colony, in SD medium that selects for the target protein (e.g. -Trp)
3. Grow for 36-48 hours (until culture becomes "milky")
4. Transfer each of these cultures to a 200 ml culture
5. Grow overnight
6. In the morning, transfer entire 200 ml into 400 ml of YPD (OD should be around 0.2 in a total volume of 600 ml)
7. Grow to an OD_{600} of about 0.7-0.75 (about 4 hours)
8. Transfer cells into 250 ml bottles (in shifts). Spin at 2200 rpm (1000 X g) for 5 min. Keep the four pools separated throughout the protocol
9. Remove supernatant and resuspend in water. Spin 2200 rpm for 5 min.
10. Remove supernatant and resuspend cells, 4 batches, with about 10 ml of 1 X TE/LiAc (made fresh) per bottle.
11. Bring to a 50 ml tube. Spin 2200 rpm for 5 min
12. Remove supernatant. Bring final volume up to 2 ml with 1 X TE/LiAc.
13. Do transformations in 4 x 10 eppendorfs containing:
    200 μl yeast cells
    8 μg library DNA
    30 μl salmon sperm DNA (10 mg/ml)
    7 μl 10 X LiAc/TE
    30 μl DMSO
    1.2 ml PEG/LiAc (freshly made !)
14. Vortex and flick to assure DNA is well distributed
15. Incubate at 30°C for 30 min while gently shaking in thermomixer, flick every 10'
16. Incubate at 42°C in thermomixer
17. Chill cells on ice (1') and spin at 2600 rpm for 1'
18. Resuspend in 300 μl 1 X TE
19. From each eppendorf, plate an equal amount (133 μl per plate) onto 3 plates lacking the appropriate amino acids: total of 4 x 30 large plates
20. Plate dilutions (1/1000, 1/10000 & 1/100000) on -Trp & -Leu plates to assay transformation efficiency, for each of the 4 competent yeast batches
21. Put plates at 30° C for 1-2 weeks

Simultaneous versus sequential transformations

The LiAc methods typically result in transformation efficiencies of 10^5 transformants per μg of DNA when using a single type of plasmid. When the yeast is simultaneously cotransformed with two plasmids having different selection markers, the efficiency is usually an order of magnitude lower due to the lower probability that a particular yeast will take up both plasmids. Yeast, unlike bacteria, can support the

propagation of more than one plasmid having the same replication origin. Although simultaneous transformation is less efficient, it might be beneficial to use this strategy. Apart from being easier, simultaneous transformation avoids a growth phase after the first transformation that might result in counterselection. Indeed, if the expressed protein is toxic or interferes with normal yeast metabolism, arising clones might have spontaneous deletions in the first plasmid which confer a growth advantage. Nevertheless, in order to saturate screening, it is advisable to perform sequential transformations which mostly results in a few million transformants.

In the inducible LexA system it is advisable to select first for transformants, (sequentially induced most of the time), and in a second step to induce the expression of the cDNA library insert. This two-step approach results in a uniform increase in the number of cells carrying each library plasmid. Before harvesting, each cell transformed with the library is allowed to form a colony of approximately the same size. One advantage of this approach is that the number of transformed yeast cells is amplified before the synthesis of cDNA encoded proteins. This ensures that yeast cells containing toxic or mildly toxic cDNA encoded proteins will not be depleted from the population.

Picking positive colonies

In the GAL4 system, transformants are plated out on plates that are selective for the two plasmids and the additional auxotrophic reporter gene *HIS3*. Colonies able to grow on this medium are in principle inducing *HIS3* as a result of protein-protein interactions that reconstitute a functional GAL4. The size of the colony might reflect the strength of the interaction. However, there is no reason to assume that weak interactions have less biological significance than strong ones. So, when picking colonies, care must be taken to evaluate the size of each colony to the background and not preferentially picking the largest colonies. The same holds for weak and strong *LacZ*$^+$ after assaying for β-galactosidase activity.

In general after 5-6 days the first colonies can be picked and transferred in e.g. 96 well plates. After 2-3 hours they can be replica plated onto plates lacking the appropriate amino acids and the titrated amount of AT (master plate), and at the same time onto plates covered with a filter membrane. Since colonies grow through the filter, it facilitates the lifting procedure and subsequent lysis in liquid nitrogen. It is a good idea to include here a positive (preferably a strong, intermediate and weak interaction) and a negative control. If not too many clones need be assayed, the 96 well procedure can be replaced by replica striking directly on a master and filter containing plate. In order to pick also slower growing colonies, this procedure is repeated after 12-15 days, largely depending upon the number of colonies obtained.

When performing a β-galactosidase test, care must be taken to the growth phase of the yeast. Again, fresh logarithmic growing yeast cells will give the best results. Strong interactions should be visible after 2 hours, intermediate ones might take 6 hours while weak interactions might need to be incubated overnight. The incorporation of both the positive and negative controls on the same filter is the only relevant reference point. It should be noted that not all *HIS3*$^+$ are *LacZ*$^+$. Due to the earlier described different sensitivities of both the *HIS3* an *LacZ* phenotypes it is usually found that *HIS3*$^+$ and *LacZ*$^+$ cells contain the most biologically relevant interaction, however, it cannot be excluded that *HIS3*$^+$ and *LacZ* cells are of potential interest.

Rescue of library plasmid

Preparation of DNA from yeast

After the difficulties to get the DNA inside the yeast, the problem of getting it out is easier to solve.. Many protocols yield little, but nevertheloss a sufficient amount of DNA that can be transformed to bacteria where it can be amplified for further analysis. One of the pitfalls is that apparently most DNA precipitates prepared from yeast contain debris of the yeast outer membranes that appears to be toxic for most competent bacteria. By diluting this debris below toxic levels, the equivalent drop in DNA concentration requires a very efficient bacterial transformation.

Separating the two-hybrids

The positive yeast host cells contains two different two-hybrid plasmids, one encoding the bait and another containing the library insert. There are several procedures to enrich for the plasmid containing the library insert. One option makes use of the CYH2 gene on the target plasmid, that confers cycloheximide sensitivity to the yeast host. By adding cycloheximide to the medium, it is possible to select for cells that have lost this plasmid. Another protocol makes use of bacteria (HB101) that are unable to grow on media lacking leucine. Only by transformation of a plasmid containing the *LEU2* gene, which happens to be the one containing the library insert, these bacteria can be rescued. It should be noted that this procedure is not trivial since HB101 grows very slowly (3 days) and DNA prepared from this strain can be hard to sequence. Alternatively, yeast DNA can be prepared and directly, without any selection procedure, transformed into bacteria (MC1061) (Protocol 3). In theory 50% of the transformed bacteria contain the library plasmid. A simple and highly efficient procedure for rescuing autonomous plasmids from yeast was described by Robzyk and Kassir (75). In the LexA system, the plasmid containing the target and the one containing the library insert contain different *E. coli* origins enabling an enrichment for the library containing plasmid by simple *E. coli* transformation.

Protocol 3 Preparation of DNA from yeast cells that can be re-transformed to chemical competent bacteria

1. Use a freshly inoculated 5-10 ml culture grown overnight
2. Centrifuge 5', 2800 rpm
3. Resuspend in 5 ml SCE (1 M Sorbitol, 0.1 M Na$_3$Citrate, 10 mM EDTA, pH 6.5-7, autoclaved)
4. Centrifuge 5', 2800 rpm
5. Resuspend in 1 ml SCEM (100 ml SCE + 214 µl β–mercaptoethanol) + lyticase, transfer into 1.5 ml eppendorf
6. Incubate 1 hour at 37 °C
7. Centrifuge 5' at 1500 rpm
8. Remove SN carefully
9. Resuspend in 400 µl Yeast Extraction Buffer (0.05 M Tris.HCl pH 7.5, 0.025 M EDTA, 0.25 M NaCl, 1% SDS, autoclaved)

10.Homogenize
11.Incubate 15' at 65 °C
12.Phenolisation using phenol/chloroform/isoamylalcohol until the interface is clear
13.Apply one extra phenolisation step using pure chloroform
14.Precipitate DNA by adding 1/10 of total volume NaAc and 2.5 volumes of icecold EtOH
15.Centrifuge 15'-30' at 14000 rpm
16.Wash in 750 µl 70% EtOH and centrifuge 15' at 14000 rpm
17.Dry 10' in speedvac
18.Dissolve in 100 µl H_2O
19.Transform 20 µl of this mixture to chemically competent bacteria

Insert analysis—Assay for specificity in two-hybrid

Depending on the quantity of positives obtained, several approaches can be taken. In the case of a very limited set (<10), a new transformation with positive and negative controls (including empty vectors) to assay the specificity of the interaction can be performed first. Between 50-100 positives, direct sequence analysis might be considered. If more than 100 positives, they are best analyzed by colony hybridization to identify identical inserts. Control transformations can then be done with fewer candidates. Insert analysis might also include restriction site analysis, to identify insert families. However, new protocols enable screening for recombinant plasmids by using PCR methods directly onto the yeast cells (40).

