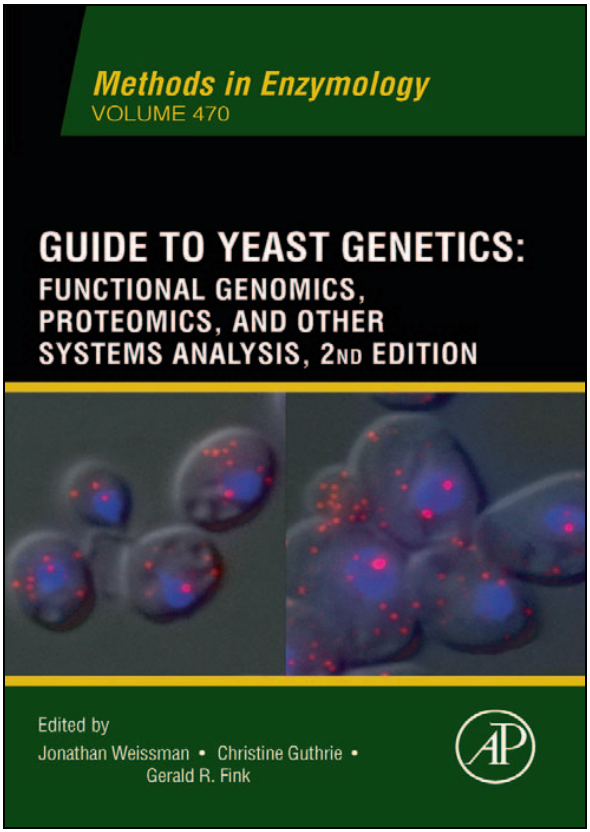

From: Matija Dreze, Dario Monachello, Claire Lurin, Michael E. Cusick, David E. Hill, Marc Vidal, and Pascal Braun, High-Quality Binary Interactome Mapping, In John Abelson and Melvin Simon editors: *METHODS IN ENZYMOLOGY,* Vol. 470,
Burlington: Academic Press, 2010, pp.281-315.
ISBN: 978-0-12-375172-0

CHAPTER TWELVE

High-Quality Binary Interactome Mapping

Matija Dreze,[*,†,‡,§] Dario Monachello,[*,†,¶] Claire Lurin,[¶] Michael E. Cusick,[*,†,‡] David E. Hill,[*,†,‡] Marc Vidal,[*,†,‡] *and* Pascal Braun[*,†,‡]

Contents

1. Introduction 282
2. High-Quality Binary Interactome Mapping 283
   2.1. Production and verification of Y2H datasets 285
   2.2. Validation of Y2H datasets to produce reliable binary interactome maps 288
   2.3. Biological evaluation of binary interactome maps 291
3. High-Throughput Y2H Pipeline 292
   3.1. Assembly of DB-X and AD-Y expression plasmids 292
   3.2. Yeast transformation 296
   3.3. Autoactivator removal and AD-Y pooling 298
   3.4. Screening and phenotyping 301
   3.5. Verification 305
   3.6. Media and plates 307
4. Validation Using Orthogonal Binary Interaction Assays 309
5. Conclusion 312
Acknowledgments 312
References 313

Abstract

Physical interactions mediated by proteins are critical for most cellular functions and altogether form a complex macromolecular "interactome" network.

* Center for Cancer Systems Biology (CCSB), Dana-Farber Cancer Institute, Boston, Massachusetts, USA
† Department of Cancer Biology, Dana-Farber Cancer Institute, Boston, Massachusetts, USA
‡ Department of Genetics, Harvard Medical School, Boston, Massachusetts, USA
§ Facultés Universitaires Notre-Dame de la Paix, Namur, Belgium
¶ Unité de Recherche en Génomique Végétale (URGV), Evry Cedex, France

Methods in Enzymology, Volume 470
ISSN 0076-6879, DOI: 10.1016/S0076-6879(10)70012-4

Systematic mapping of protein–protein, protein–DNA, protein–RNA, and protein–metabolite interactions at the scale of the whole proteome can advance understanding of interactome networks with applications ranging from single protein functional characterization to discoveries on local and global systems properties. Since the early efforts at mapping protein–protein interactome networks a decade ago, the field has progressed rapidly giving rise to a growing number of interactome maps produced using high-throughput implementations of either binary protein–protein interaction assays or co-complex protein association methods. Although high-throughput methods are often thought to necessarily produce lower quality information than low-throughput experiments, we have recently demonstrated that proteome-scale interactome datasets can be produced with equal or superior quality than that observed in literature-curated datasets derived from large numbers of small-scale experiments. In addition to performing all experimental steps thoroughly and including all necessary controls and quality standards, careful verification of all interacting pairs and validation tests using independent, orthogonal assays are crucial to ensure the release of interactome maps of the highest possible quality. This chapter describes a high-quality, high-throughput binary protein–protein interactome mapping pipeline that includes these features.