Much information can be derived from sequence data. By running the appropriate algorithms on the obtained sequences, one is able to easily classify inserts into two main groups; the ones of potential interest, and the others. The latter category includes all known genes that are not in frame with the activation domain and those with unknown genes that show a polyAAA tail indicating that the gene is probably inverted. Other genes that might be omitted from further analysis do not encode proteins but RNA or retroposons. Some proteins can also be excluded if they are part of a list of known false positives that are frequently found, independently of the target used.

Table 5: Overview of most commonly encountered false positives. Adapted from the WWW-site of E. Golemis (44-45)

Proteins	Found as false positives	Believed to date to be real interactors
hsp's	16	5
ribosomal proteins	14	1
cytochrome oxidase	5	-
mitochondrial proteins	3	1
proteasome subunits	4	3
ferritin	4	-
tRNA synthase	3	-
collagen-related proteins	3	-
Zn finger proteins	3	4
vimentin	2	-

In general a second more interesting category persists in the algorithmic screening, containing both known and unknown genes. Although the choice of which interactions to characterize further is always arbitrary, some guidelines can be used. For *known proteins*, the potential biological relevance of the interaction is the most important criteria. If several clones containing portions of the same proteins are found, one can try to establish a minimum interaction domain. By using physicochemical predictions or three-dimensional data, the nature of interaction can be evaluated *in silico*, but remains to be validated by additional biological experiments.

For *new proteins* that show no homology with any known sequence in DNA nor protein databases, the above selection becomes less trivial. The redundancy at which the interaction is found might be a clue for the importance of the interaction. Noteworthy is that this redundancy is not related to the strength of the interaction. Even with new proteins, mostly several EST that contain parts of the coding sequence can be found. By performing extensive EST searches, in cycle order, and using multiple sequence alignment algorithms, one can sometimes reconstitute full-length cDNA clones. Using these sequence data, one can perform many different algorithms to predict subcellular localization, transmembrane domains, conserved functional patterns, etc. Of course, these predictions are never "bullet-proof" evidence, but can be the starting point for more conclusive experiments. Apart from the false positives that often arise, false negatives can never be excluded. Indeed a negative result in a two-hybrid system does not necessarily mean that the two proteins tested do not interact under all conditions. Several explanations can be given for lack of interaction (see Why not use the two-hybrid system?).

To assay the specificity of a two-hybrid interaction, a series of additional transformations is needed. In general, these include cotransformation of the found bait with empty vector or with a vector encoding an irrelevant protein. By using mating assays, these specificity tests can be performed in a very short period of time. Mating takes advantage of the fact that haploid cells of the opposite mating type will fuse to form diploids when brought into contact with each other (41). In the mating assay, the target is expressed in one yeast strain and the activation domain fusion is expressed in a second strain of opposite mating type. Using pre-transformed panels of yeast cells one can easily check lots of combinations that would otherwise require long cotransfection experiments (42-44).

Take the interaction outside the yeast system

Confirmation of interaction

In general, people like to see a confirmation of the interaction outside the yeast cell. Therefore several assays are available, such as *in vitro* "pull-down" assays and co-immunoprecipitation. Most of these assays use epitope tagged versions of the proteins involved. Alternatively a co-localization experiment can be used to validate the two-hybrid interaction couple. In most cases, interactions found in two-hybrid can be confirmed although it may take time to find the best conditions since most of the protocols include empirical steps that need to be optimized for every interacting couple. Probably the most conclusive experiments are performed by co-immunoprecipitating endogeneous, and thus physiological amounts of, proteins. A new vector, pWITCH, was designed to facilitate the characterization of proteins encoded by novel cDNAs that were isolated using either a one- or two-hybrid assay. Expression of directionally cloned cDNAs is obtained *in vivo* in *Saccharomyces cerevisiae* from the inducible GAL10

promoter and *in vitro* from the T7 promoter. In addition, translation of the cloned cDNAs results in proteins that are tagged *in vitro* with a specific epitope tag, and *in vivo* with both the epitope tag and the VP16 transcription activation domain (46-47).

Functional assays to test relevance of interaction

One of the most challenging and most difficult tasks is to determine the actual function of the interaction. The way this is achieved is largely dependent upon the signal transduction pathway in which the interaction is thought to be functionally relevant. In most cases the effect of inhibitors, mutants or dominant negative versions of the involved proteins can be evaluated. Using two-hybrid technology, a very detailed but time-consuming dissection of signal transduction can be made by using the so-called interaction suppression approach (see Interaction suppression). In almost all cases, a phenotype that can be assayed is required. It is said that the two-hybrid is an easy tool to generate many new hypotheses. The validation of these becomes in many cases the next time-limiting step.

NEW DEVELOPMENTS

Improvements

Refinements of vectors for use in the yeast two-hybrid system are reviewed by (48). A new host strain was developed that is extremely sensitive to weak interactions and eliminates nearly all false positives using simple plate assays. This yeast strain contains three reporter genes: *HIS3*, *ADE2* and *LacZ*, each under control of a different promoter (GAL1, GAL2, and GAL7 respectively) that respond to the same activator GAL4 (76). In another system, the target protein is expressed as a fusion with the DB domain of the human estrogen receptor (ER) in a yeast strain containing an integrated *URA3* reporter gene driven by one or three ER response elements (*EREs*). In contrast to the *HIS3*- and *LEU2*-based systems, it can be assayed quantitatively, by determining the activity of the reporter gene product orotidine-5'-monophosphate decarboxylase in cell-free extracts, and enables a negative selection scheme by using 5-fluoroorotic acid (see Reverse Hybrid System) (49). An important advantage of this system is the use of a target promoter that drives *URA3* reporter gene expression, which is different from those commonly used in two-hybrid systems (mostly the *GAL1* promoter). Combination of this *URA3/ER* with *HIS3/GAL4* or *LEU2/LexA* might provide elegant strategies to eliminate false positives. An additional review on progress and variations in two-hybrid and three-hybrid technologies is available (50).

Reverse Hybrid System

The two-hybrid system does not allow genetic selection of events responsible for dissociation of particular interactions, and hence, its use in the characterization and manipulation of protein-protein interactions has been limited. All reverse systems make use of yeast strains in which expression of interacting hybrid proteins increases the expression of a counterselectable marker that is toxic under particular conditions. Under these conditions, dissociation of an interaction provides a selective advantage, thereby facilitating detection: a few growing yeast colonies in which hybrids fail to interact can be identified among millions of non-growing colonies expressing interacting proteins. Several

reverse hybrid systems are proposed. The first reverse two-hybrid system uses a yeast strain that is resistant to cycloheximide due to the presence of a mutant *CYH2* gene. This strain also contains the wild-type CYH2 allele under the transcriptional control of the *GAL1* promoter. Expression of the wild-type GAL4 protein is sufficient to restore growth sensitivity to cycloheximide. Growth sensitivity towards cycloheximide is also restored by the co-expression of the avian c-Rel protein and its IκB-α counterpart, p40, as GAL4 fusion proteins. Restoration of growth sensitivity towards cycloheximide requires the association of c-Rel and p40 at the *GAL1* promoter and correlates with the ability of the c-Rel/p40 interaction to activate expression from the *GAL1* promoter (51).