1. Introduction

Interactions mediated by proteins and the complex ''interactome'' networks resulting from these interactions are essential for biological systems. Mapping protein–protein, protein–DNA, protein–RNA, and protein–metabolite interactions that form ''interactome'' networks is a major goal of functional genomics, proteomics, and systems biology (Vidal, 2005). Information obtained from large-scale efforts to identify protein interaction partners yields crucial biological insights throughout a range of applications. At the single protein level, interactome maps have helped assign functions to both uncharacterized and well-studied gene products (Oliver, 2000). At the systems level, interactome maps have enabled investigations of how regulatory circuits and global cellular network properties relate to biological functions (Han *et al.*, 2004; Jeong *et al.*, 2001; Milo *et al.*, 2002; Yu *et al.*, 2008).

The two major high-throughput strategies used so far to delineate protein–protein interactome networks are: (i) binary protein–protein interaction assays, which detect direct pairwise interactions, and (ii) affinity purification followed by mass spectrometry (AP–MS) approaches, which detect biochemically stable, copurifying protein complexes containing both direct and indirect protein associations. Classically, binary interaction assays have been based on the yeast two-hybrid (Y2H) system developed 20 years ago (Fields and Song, 1989), and which has been improved over time to

increase efficiency and quality (Durfee *et al.*, 1993; Gyuris *et al.*, 1993; Vidal *et al.*, 1996). Of late, alternative approaches have been developed to detect binary interactions, such as protein arrays, protein complementation assays, and the split ubiquitin method (Miller *et al.*, 2005; Tarassov *et al.*, 2008; Zhu *et al.*, 2001).

Until recently high-throughput methods were regarded as more likely to produce lower quality information than low-throughput experiments. It has now been shown that highly reliable interactome datasets can be obtained at the scale of the whole proteome (Braun *et al.*, 2009; Cusick *et al.*, 2009; Simonis *et al.*, 2009; Venkatesan *et al.*, 2009) provided that all experimental steps are thorough and all necessary controls and quality standards are included. Lastly, careful verification of all candidate interactions and experimental validation using independent interaction assays are necessary to ensure the release of interactome maps of the highest possible quality.

Even when highly reliable, interactome maps should be considered as network models of interactions that *can* happen between all proteins encoded by the genome of an organism of interest. As such, they correspond to static representations of collapsed time-, space-, and condition-dependent interactions that dynamically regulate the behavior and developmental fate of diverse tissues. Thus, interactome maps should be used as static scaffold-like information from which the dynamic features of biologically relevant interactions, that is, those that *do* happen *in vivo*, can be modeled by integrating additional functional information such as transcriptional and phenotypic profiling data (Ge *et al.*, 2001, 2003; Gunsalus *et al.*, 2005; Vidal, 2001). Ultimately, novel potentially insightful interactions need to be evaluated for their biological significance using genetic experiments, where specific *cis*-acting interaction-defective alleles (IDAs) of one or both proteins or *trans*-acting disruptors are tested functionally (Dreze *et al.*, 2009; Endoh *et al.*, 2002; Vidal and Endoh, 1999; Vidal *et al.*, 1996; Zhong *et al.*, 2009).

2. High-Quality Binary Interactome Mapping

The quality of any dataset can be affected by a high rate of "false positives" and need to be addressed in two fundamentally different contexts. One relates to avoidable experimental errors leading to wrong information, and the other relates to as yet undiscovered fundamental properties of proteins (Fig. 12.1). Our binary interactome mapping strategy is designed to differentiate between these two classes of issues designated "technical" and "biological" false positives.

- *Technical false positives*

All techniques used to map protein interactions can give rise to artifacts. It goes without saying that artifacts or technical false positives should be