Another reverse hybrid system makes use of the most widely used counter-selectable marker in yeast genetics, *URA3*, which encodes orotidine-5'-phosphate decarboxylase, an enzyme required for the biosynthesis of uracil. Yeast cells that contain wild-type *URA3*, either on a plasmid or integrated in the genome, grow on media lacking uracil (*URA3*+ phenotype). However, the *URA3*-encoded decarboxylase can also catalyze the conversion of a non-toxic analogue, 5-fluoroorotic acid (FOA) into a toxic product, 5-fuorouracil (52).
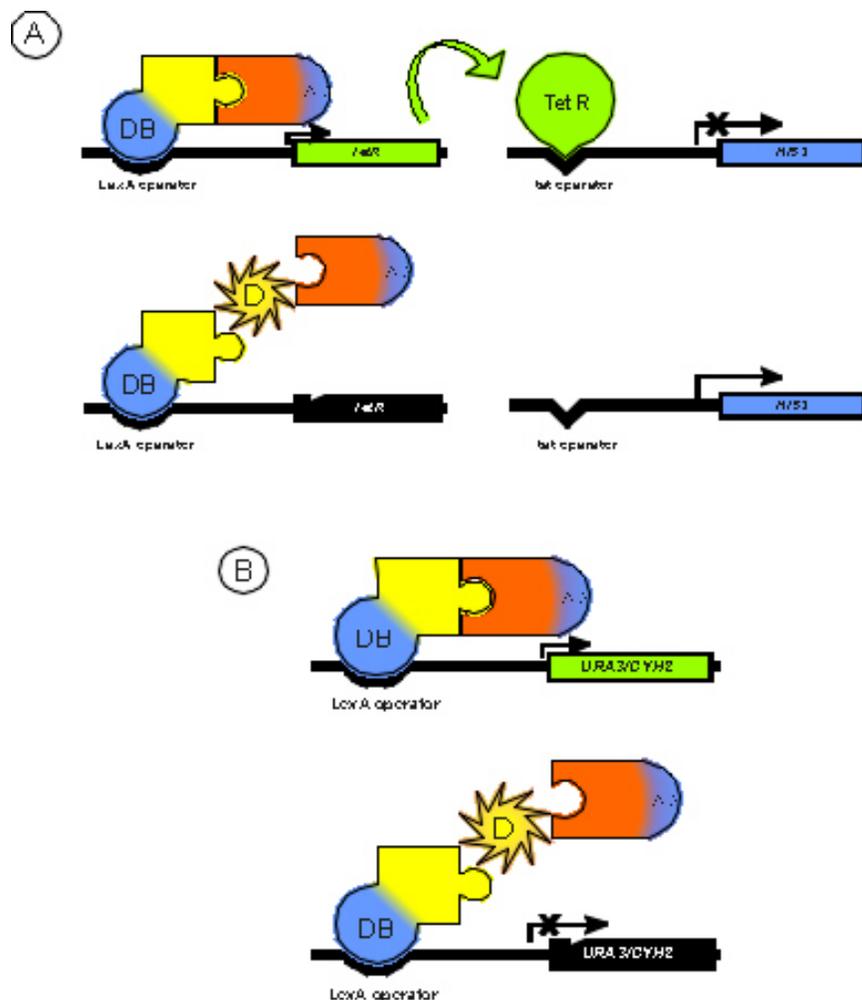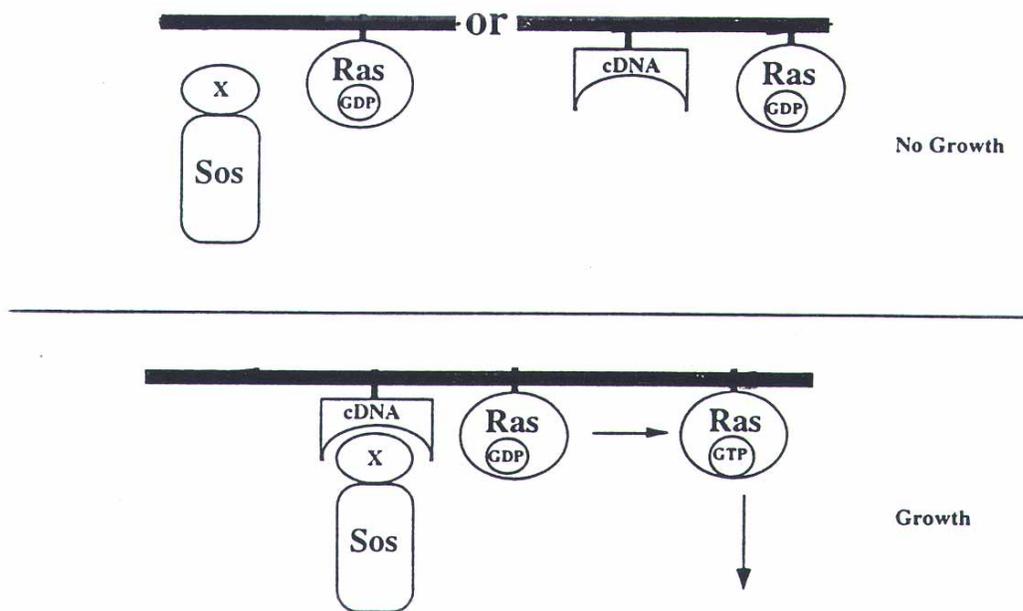

Figure 4: Principle of the reverse two-hybrid system. (A) the "split-hybrid" system is based on the E. coli TN10-encoded tet-repressor (TetR)/operator system: interaction of the hybrid proteins X and Y activates the transcription of a TetR; the TetR protein then represses HIS3 expression, resulting in a no-growth phenotype on plates without histidine in the growth medium (53). Abrogation of the interaction, either by mutating Y or by introducing a dissociator (D), shuts down TetR expression, leads to HIS3 expression and results in growth on selective plates. **(B)** The reverse two-hybrid system takes advantage of counterselectable yeast reporter genes (51,54-55). URA3 expression sensitizes yeast cells to 5-fluoorotic acid (5-Foa), CYH2 expression to cycloheximide. Abrogation of the interaction enables the cells to grow in the presence of these drugs.

Hence, mutations that prevent an interaction can be selected from large libraries or randomly generated alleles. Similarly, molecules that dissociate or prevent an interaction could be selected from large libraries of peptides or compounds (54-55).

SRS system

A new Sos-recruitment system (SRS) was described by Aronheim (57). This SRS is based on the observation that the mammalian GDP-GTP exchange factor (GEF) hSos can only activate Ras when hSos is localized to the plasma membrane. In mammalian cells, this occurs by recruitment to the cytoplasmatic tail of activated growth factor receptors. The yeast *Saccharomyces cerevisiae* requires a functional Ras signaling pathway for cell viability. A yeast strain containing a point mutation in the yeast Ras GEF (Ras guanyl nucleotide exchange factor), cdc25-2, shows temperature-sensitive growth, but expression of hSos artificially targeted to the membrane by myristoylation or farnesylation can restore growth at the non-permissive temperature.

The SRS uses the requirement of hSos to be recruited to the membrane to rescue growth as a method to detect protein-protein interactions. The target protein, is fused to hSos, and the partner protein or cDNA to be screened is fused to a membrane-localization signal. When these hybrid proteins are co-expressed in a cdc25-2 yeast strain, the cells only grow at the nonpermissive temperature if the two fusion proteins interact, allowing the recruitment of hSos to the membrane. Mammalian Ras can bypass the requirement for a functional Ras GEF and represents a predictable false positive in this system. In a second publication, the author proposed that introduction of a mammalian GTPase activating protein (mGAP) reduces the amount of false positives, thereby significantly enhancing the efficiency of the SRS system