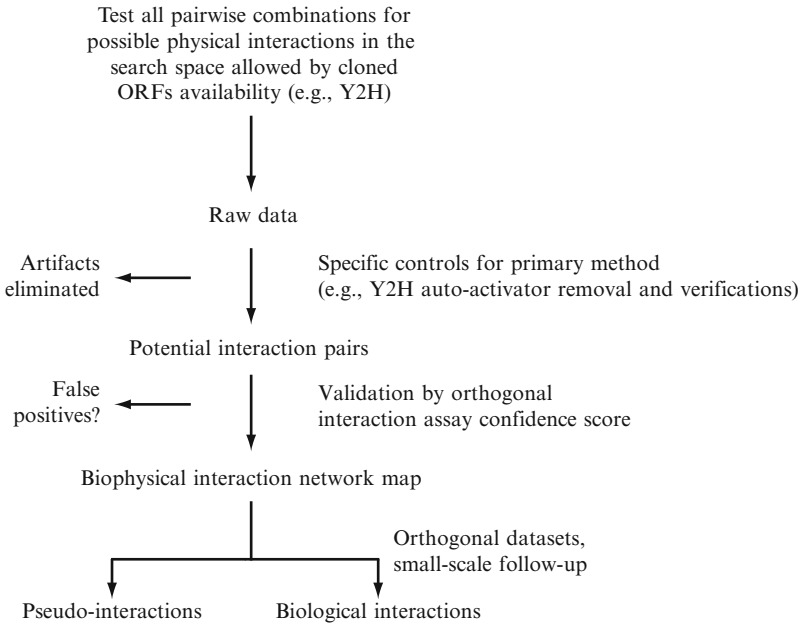

Figure 12.1 **General strategy to map binary interactome networks. All possible pairs of a search space are tested using a large-scale binary interaction detection assay such the yeast two-hybrid (Y2H) system. First-round positives constitute the raw dataset in which artifacts need to be identified and eliminated. The resulting set of putative interactions is then validated using alternative binary interaction detection assays. This step allows determination of the dataset precision or experimentally determined confidence scores for all individual interactions. Overlap of biophysical interactions with other types of datasets, such as coexpression or phenotypic profiles, or small-scale experimental follow-up, allows the identification of biologically relevant binary interactions.**

identified and removed as much as possible with appropriately designed experimental conditions and controls. Potential artifacts are different for every method and can arise systematically or sporadically. Often it takes several years of collective use, after the original description of a method, for systematic artifacts to be understood and thus become avoidable.

In biochemical AP–MS experiments, or in the design and use of antibodies, nonspecific binding by abundant proteins or contaminant proteins introduced while carrying out experiments represent technical false positives that can and should be removed. Y2H is based on a set of growth selections designed to identify the reconstitution of a transcription factor mediated by two hybrid proteins. Although powerful, the system needs to be carefully controlled because unrelated spontaneous genetic suppressors can appear during these growth selections. Such artifacts can reliably be removed by thorough implementation of the methods described below.

However careful the execution of Y2H mapping experiments or any other high-throughput methodology is, the precision of the obtained dataset (i.e., the inverse of the false discovery rate (FDR)) still needs to be determined to estimate both systematic and sporadic technical false positives that might remain undetected (Fig. 12.1). We advocate below further rigorous experimental *verifications* of all interacting pairs using the Y2H version used to produce a dataset, followed by careful *validation* using orthogonal protein interaction assays to determine overall quality. Once these steps have been implemented the result is a set of well-demonstrated interactions, proven to physically interact. We refer to such protein pairs as ''biophysical interactors''.

- *Biological false positives*

While it is plausible that most biophysical interactions are biologically relevant, their relevance, and the mechanism by which biophysically demonstrated protein interactions affect the physiology of an organism, remains to be demonstrated in subsequent, often laborious experiments (Fig. 12.1). It is theoretically possible that a subset of biophysical interactions might be biologically inconsequential because, among other possibilities, they remain either spatially or temporally separated throughout the lifetime of an organism. Such ''pseudo-interactions'' can be viewed as biological false positives that need to be eliminated or, alternatively, might represent interesting evolutionary remnants similar to the existence of pseudo-genes in many organisms (Venkatesan *et al.*, 2009).

2.1. Production and verification of Y2H datasets

Fields and Song (1989) first described the Y2H system as the reconstitution of a transcription factor through expression of two hybrid proteins, one fusing the DNA-binding (DB) domain to a protein X (DB-X) and the other fusing an activation domain (AD) to a protein Y (AD-Y). In the last 20 years much has been learned about possible artifacts and appropriate controls, so that today Y2H can be considered not only one of the most efficient, but also one of the most reliable binary interaction assays available for small-, medium-, and large-scale interaction mapping. We next discuss specific artifacts of the Y2H system and the appropriate controls developed to detect and remove them.

2.1.1. Autoactivators

A common artifact of the Y2H system is autoactivation of Y2H-inducible reporter genes. This occurs when DB-X (where X is a full-length protein or a protein fragment) activates transcription of Y2H reporter genes irrespective of the presence of any AD-Y. Three classes of autoactivators need to considered: (i) genuine transcription factors that contain a *bona fide* AD and consequently

will likely score as autoactivators when fused to DB, (ii) proteins that are not transcription factors in their natural context but can behave as autoactivators because they contain a cryptic AD (cognate autoactivators), and (iii) nontranscription factor proteins that contain one or more cryptic ADs that are only functional as truncated fragments and not when expressed in the context of full-length proteins (*de novo* autoactivators).

Both genuine transcription factors and cognate DB-X autoactivators can be identified and removed by performing prescreens for reporter gene activation either with AD expressed alone (i.e., in the absence of any Y fused protein) or even with no AD at all.