Figure 5: Principle of the SRS system. In cells that express a chimera of Sos-X (where X is a target protein) along with a membrane-localized version of a protein that interacts with X (cDNA), the interaction between X and its partner recruits Sos to the plasma membrane, resulting in Ras activation (56)
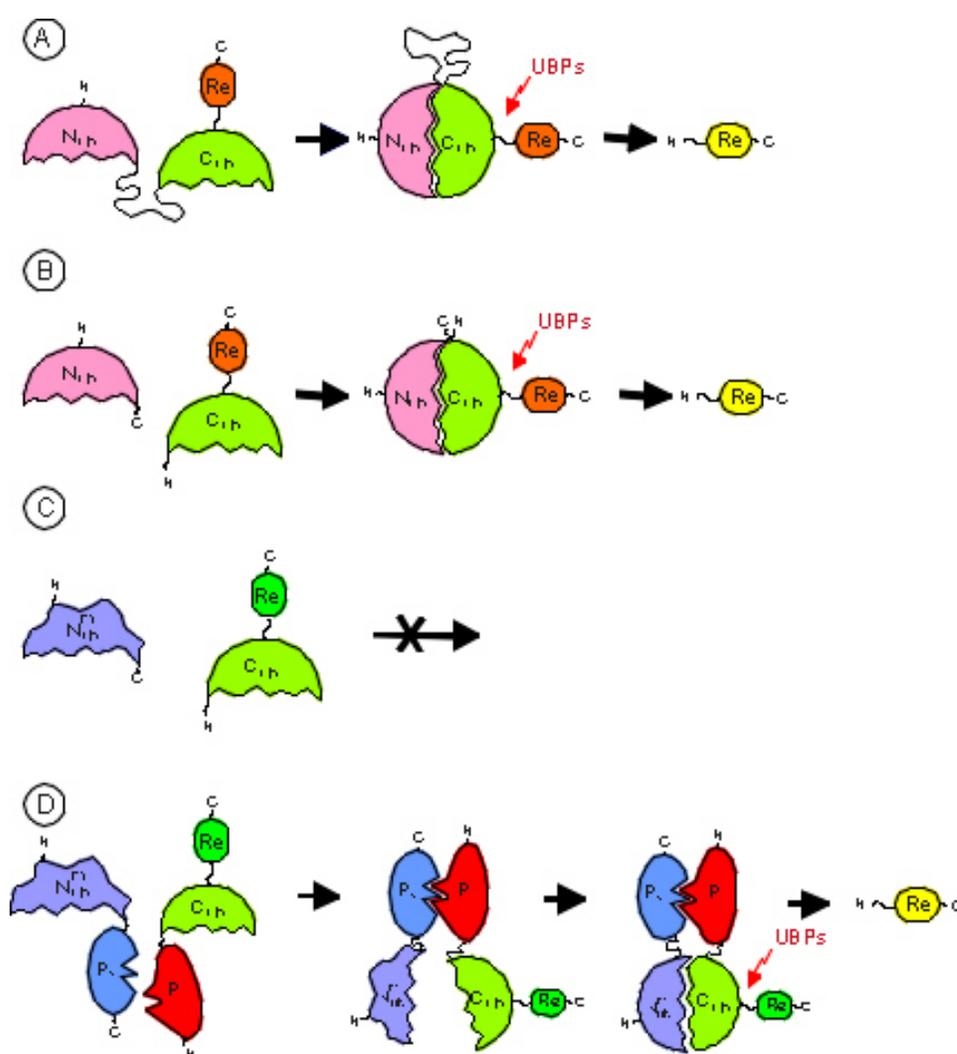
(56). SRS may be the method of choice in studying interactions that involve transcriptional activators or repressors, because it is not based on a transcriptional readout. In addition, certain proteins may function more physiologically when expressed in the cytoplasm rather than the nuclear milieu. For example, the SRS should be more suitable for examining interactions between proteins that require modification by cytoplasmic or membrane-associated enzymes.

USPS system

The ubiquitin-based split-protein sensor system (USPS) is based on the fact that newly formed fusions between ubiquitin and proteins are rapidly cleaved by ubiquitin-specific proteases in eukaryotes.

Figure 6: Principle underlying the ubiquitin-based split-protein system (USPS) (A) A Ub moiety bearing an insertion (wavy red line) between its N-terminal (N_{ub}; pink) and C-terminal (C_{ub}; green) subdomains is linked to a reporter protein (Re; orange, yellow). The insertion does not detectably interfere with the Ub folding, which is required for the in vivo cleavage of the fusion by Ub-specific proteases (UBPs; red lightning arrow), yielding the free reporter. **(B)** When N_{ub} and C_{ub} are co-expressed as separate fragments, with C_{ub} still linked to the reporter, significant in vivo reconstitution is observed. **(C)** In vivo reconstitution of the Ub from its separate, co-expressed fragments does not occur with a mutant N_{ub} fragment, denoted as N_{ub}^{m}. **(D)** N_{ub}^{m} supports reconstitution if the two Ub fragments are linked to polypeptides P_1 and P_2 (red and blue, respectively) that interact in vivo (58).

Indeed, protein fusions containing ubiquitin, a 76-residue single-domain protein, are rapidly cleaved in vivo by ubiquitin-specific proteases (UBP) which recognize the folded conformation of ubiquitin. When a C-terminal fragment of ubiquitin (C_{ub}) is expressed as a fusion to a reporter protein, the fusion is cleaved only if an N-terminal fragment of ubiquitin (N_{ub}) is also expressed in the same cell. This reconstitution of native ubiquitin from its fragments, detectable by the *in vivo* cleavage assay, is not observed with a mutationally altered N_{ub}.

However, if C_{ub} and the altered N_{ub} are each linked to polypeptides that interact *in vivo*, the cleavage of the fusion containing C_{ub} is restored, yielding a generally applicable assay for kinetic and equilibrium aspects of *in vivo* protein interactions. This method was termed USPS which stands for ubiquitin-based split protein sensor (58).

To higher eukaryotes: Mammalian Two-hybrid

Here the gene of interest is cloned into pM, where it is expressed as a fusion with the GAL4 DNA-BD. The second gene is cloned into pVP16, where it is expressed as a fusion with VP16 AD. A third vector, pG5CAT, contains the CAT reporter gene downstream of a GAL4 responsive UAS. These three vectors are cotransfected into a mammalian cell line. Interaction between proteins X and Y is assayed by measuring CAT gene expression by any standard method (e.g. ELISA). The system might be meaningful to confirm interactions (59). An inverse mammalian two-hybrid system for secretase was described (60).

To determine if a protease contained "secretase-like" activity, a 20 amino-acid cleavage cassette for secretase (from amyloid precursor protein "APP") was inserted into the GAL4 gene of *S. cerevisiae*. When this GAL4/APP fusion construct is transiently transfected into 293 cells and indirectly assayed for its functional capabilities through chloramphenicol acetyltransferase (CAT) transactivation, this protein retained functional activity confirming that no endogenous secretase activity was present and that the fusion protein retained its transactivation activity. After cotransfection of cDNAs encoding "APP" cleaving proteases, the ability of GAL4 to transactivate the reporter CAT gene was indeed reduced (60).
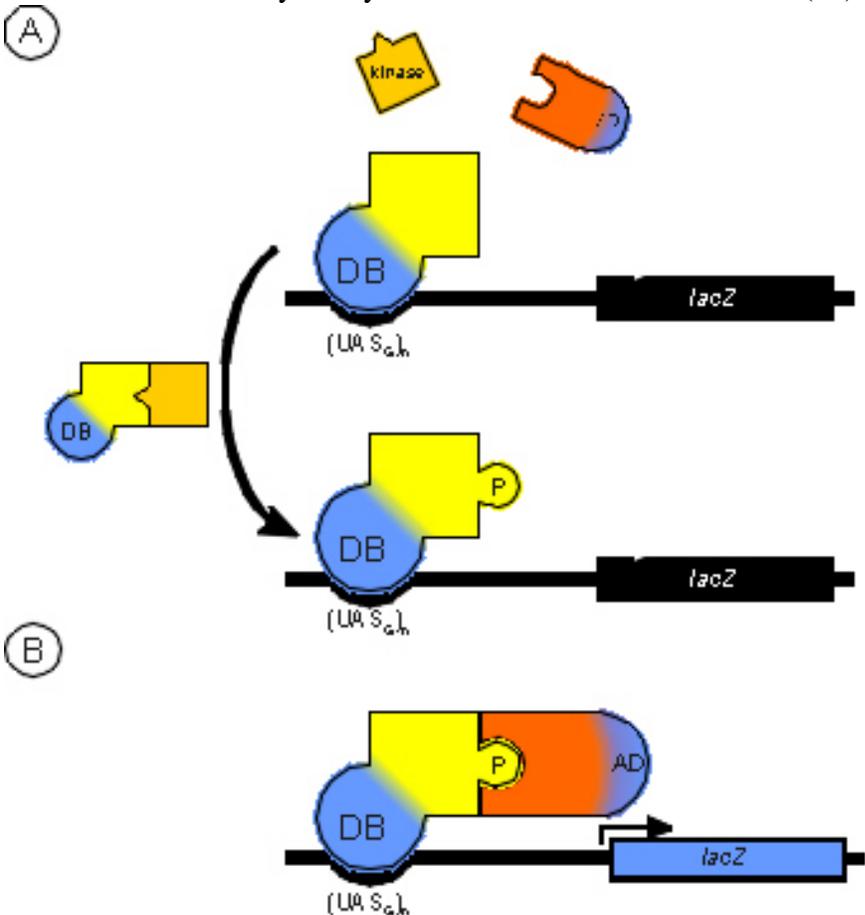

Figure 7: Principle of the kinase tri-brid system. The protein target needs to be phosphorylated (P) by a co-expressed kinase before interaction occurs.

Three-hybrid systems

Several recent studies probed more complex macromolecular interactions of at least three partners.

Kinase three-hybrid system (tri-brid)

This kinase three-hybrid system called tri-brid system can detect protein-protein interactions that depend on post translational modifications. Indeed, one of the limitations of the original version of the two-hybrid system has been the lack of some crucial post-translational modifications (such as tyrosine phosphorylation) which do not occur in *Saccharomyces cerevisiae*. Higher eukaryotic cells utilize tyrosine phosphorylation as a major modification in signal transduction pathways in response to extracellular stimuli. Although *S. cerevisiae* do possess dual-specificity kinases, to date, no mono-specific tyrosine kinases have been described. This was solved by introduction of a third component, a cytosolic tyrosine kinase, which then phosphorylated substrates in the yeast cell. In principle, it should be possible to incorporate any posttranslational or allosteric regulation that is desired. However, care must be taken that the posttranslational modification is correctly carried out inside the yeast cell (61).
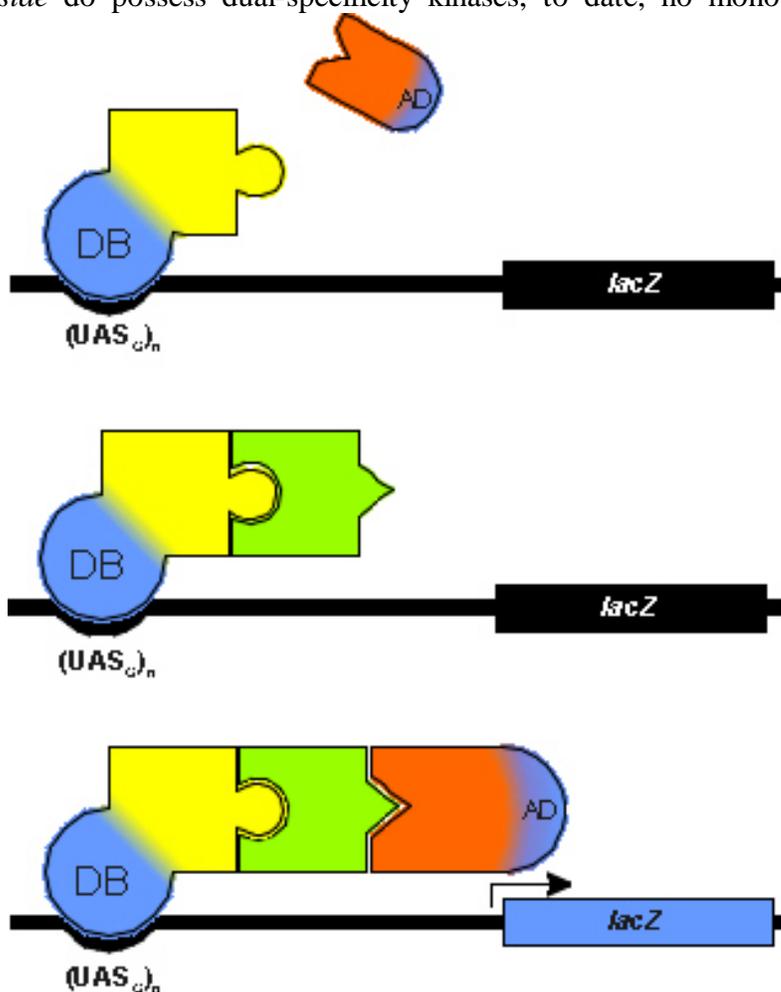
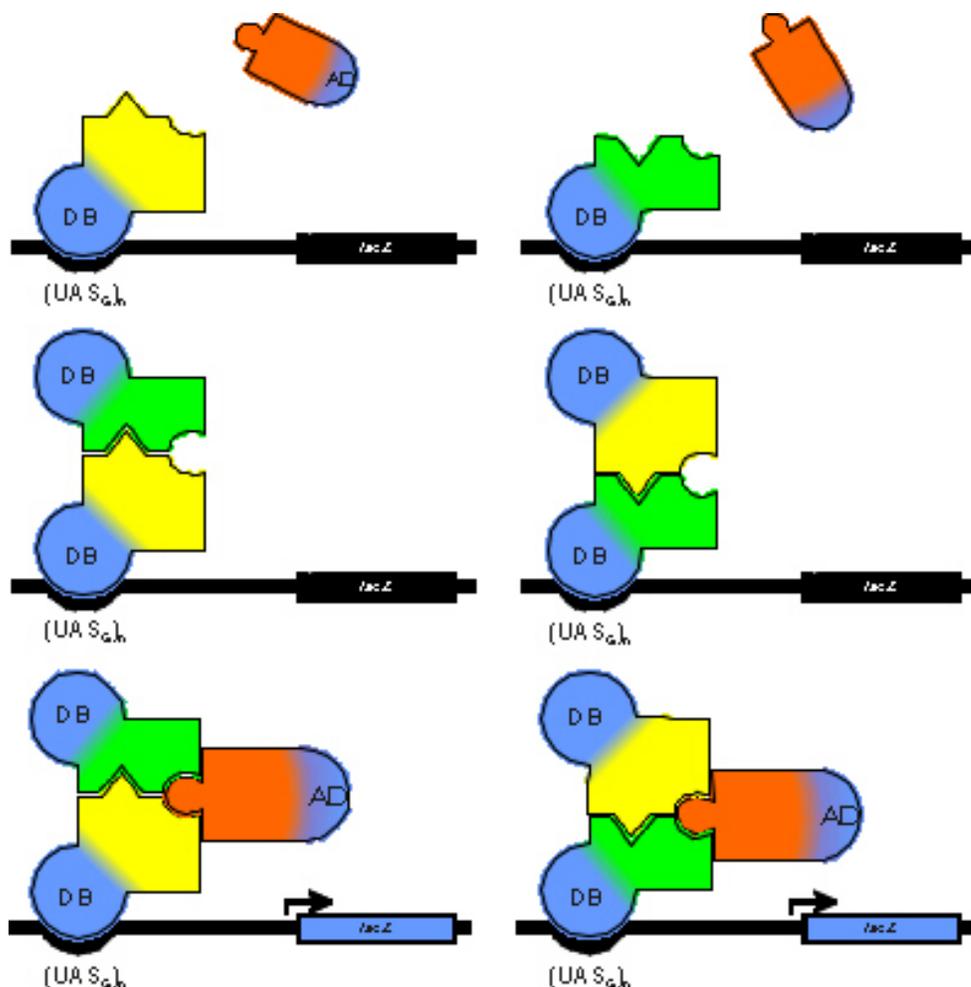*Protein three-hybrid system*

The protein three-hybrid system represents a natural extension of the two-hybrid system in that stable interaction of proteins X and Y may rely on the presence of a third protein Z. Protein Z either mediates the interaction or induces a conformational change in one of the proteins (e.g. X) so that it promotes interaction with another protein Y (62). Zhang used this system to demonstrate the Grb2-mediated interaction of the cytoplasmic domain of the epidermal growth factor receptor with the Sos protein.

Figure 8: Principle that forms the basis of all multiple hybrid systems that make use of three components at the UAS$_G$. The third "hybrid" is colored in green.