De novo autoactivators are more difficult to detect than transcription factors and cognate autoactivators. The Y2H system is based on positive growth selections for potentially rare events, such as the finding of a single cDNA out of a complex library. The Y2H system can just as rigorously select for mutations that occur during the course of a screen and which convert a nonactivator protein into a *de novo* autoactivator. Such events are relatively frequent and some early Y2H datasets may have been inadvertently overpopulated by spontaneous autoactivators (Ito *et al.*, 2001; Yu *et al.*, 2008). A method to systematically remove these artifacts (Walhout and Vidal, 1999) employs a counter-selectable marker *CYH2* present on the AD-Y coding plasmid together with control plates that contain cycloheximide (CHX). At every stage of the interactome mapping pipeline reporter gene activity is evaluated in parallel both on regular selective plates and on selective plates that contain CHX. The *CYH2* marker allows the selection of yeast cells that do not contain any AD-Y and thus the convenient identification of DB-X autoactivators.

2.1.2. Retesting to verify candidate interactions

In addition to autoactivating mutations in the DB-X protein, other genetic changes can occur during a screen. Mutations of the full-length DB-X or AD-Y protein might permit interactions that are otherwise undetectable or inhibited. Other mutations, such as *cis*-acting mutations in reporter genes and *trans*-acting mutations at unlinked genetic loci, could lead to reporter gene activation in the absence of any physical interaction between DB-X and AD-Y. To identify and remove such artifacts, all interaction candidates are systematically verified using yeast transformants freshly thawed from DB-X and AD-Y archival stocks. Haploid yeast cells of opposite mating-type, each containing DB-X or AD-Y expression plasmids, are mated according to the interacting pairs identified in the original screens and are tested for reproducible Y2H phenotypes to confirm reporter gene activation. Usually $\sim$50% of interaction candidates can be successfully verified, which suggest that perhaps half of all primary Y2H positives belong to the classes of artifacts described above.

2.1.3. A high-quality Y2H implementation

Besides the precautions already mentioned, the Y2H version we have developed presents the following features that ensure high data quality.

Low DB-X and AD-Y hybrid protein expression The use of low copy number yeast expression vectors together with the presence of weak promoters expressing DB-X and AD-Y hybrid proteins leads to low expression, which minimizes artifactual interactions driven by mass action. Use of high copy number vectors can increase DB-X and AD-Y protein expression and increase the sensitivity of the assay. This comes at the cost of increasing the detection of unspecific interactions (Braun *et al.*, 2009). The use of high copy number vectors should be accompanied by rigorous quality control and validation of every individual interaction with multiple assays.

Yeast strains We have used two different Y2H strain backgrounds over the years (Vidal *et al.*, 1996; Yu *et al.*, 2008). The protocols described are applicable to Y8800 and Y8930, *MAT***a** and *MAT*α, respectively, two strains derived from PJ69-4 (James *et al.*, 1996) which harbor the following genotype: *leu2-3,112 trp1-901 his3-200 ura3-52 gal4*Δ *gal80*Δ *GAL2-ADE2 LYS2::GAL1-HIS3 MET2::GAL7-lacZ* $cyh2^R$. The availability of two haploid strains of opposite mating types enables the use of mating to efficiently combine large collections of DB-X and AD-Y constructs. By convention the Y8800 *MAT***a** and Y8930 *MAT*α strains are transformed with AD-Y and DB-X constructs, respectively.

Y2H-inducible reporter genes The reporter genes *GAL2-ADE2* and *LYS2::GAL1-HIS3* are integrated into the yeast genome. Expression of the *GAL1-HIS3* reporter gene should be tested with 1 m*M* 3AT (3-amino-1,2,4-triazole, a competitive inhibitor of the *HIS3* gene product). When dealing with DB-X autoactivators, higher 3AT concentrations can be used to circumvent autoactivator-dependent activity of *GAL1-HIS3*. Interactions identified at higher 3AT concentrations should be accompanied by rigorous quality control and validation of every individual interaction using multiple assays.

Y2H controls Y2H-inducible reporter gene expression levels can vary from weak to very strong, although these levels may not reflect the actual affinity of protein–protein interactions as they take place in their native environment. To help determine which candidate clones likely represent genuine biophysical interactors, six controls are added systematically to the master plates of Y2H experiments (Walhout and Vidal, 2001). This collection of diploid control strains contains plasmid pairs expressing DB-X and AD-Y hybrid proteins across a wide spectrum of interaction read-outs. For each control strain, a short description of plasmids and DB-X and AD-Y

hybrid proteins are provided in Table 12.1 and expected phenotypes are shown in Fig. 12.2.

Please email "pascal_braun@dfci.harvard.edu" to request strains, plasmids, and controls.

2.2. Validation of Y2H datasets to produce reliable binary interactome maps

Despite the rigorous implementation of controls for identification of technical artifacts, a fraction of technical false positives can still be recovered in large-scale datasets. Well-described artifacts might have escaped detection, or it is possible that certain classes of artifacts have not been identified yet and consequently no controls are available to detect and remove them. Therefore, the quality of any dataset must be further assessed before it can be used as a reliable interactome map.

In earlier attempts at addressing this question for Y2H, protein pairs that activated two or more distinct reporters or pairs that were detected in two or more configurations (e.g., DB-X/AD-Y and DB-Y/AD-X) were

Table 12.1 Y2H controls

	Plasmid pairs	Protein	Interaction strength
Control 1	pDEST-AD pDEST-DB	No insert No insert	None, background
Control 2	pDEST-AD-E2F1 pDEST-DB-*CYH2*-pRB	Human E2F1 aa 342–437 Human pRB aa 302–928	Weak (control for CHX control plates)
Control 3	pDEST-AD-Jun pDEST-DB-Fos	Mouse Jun aa 250–325 Rat Fos aa 132–211	Moderately strong
Control 4	pDEST-AD pDEST-DB-Gal4	No insert Yeast Gal4 aa 1–881	Very strong
Control 5	pDEST-AD-dE2F1 pDEST-DB-dDP	*Drosophila* E2F aa 225–433 *Drosophila* DP aa 1–377	Strong
Control 6	pDEST-AD-*CYH2*-dE2F1 pDEST-DB-dDP	*Drosophila* E2F aa 225–433 *Drosophila* DP aa 1–377	Strong (control for CHX plates)

Identities and description of expected phenotypes for the six controls used in every Y2H experiment.

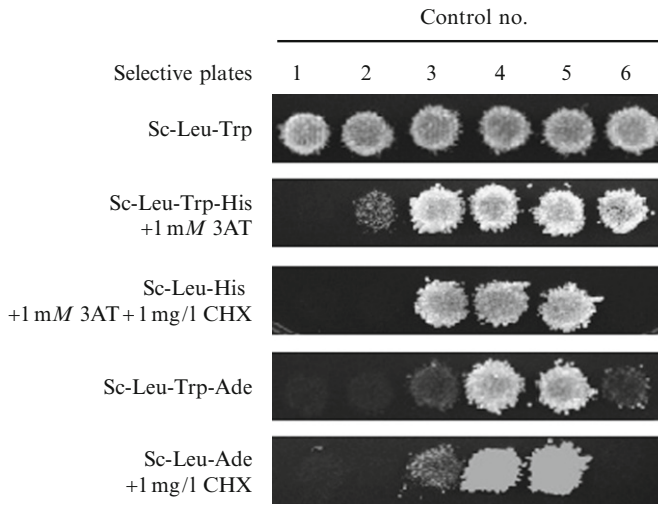

Figure 12.2 Phenotypes of Y2H controls. Six strains referred to as Control 1–6, each containing a different pair of DB-X and AD-Y hybrid proteins, are spotted on media selecting for the presence of both plasmids (top row) and, after an overnight incubation, replica-plated onto media selecting for Y2H-dependent reporter activation (rows 2 and 4). The six strains express DB-X/AD-Y pairs that result in reporter gene activation at various intensities. DB-X autoactivation is tested on plates that select for the loss of the AD-Y plasmid (rows 3 and 5).

considered to be of "higher quality", that is, more likely to be real biophysical interactors, than those pairs that activated only one reporter or were only found in a single orientation. Historically, and especially with cDNA screens, these criteria did indeed offer limited protection against artifacts, and enabled identification of more likely "true" interactions (Vidalain *et al.*, 2003). Today, however, such artifacts can be removed more systematically and more reliably by the controls described in Sections 2.1.1 (CHX control) and 2.1.2 (verification). All interactions that pass these controls are considered high-quality Y2H interactions, irrespective of whether or not they are detected in only one orientation or if they activate only one reporter.

Many "true" interaction pairs activate only one Y2H reporter or are detected in only one configuration. This is due to effects that are unrelated to the interaction capacity of the two examined proteins. The genomic context of the different reporters or use of promoters that require different levels of reconstituted transcription factor can lead to differential reporter activation. Similarly, the use of hybrid proteins imposes steric constraints on proteins that can interfere with detection of many interactions in at least one configuration. This was shown by testing a set of well documented positive control interaction pairs in Y2H and four other binary interaction assays. Consistently, only half of the positive scoring controls were detected in

both orientations in any of the five assays (Braun *et al.*, 2009). Thus, while activation of multiple reporters and detection of interactions in multiple configurations can be comforting, these attributes are neither necessary nor sufficient requirements for high quality interactions.

Various experimental methods and computational approaches have been described to evaluate the quality of large-scale interactome datasets. Most computational methods estimate the correlation between physical interaction data and secondary data, such as expression profiling or types of functional annotations (Bader *et al.*, 2004; Deane *et al.*, 2002). Determination of data quality using this approach can effectively lead to filtered datasets that might be biased for particular classes of interactors, such as those with strong coexpression correlation. Such correlative data evaluation approaches make implicit assumptions about the nature of protein–protein interactions, which can potentially lead to erroneous conclusions (Yu *et al.*, 2008). Interactome maps can be productively integrated with orthogonal datasets to gain novel insights into biology (Pujana *et al.*, 2007; Vidal, 2001). If interaction datasets have been prefiltered using orthogonal data then such higher level analysis becomes less informative.