Recently a variant of this system was proposed by Tirode *et al.* (32). This system uses both LexA and GAL4 plasmids and is based on the addition of a Met25 promoter for the third partner. This "third man" is cloned on the same plasmid that already encodes the GAL4 and LexA fusion protein. By doing this, a selection marker is saved. Transcription of the third protein is repressed upon addition of methionine in the culture medium, and provides an elegant way to perform negative controls (32). In order to adapt

the two-hybrid system to the needs of caspases we devised in parallel our own three-hybrid system. This system can be used to study any hetero-dimeric target (Figure 9)(68).

Studies with extracellular domains of transmembrane receptors (10,11), have led to the development of a peptide ligand three-hybrid system (63). In this system, the extracellular domain of a transmembrane receptor is fused to the GAL4BD and separately to the GAL4AD. Expression of the native peptide ligand for the receptor leads to receptor dimerization via peptide ligand binding and results in a transcriptional readout. This was demonstrated for both the growth hormone receptor and the receptor for vascular endothelial growth factor flk1/KDR (63).

*Peptide ligand and small ligand three-hybrid system*

Figure 9: Use of three-hybrid to screen with hetero-dimeric targets. Here the target consists of two-fusion protein (yellow and green) (68).

The small ligand variant is based on the previously discussed concept and exploits the use of a heterodimer of covalently linked small organic ligands, in this case dexamethasone and FK506, as inducers of dimerization. Dexamethasone interacts with a rat glucocorticoid receptor while FK506 interacts with its binding protein, FKBP12. Yeast expressing fusion proteins of the hormone binding domain of the rat glucocorticoid receptor fused to LexA and FKBP12 fused to a transcriptional activation domain activated reporter genes when plated on medium containing the dexamethasone-FK506 heterodimer. The reporter gene activation is completely abrogated in a competitive manner by the presence of excess FK506. Since small ligand-receptor interactions underlie many fundamental processes and form the basis for pharmacological intervention this system might gain interest. However, limitations of the assay include those of the two-hybrid system in general and in particular the question

of how extensively transmembrane receptors function in yeast. Since this system depends on diffusion of the hybrid ligand into the yeast cells, permeability problems may also limit its general applicability (64).

RNA three-hybrid system

The RNA three-hybrid system requires two hybrid proteins and one hybrid RNA. One part of the hybrid RNA acts in a known interaction, whereas the other part can be used to screen for orphan RNA-binding proteins. In this three-hybrid system, the DNA binding and transcription activation domains of the yeast transcriptional activator GAL4 are brought together via the interaction of recombinant fusion proteins with a recombinant RNA. The method provides a system for studying RNA-protein interactions with the genetic advantages of the two-hybrid system. It may be used to detect specific RNA-binding proteins or target RNAs from a library of cDNAs, or to analyze the structural specificity of identified RNA-protein interactions. By use of this system it was shown that when HIV-1 RevM10 is fused to the GAL4DB it tightly binds to its target RNA, the Rev responsive element (RRE) (65). Other examples are the binding of the iron regulatory protein 1 (IRP1) to the iron response element (IRE), and of HIV trans-activator protein (Tat) to the HIV trans-activation response element (TAR) RNA sequence (67). This RNA three-hybrid system may have broad application in the identification of RNA-binding proteins and RNAs, as well as in the detailed analysis of their interactions although it is not known how strong the RNA-protein has to be in order to be detected.
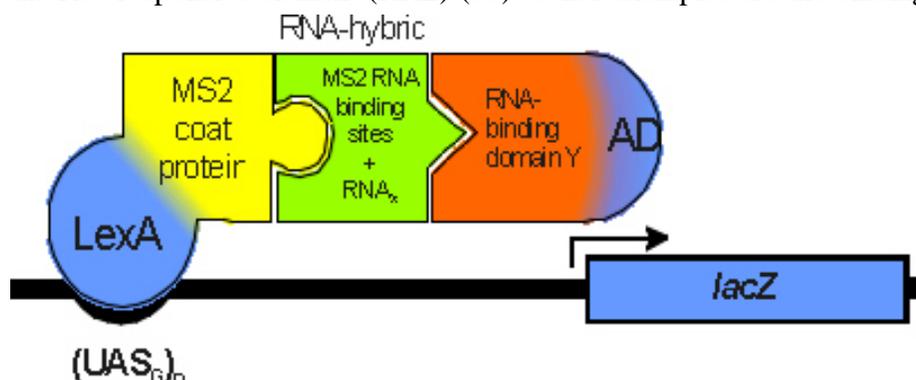
Figure 10: Reconstitution of LacZ transcription by a hybrid RNA molecule. LexA is fused to the MS2 coat protein to form hybrid protein 1. Hybrid protein 2 consists of the GAL4AD linked to the RNA binding domain, Y, you wish to test. The third hybrid consists of two MS2 RNA binding sites and the RNA sequence you wish to test, RNA_X.

NEW APPLICATIONS OF YEAST TWO-HYBRID

Introduction

During the past decade, together with the evolution of the technique, new applications became apparent. Apart from the molecular dissection of known interactions and the identification of new potential interacting partners, several new problems can now be assayed by use of two-hybrid based technology.

Interaction suppression

Interaction suppression is one of the few techniques available to evaluate the biological significance of an interaction between two proteins. This method makes extensive use of the two-hybrid technique in a very original manner. In a first step one screens for mutations that affect binding of a target protein to its partners (B). Once a mutant version of the bait (A) is found, one can study the influence of this mutation on the phenotype. However, if one finds mutants of protein A that do not interact with protein B, one cannot exclude that also interaction with an unknown protein C is abolished and accounts for the change in phenotype. Therefore, in a second step one has to look for binding suppressors, meaning that one makes a mutant version of B that is able to interact with the mutated version of A and restores the altered phenotype. This approach was taken to study the Ras/Raf pathway (69-71).

Protease Trap

Here the system relies entirely on the nuclear localization of the constructed fusion protein. A functional transcription factor is fused to a domain preventing it from going to the nucleus. If an orphan protease site is cloned in between the transcription factor and this domain, one can screen a protease which cuts off the anti-nuclear localization domain enabling nuclear localization of the transcription factor and finally results in reporter gene transcription. This approach could also be applied to screen for target sequences for a known protease by cloning random sequences in between the two functional domains (72). It should be noted that GAL4 may also be cleaved resulting in both unexpected results and novel applications (73).

Whole genome approaches using the two-hybrid system

One of the most spectacular and ambitious applications consists of establishing so-called protein linkage maps (PLM). These maps consist of all possible protein interactions that occur during the entire lifespan of a cell. PLMs might provide new functions of well-studied proteins by identification of unexpected interactions, clarify cross-connections between pathways, result in new drug-targets and the identification of new functions and gain some insight in the overall cell complexity.

Initially random libraries fused to the DNA-binding domain and random activation domain hybrids are simply used to screen for all possible interactions. This approach was applied to the *E. coli* bacteriophage T7. The T7 genome of ~40 kb, coding for ~55 proteins was used to generate the two libraries, which were then transformed into yeast strains of opposite mating type. By doing this, a total set of interactions was identified and termed a protein linkage map. What was interesting about this approach, apart from the biological implications of the interactions found, is that several positives were interactions between adjacent domains of the same polypeptide. These interactions are presumably due to the random fragmentation of protein-coding regions in the library construction, exposing interior surfaces that are capable of finding their complementary surfaces, as must occur during protein folding (77).

Although the genome of the yeast *Saccharomyces cerevisiae* is completely sequenced and despite successful genetic work, 60% of the yeast gene have no assigned function and half of those encode

putative proteins without any homology with known proteins. By performing exhaustive two-hybrid screens the yeast PLM is currently under construction. For each bait, a selected limited set of interacting preys is classified in categories of distinct heuristic values. From this classification, new baits are chosen and used in second-round screens. Repeating this procedure several times has led to the characterization of a network of interactions (74). The approach taken for T7, using 1 copy centromeric plasmid vector, showed that 4% of all constructs was able to auto-activate. Scaling up this approach to an organism like yeast with approximately 7000 proteins becomes not trivial since, false positives and negatives might seriously interfere with subsequent analysis. Therefore, instead of using random libraries as in the above mentioned exhaustive cyclic screening procedure, a more labor intensive route is taken by performing a "double-PCR-*in-vivo*-cloning"- method in order to clone every full length ORF in the correct reading frame in both two-hybrid vectors. Using robots and a mating procedure every possible interaction couple, around 50 million, can be assayed (66). Defining the biological relevance of identified interactions will be the next rate-limiting step in deciphering the molecular mechanisms of biological processes.