Another approach is to overlap the information from different interaction datasets to assess the FDR. In these analyses crucial details of the underlying experiments used in the respective screens are often ignored. Four critical parameters have to be taken into consideration (Venkatesan *et al.*, 2009): (i) the number and identity of ORFs used in each screen (*completeness*), (ii) the detection limitations of the assays used (*assay sensitivity*) (affected by many parameters like strains, location of protein tags, detection methods), (iii) the extent of incomplete sampling in each search space (*sampling sensitivity*), and (iv) the potential presence of technical false positives (*precision*). Without knowledge of these parameters for each dataset, any conclusion about data quality based on their overlap is meaningless. Thus, given the inherent limitations of computational approaches for quality control, experimental methods involving alternative protein interaction assays are strongly preferred.

2.2.1. Quality control I: Experimental assessment of dataset *precision*

One experimental approach to validate dataset quality consists in testing a representative sample of potential interactions from a given dataset with an orthogonal interaction assay. Since there is apparently not a single assay capable of detecting all protein–protein interactions tested, and considering that the subset of interaction pairs scoring positive in any two assays is rarely identical (Braun *et al.*, 2009), it is to be expected that only a fraction of interactions from a particular dataset will be detected by a validation assay. This is a consequence of the nature of interaction assays and the biochemical diversity of interactions, and not *per se* an indication of the quality of

the original dataset. Characterizing the validation assay of choice using a positive reference set (PRS) and random reference set (RRS) of well-documented and random protein–protein interactions, respectively (Braun *et al.*, 2009; Cusick *et al.*, 2009; Venkatesan *et al.*, 2009; Yu *et al.*, 2008) provides an estimate of the *assay sensitivity* and the background of the validation assay. If desired, the stringency of the validation assay can be adjusted to decrease the background (and *assay sensitivity*) or to increase the *assay sensitivity* (and background). The validation assay results of the dataset sample relative to the PRS/RRS benchmark data enables estimation of the dataset *precision* (Braun *et al.*, 2009; Venkatesan *et al.*, 2009; Yu *et al.*, 2008).

2.2.2. Quality control II: Experimental confidence scores for individual interactions

When dataset precision is determined using a single assay, validation rates between 20% and 40% can be expected for both PRS and high-quality datasets under conditions in which the RRS detection rate is below 5% (Braun *et al.*, 2009). In the long term, it will be highly desirable to not only estimate the overall precision of a dataset but to validate all protein–protein interactions individually. Validation for individual interactions in a dataset can be made stronger if multiple complementary assays are used to test the interactions (Braun *et al.*, 2009). The concept of calibrating and benchmarking assay performance with the PRS/RRS can be applied to multiple assays and can be used to calculate a confidence score for individual biophysical protein–protein interactions. Multiple interaction assays are first benchmarked against common PRS and RRS reference sets to obtain comparable calibrations of *assay sensitivity* and background. Then, all interactions identified in a large-scale interactome screen are characterized using the same assay implementations. After the results from all assays have been collected for any interaction pair, a confidence score can be calculated based on prior PRS/RRS calibration of the assays and the validation results of the respective interaction (Braun *et al.*, 2009). PRS/RRS clones for several organisms are available upon request (Braun *et al.*, 2009; Simonis *et al.*, 2009; Venkatesan *et al.*, 2009; Yu *et al.*, 2008).

2.3. Biological evaluation of binary interactome maps

The identification of high-confidence biophysical interactions is an important first step towards answering many biological questions both at small and large scale. However, even robustly demonstrated biophysical interactions might be biological false positives, or pseudo-interactions, that never occur *in vivo*.

Biological relevance of protein–protein interactions has been inferred from network analyses or by combining interactome information with systematic genetic data (Collins *et al.*, 2007a,b; Pujana *et al.*, 2007). Despite some success, these approaches remain constrained by the availability of high-quality datasets, and are limited as they are predictions.

Until demonstrated by thorough mechanistic studies of all proteins involved, the biological role of protein–protein interactions remains elusive. Such mechanistic studies are typically carried out at small scale, so this approach is unsustainable and cost prohibitive for characterizing hundreds of thousands of soon to be discovered human protein interactions (Venkatesan *et al.*, 2009).

Biological relevance of a biophysical protein–protein interaction may be derived from an observed phenotype following genetic disruption of this specific interaction *in vivo*. Such IDAs can occur naturally, as has been found for some inherited Mendelian diseases (De Nicolo *et al.*, 2009; Zhong *et al.*, 2009). For these alleles the causal link between disruption of the biophysical interaction and the observable phenotype must be demonstrated. Alternatively, IDAs can be generated experimentally using a reverse two-hybrid approach (Dreze *et al.*, 2009; Vidal *et al.*, 1996). For defining a biological function of a biophysical interaction using such experimentally generated IDAs, a critical step is the identification of a phenotype and subsequent demonstration of causality.

Certain interactions may have subtle or modifying roles in the regulation of cellular functions. Disruption of such interactions individually may lead to subtler, easy to overlook phenotypes. Disruption of such interactions in the presence of other genetic or environmental perturbations may produce more observable systems alterations. For those, quantitative mathematical modeling may be useful for analyzing small or synergistic phenotypic consequences.

3. High-Throughput Y2H Pipeline

3.1. Assembly of DB-X and AD-Y expression plasmids

The first step towards binary interactome mapping is the generation of expression plasmids. For high-throughput experiments it is preferable to use sequence independent recombinational subcloning technologies such as Gateway cloning (Walhout *et al.*, 2000). Large resources containing thousands of distinct ORFs in Gateway entry vectors are available for a few organisms (Lamesch *et al.*, 2004, 2007; Reboul *et al.*, 2003; Rual *et al.*, 2004). These ORFs can be transferred into Gateway-compatible expression vectors in a simple single-step reaction (Fig. 12.3). Albeit not mandatory, linearizing the destination vectors by restriction digestion improves recombination efficiency and decreases background as well as chances of obtaining incorrect LR recombination clones. The restriction enzyme should be chosen so that the destination vector is digested only once between the two Gateway recombination sites. The Gateway LR reaction, carried out using enzyme and buffer concentrations optimized by titration, gives best yields at 25 °C for 18 h but can also be carried out for ~2 h at room

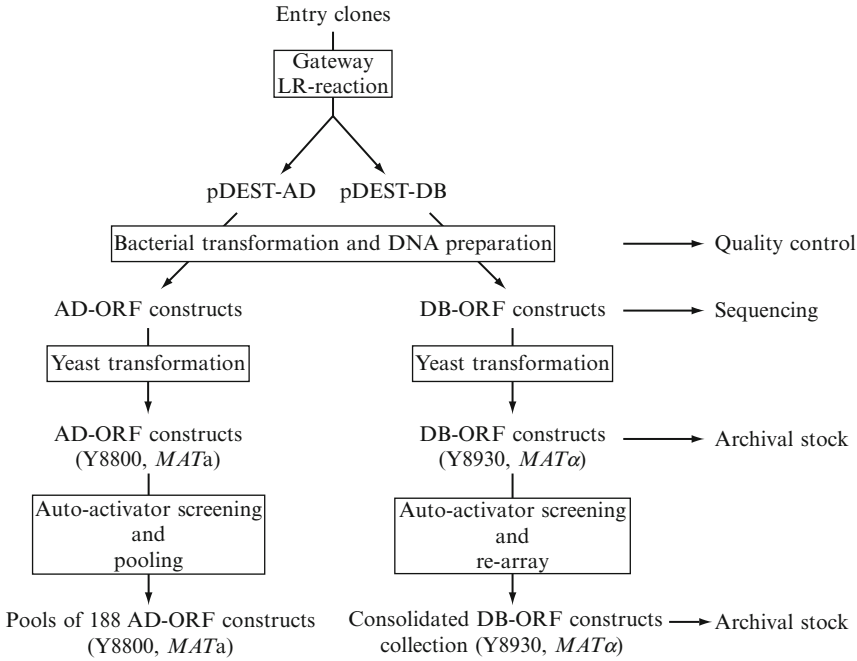

Figure 12.3 **Pipeline for preparation of Y2H reagents. The pipeline from producing Gateway entry clones to transformation and quality control of yeast strains used in Y2H screens. Protein-encoding ORFs are first transferred by Gateway LR reactions into pDEST-AD and pDEST-DB, and amplified in bacteria. DNA is then extracted for yeast transformations. After transformation DB-X hybrid proteins are tested for autoactivator phenotypes and then rearrayed before screening. AD-Y hybrid proteins are combined into minipools of 188 different clones per pool.**

temperature. Completed recombination reactions are transformed into *Escherichia coli*, grown for 18 h, and plasmids are isolated. This step can be done manually or by using liquid handling robots. Because all steps are carried out in 96-well microtiter plates, protocols are provided for the equivalent of one 96-well plate.

Protocol 1: Restriction digestion of Y2H destination vectors

1. Combine in one tube:
   - 11 μg destination vector (pDEST-AD or pDEST-DB).
   - 11 μl of 10× restriction enzyme buffer.
   - 2.5 μl of *Sma*I restriction enzyme (50 units).
   - 85.5 μl filter-sterilized water.
2. Mix well by pipetting up and down several times.
3. Incubate at 25 °C for 12–16 h.
4. Incubate at 65 °C for 20 min to heat-inactivate the restriction enzyme.

Display 500 ng of digested destination vector on a 1% agarose gel alongside 500 ng of undigested destination vector to confirm complete digestion. The heat-inactivated reaction mix can be used for Gateway LR reactions without further purification.