REFERENCES

1. Choi K.Y., Satterberg B., Lyons D.M., Elion E.A. 1994. Ste5 tethers multiple protein kinases in the MAP kinase cascade required for mating in S. cerevisiae. *Cell* **78**, 499-512.
2. Kischkel F.C., Hellbardt S., Behrmann I., Germer M., Pawlita M., Krammer P.H., Peter M.E. 1995. Cytotoxicity-dependent APO-1 Fas/CD95-associated proteins form a death-inducing signaling complex DISC with the receptor. *EMBO J.* **14**, 5579-5588.
3. Phizicky E.M., Fields S. 1995. Protein-protein interactions: methods for detection and analysis. *Microbiol.Rev.* **59**, 94-123.
4. Hope I.A., Struhl K. 1986. Functional dissection of a eukaryotic transcriptional activator protein, GCN4 of yeast. *Cell* **46**, 885-894.
5. Keegan L., Gill G., Ptashne M. 1986. Separation of DNA binding from the transcription-activating function of a eukaryotic regulatory protein. *Science* **231**, 699-704.
6. Brent R., Ptashne M. 1985. A eukaryotic transcriptional activator bearing the DNA specificity of a prokaryotic repressor. *Cell* **43**, 729-736.
7. Ma J., Ptashne M. 1988. Converting a eukaryotic transcriptional inhibitor into an activator. *Cell* **55**, 443-446.
8. Fields S., Song O.K. 1989. A novel genetic system to detect protein-protein interactions. *Nature* **340**, 245-246.
9. Fields S., Sternglanz R. 1994. The two-hybrid system: an assay for protein-protein interactions. *Trends Genet.* **10**, 286-292.
10. Young K.H., Ozenberger B.A. 1995. Investigation of ligand binding to members of the cytokine receptor family within a microbial system. *Ann.N.Y.Acad.Sci.* **766**, 279-281.
11. Kajkowski E.M., Price L.A., Pausch M.H., Young K.H., Ozenberger B.A. 1997. Investigation of growth hormone releasing hormone receptor structure and activity using yeast expression technologies. *J.Recept.Signal.Transduct.Res.* **17**, 293-303.
12. Denis C.L., Ferguson J., Young E.T. 1983. mRNA levels for the fermentative alcohol dehydrogenase of Saccharomyces cerevisiae decrease upon growth on a nonfermentable carbon source. *J.Biol.Chem.* **258**, 1165-1171.

13. Tornow J., Santangelo G.G. 1990. Efficient expression of Saccharomyces cerevisiae glycolytic gene ADH1 is dependent upon a cis-acting regulatory element UAS-PRG found initially in genes encoding ribosomal proteins. *Gene* **90**, 79-85.

14. Ruohonen L., Penttila M., Keranen S. 1991. Optimization of Bacillus amylase Production by Saccharomyces cerevisiae. *Yeast* **7**, 337-346.

15. Legrain P., Dokhelar M.C., Transy C. 1994. Detection of protein-protein interactions using different vectors in the two-hybrid system. *Nucleic Acids Res*. **22**, 3241-3242.

16. Fradet Y., Tardif M., Parent-Vaugeois C. 1985. The use of multiparameter flow cytometry in the detection and evaluation of human bladder tumors. *Union.Med.Can.* **114**, 778-780.

17. Parent S.A., Fenimore C.M., Bostian K.A. 1985. Vector systems for the expression, analysis and cloning of DNA sequences in S. cerevisiae. *Yeast* **1**, 83-138.

18. Green N., Alexander H., Olson A., Alexander S., Shinick T.M., Sutcliffe J.G., Lerner R.A. 1982. Immunogenic structure of the influenza virus hemagglutinin. *Cell* **28**, 477-487.

19. Kaufer A.F., Fried H.M., Schwindinger W.F., Jasin M., Warner J.R. 1983. Cycloheximide resistance in yeast: the gene and its protein. *Nucleic Acids Res.* **11**, 3123

20. Langlands K., Prochownik E.V. 1997. A rapid method for the preparation of yeast lysates that facilitates the immunodetection of proteins generated by the yeast two-hybrid system. *Anal.Biochem.* **249**, 250-252.

21. Beranger F., Aresta S., de Gunzburg J., Camonis J. 1997. Getting more from the two-hybrid system: N-terminal fusions to LexA are efficient and sensitive baits for two-hybrid studies. *Nucleic Acids Res.* **25**, 2035-2036.

22. Johnston S.A., Zavortink M.J., Debouck C., Hopper J.E. 1986. Functional domains of the yeast regulatory protein GAL4. *Proc.Natl.Acad.Sci.U.S.A*. **83**, 6553-6557.

23. Marmorstein R., Carey M., Ptashne M., Harrison S.C. 1992. DNA recognition by GAL4: structure of a protein-DNA complex. *Nature* **356**, 408-414.

24. Carey M., Kakidani H., Leatherwood J., Mostashari F., Ptashne M. 1989. An amino-terminal fragment of GAL4 binds DNA as a dimer. *J.Mol.Biol.* **209**, 423-432.

25. Silver P.A., Keegan L.P., Ptashne M. 1984. Amino terminus of the yeast GAL4 gene product is sufficient for nuclear localization. *Proc.Natl.Acad.Sci.U.S.A.* **81**, 5951-5955.

26. Silver P.A., Chiang A., Sadler I. 1988. Mutations that alter both localization and production of a yeast nuclear protein. *Genes and Dev.* **2**, 707-717.

27. Silver P.A., Brent R., Ptashne M. 1986. DNA binding is not sufficient for nuclear localization of regulatory proteins in Saccharomyces cerevisiae. *Mol.Cell.Biol.* **6**, 4763-4766.

28. Brent R., Ptashne M. 1984. A bacterial repressor protein or a yeast transcriptional terminator can block upstream activation of a yeast gene. *Nature* 612-615.

29. Dagher M.C., Filhol Cochet O. 1997. Making hybrids of two-hybrid systems. *Biotechniques* **22**, 916-8,920-2.

30. Guthrie C., Fink G.R. 1991. Guide to yeast genetics and molecular biology. *Methods Enzym.* **194**, 1-932.

31. Giniger E., Ptashne M. 1988. Cooperative DNA binding of the yeast transcriptional activator GAL4. *Proc.Natl.Acad.Sci.U.S.A.* **85**, 382-386.

32. Tirode F., Malaguti C., Romero F., Attar R., Camonis J., Egly J.M. 1997. A conditionally expressed third partner stabilizes or prevents the formation of a transcriptional activator in a three-hybrid system. *J.Biol.Chem.* **272**, 22995-22999.

33. Ebina Y., Takahara Y., Kishi F., Nakazawa A. 1983. LexA protein is a repressor of colicin E1 gene. *J.Biol.Chem.* **258**, 13258-13261.

34. Estojak J., Brent R., Golemis E.A. 1995. Correlation of two-hybrid affinity data with in vitro measurements. *Mol.Cell.Biol.* **15**, 5820-5829.

35. West R.W.Jr., Yoccum R.R., Ptashne M. 1984. Saccharomyces cerevisiae GAL1-GAL10 divergent promotor region: location and function of the upstream activating sequence UAS_G. *Mol.Cell.Biol.* 2467-2478.

36. Gietz R.D., Schiestl R.H. 1991. Applications of high efficiency lithium acetate transformation of intact yeast cells using single-stranded nucleic acids as carrier. *Yeast* **7**, 253-263.

37. Gietz R.D., St.Jean A., Woods R.A., Schiestl R.H. 1992. Improved method for high efficiency transformation of intact yeast cells. *Nucleic Acids Res.* **20**, 1425

38. Gietz R.D., Schiestl R.H., Willems A.R., Woods R.A. 1995. Studies on the transformation of intact yeast cells by the LiAc/SS-DNA/PEG procedure. *Yeast* **11**, 355-360.

39. Gietz   R.D.   1997.   The   Gietz   Lab   Yeast   Transformation   Home   Page   . *http://www.umanitoba.ca/faculties/medicine/human_genetics/gietz/Trafo. html*

40. Schenk J.A., Heymann S., Peters L.E., Micheel B. 1996. Screening for recombinant plasmids in yeast colonies of the two-hybrid system using PCR. *Biotechniques* **20**, 812-4, 816.

41. Herskowitz I. 1988. Life cycle of the budding yeast Saccharomyces cerevisiae. *Microbiol.Rev.* **52**, 536

42. Bartel P., Chien C.T., Sternglanz R., Fields S. 1993. Elimination of false positives that arise in using the two-hybrid system. *Biotechniques* **14**, 920-924.

43. Golemis E.A., Khazak V. 1997. Alternative yeast two-hybrid systems. The interaction trap and interaction mating. *Methods Mol.Biol.* **63**, 197-218.

44. Hengen P.N. 1997. False positives from the yeast two-hybrid system. *Trends Biochem.Sci.* **22**, 33-34.

45. Golemis E. 1997. Lab of Erica Golemis. *http://chaos.fccc.edu/research/labs/ golemis*

46. Yavuzer U., Goding C.R. 1995. pWITCH: a versatile two-hybrid assay vector for the production of epitope/activation domain-tagged proteins both in vitro and in yeast. *Gene* **165**, 93-96.

47. Wong C., Naumovski L. 1997. Method to screen for relevant yeast two-hybrid-derived clones by coimmunoprecipitation and colocalization of epitope-tagged fragments, application to Bcl-X_L. *Anal.Biochem.* **252**, 33-39.

48. Roder K.H., Wolf S.S., Schweizer M. 1996. Refinement of vectors for use in the yeast two-hybrid system. *Anal.Biochem.* **241**, 260-262.

49. Le Douarin B., Pierrat B., vom Baur E., Chambon P., Losson R. 1995. A new version of the two-hybrid assay for detection of protein-protein interactions. *Nucleic Acids Res.* **23**, 876-878.

50. Drees BL. 1999. Progress and variations in two-hybrid and three-hybrid technologies. *Curr Opin Chem Biol.* **3**, 64-70.

51. Leanna C.A., Hannink M. 1996. The reverse two-hybrid system: a genetic scheme for selection against specific protein/protein interactions. *Nucleic Acids Res.* **24**, 3341-3347.

52. Boeke J.D., LaCroute F., Fink G.R. 1984. A positive selection for mutants lacking orotidine-5'-phosphate decarboxylase activity in yeast: 5-fluoro-orotic acid resistance. *Mol.Gen.Genet.* **197**, 345-346.

53. Shih H.M., Goldman P.S., DeMaggio A.J., Hollenberg S.M., Goodman R.H., Hoekstra M.F. 1996. A positive genetic selection for disrupting protein-protein interactions: identification of CREB mutations that prevent association with the coactivator CBP. *Proc.Natl.Acad.Sci.U.S.A.* **93**, 13896-13901.

54. Vidal M., Brachmann R.K., Fattaey A., Harlow E., Boeke J.D. 1996a. Reverse two-hybrid and one-hybrid systems to detect dissociation of protein-protein and DNA-protein interactions. *Proc.Natl.Acad.Sci.U.S.A.* **93**, 10315-10320.

55. Vidal M., Braun P., Chen E., Boeke J.D., Harlow E. 1996b. Genetic characterization of a mammalian protein-protein interaction domain by using a yeast reverse two-hybrid system. *Proc.Natl.Acad.Sci.U.S.A.* **93**, 10321-10326.

56. Aronheim A. 551997a. Improved efficiency Sos recruitment system: expression of the mammalian GAP reduces isolation of Ras GTPase false positives. *Nucleic Acids Res.* **25**, 3373-3374.

57. Aronheim A., Zandi E., Hennemann H., Elledge S.J., Karin M. 1997b. Isolation of an AP-1 Repressor by a Novel Method for Detecting Protein-Protein Interactions. *Mol.Cell.Biol.* **17**, 3094-3102.

58. Johnsson N., Varhavsky A. 1994. Split ubiquitin as a sensor of protein interactions in vivo. *Proc.Natl.Acad.Sci.U.S.A.* **91**, 10340-10344.

59. Buchert M., Schneider S., Adams M.T., Hefti H.P., Moelling K., Hovens C.M. 1997. Useful vectors for the two-hybrid system in mammalian cells. *Biotechniques* **23**, 396-8, 400, 402.

60. Dixon E.P., Johnstone E.M., Liu X., Little S.P. 1997. An inverse mammalian two-hybrid system for secretase and other proteases. *Anal.Biochem.* **249**, 239-241.

61. Osborne M.A., Dalton D., Kochan J.P. 1995. The yeast tribrid system-genetic detection of tans-phosphorylated ITAM-SH2-interactions. *Biotechnology* **13**, 1474-1478.

62. Zhang J., Lautar S. 1996. A yeast three-hybrid method to clone ternary protein complex components. *Anal.Biochem.* **242**, 68-72.

63. Ozenberger B.A., Young K.H. 1995. Functional interaction of ligands and receptors of the hematopoietic superfamily in yeast. *Mol.Endocrinol.* **9**, 1321-1329.

64. Licitra E.J., Liu J.O. 1996. A three-hybrid system for detecting small ligand-protein receptor interactions. *Proc.Natl.Acad.Sci.U.S.A.* **93**, 12817-12821.

65. Putz U., Skehel P., Kuhl D. 1996. A tri-hybrid system for the analysis and detection of RNA-protein interactions. *Nucleic Acids Res.* **24**, 4838-4040.

66. Sengupta D.J. 1996. Application of two-hybrid and three-hybrid technologies. *EMBO Two-Hybrid Course,* October 12, Munich.

67. Brachmann R.K., Boeke J.D. 1997. Tag games in yeast: the two-hybrid system and beyond. *Curr.Op.Biotech.* 561-568.

68. Van Criekinge W, van Gurp M, Decoster E, Schotte P, Van de Craen M, Fiers W, Vandenabeele P, Beyaert R 1998 Use of the yeast three-hybrid system as a tool to study caspases. *Anal Biochem.* **263**, 62-64.

69. Van Aelst L., White M.A., Wigler M.H. 1994. Ras partners. *Cold.Spring.Harb.Symp.Quant. Biol.* **59**, 181-186.

70. White M.A., Nicolette C., Minden A., Polverino A., Van Aelst L., Karin M., Wigler M.H. 1995. Multiple Ras functions can contribute to mammalian cell transformation. *Cell* **80**, 533-541.

71. Khosravi-Far R., White M.A., Westwick J.K., Solski P.A., Chrzanowska-Wodnicka M., Van Aelst L., Wigler M.H., Der C.J. 1996. Oncogenic Ras activation of Raf/mitogen-activated protein kinase-independent pathways is sufficient to cause tumorigenic transformation. *Mol.Cell.Biol.* **16**, 3923-3933.

72. Jessen T. 1996. Pharmaceutical applications of the interaction trap. *EMBO Two-Hybrid Course,* October 10, Munich.

73. Van Criekinge W., Cornelis S., Van de Craen M., Vandenabeele P., Fiers W. & Beyaert R. 1999. GAL4 is a substrate for caspases: implications for two-hybrid screening and other GAL4-based assays. *Mol. Cell Biol. Res. Commun.* **1**, 158-161.

74. Fromont-Racine M., Rain J.C., Legrain P. 1997. Towards a functional analysis of the yeast genome through exhaustive two-hybrid screens. *Nature Gen.* 277-281.

75. Robzyk K, Kassir Y. 1992. A simple and highly efficient procedure for rescuing autonomous plasmids from yeast. *Nucleic Acids Res* **20**, 3790.

76. James P, Halladay J, Craig EA 1996. Genomic libraries and a host strain designed for highly efficient two-hybrid selection in yeast. *Genetics* **144**, 1425-36.

77. Bartel PL, Roecklein JA, SenGupta D, Fields S 1996. A protein linkage map of Escherichia coli bacteriophage T7. *Nat Genet* **12**, 72-77.