Protocol 2: High-throughput Gateway LR recombinational cloning

1. Combine in one tube:
   - 110 μl of *Sma*I digested destination vector (11 μg).
   - 110 μl of LR clonase buffer 5×.
   - 55 μl of TE 1×.
   - 55 μl of LR clonase enzyme mix (Invitrogen) (keep this mix on ice).
2. Homogenize by gently pipetting up and down.
3. With a multichannel pipette, distribute 3 μl of this solution into every well of a 96-well microtiter plate.
4. Add 2 μl of entry clone per well.
5. Centrifuge briefly.
6. Incubate at 25 °C for 18 h.

Protocol 3: Bacterial transformation

The following protocol is used to transform, amplify, and isolate Gateway LR reaction products:

1. Thaw 1 ml of competent DH5α-T1^{R} (Invitrogen) cells on ice (with a transformation efficiency greater than 5×10^{7} antibiotic resistant colonies per μg of input DNA).
2. Add 10 μl of competent cells per well directly into a 96-well plate containing 5 μl Gateway LR reaction mix in each well.
3. Seal the plate with adhesive foil.
4. Incubate on ice for 20 min.
5. Heat shock at 42 °C in a standard thermocycler for 1 min.
6. Incubate on ice for 2 min.
7. Add 100 μl of prewarmed (37 °C) SOC media per well. Seal the plate with adhesive foil to avoid contamination.
8. Incubate at 37 °C for 1 h.
9. Transfer the transformation mix into a 96-well deep-well plate containing 1 ml of LB media with 100 μg/ml of ampicillin.
10. Incubate on a 96-well plate shaker at 37 °C for 20 h.
11. Remove 5 μl for subsequent analysis by PCR (Protocol 4).
12. Remove 80 μl of the overnight culture, mix with 80 μl of 40% (w/v) autoclaved glycerol and store at −80°C.
13. Use the remainder of the overnight culture for plasmid isolation.

SOC medium 0.5% yeast extract, 2% tryptone, 10 m*M* NaCl, 2.5 m*M* KCl, 10 m*M* $MgCl_2$, 10 m*M* $MgSO_4$, 20 m*M* glucose. Add glucose after autoclaving the solution with the remaining ingredients, and let cool down. Sterilize the final solution by passing it through a 0.2 μm filter. SOC medium can be stored at room temperature.

Transformation controls It is good practice to systematically control for media contamination (no cells), competent cells contamination (cells only), Gateway LR reaction contamination (negative control of LR reaction), and transformation efficiency (10 pg of pUC19). If the four controls indicate clean and successful transformation, proceed to the next quality control step.

Recombination control To confirm that Gateway LR reactions occurred properly, analyze recombination products by bacterial culture PCR using destination vector specific primers (Protocol 4). For each transformation plate select one row for PCR.

Protocol 4: Bacterial culture PCR

Dilute 5 μl of bacterial culture into 95 μl of sterile water and mix by pipetting up and down. Keep bacterial cultures at 4 °C until PCR results are determined.

For one 96-well plate of PCR, prepare in a tube on ice:

- 330 μl of HiFi Platinum Taq polymerase buffer 10× (Invitrogen).
- 120 μl of 50 m*M* $MgSO_4$ (final concentration 1.8 m*M*).
- 33 μl of 40 μ*M* dNTPs (final concentration 400 n*M*).
- 3.3 μl of 200 μ*M* AD or DB forward primer (final concentration 180 n*M*).
- 3.3 μl of 200 μ*M* Term reverse primer (final concentration 180 n*M*).
- 20 μl of HiFi Platinum Taq polymerase (Invitrogen).
- 2.5 ml of filter-sterilized water.

Aliquot 30 μl of the reaction mix into every well of a soft shell, V-bottom 96-well microtiter plate. Keep on ice. Add 3 μl of the diluted bacterial culture per well as DNA template. Wells G12 and H12 are used as negative control (water as template) and positive control (10 ng of empty pDEST-AD or pDEST-DB), respectively.

Place the PCR plate on a thermocycler and run the following program:

Step 1: Denaturation at 94 °C for 4 min.
Step 2: Denaturation at 94 °C for 30 s.
Step 3: Annealing at 58 °C for 30 s.
Step 4: Elongation at 68 °C for 3 min.
Repeat Steps 2-3-4, 34 times.
Step 5: Final elongation at 68 °C for 10 min.
Step 6: Hold at 10 °C.

Primer sequences The primers are designed such that the 5′-primer confers AD and DB vector specificity, respectively, whereas the Term 3′-primer is identical for both vectors.

AD: 5′-CGCGTTTGGAATCACTACAGGG-3′
DB: 5′-GGCTTCAGTGGAGACTGATATGCCTC-3′
Term: 5′-GGAGACTTGACCAAACCTCTGGCG-3′

Once PCR reactions are completed, analyze 5 μl of PCR product on a 1% agarose gel by comparing sizes to that of the control from well H12 (the PCR amplicon from a destination vector containing the Gateway cassette has an expected size $\sim$1.9 kb). The H12 control serves simultaneously as a positive control for the PCR and as a negative control for the LR recombination reaction. PCR failure is indicated by the absence of the H12 product, and failure of the LR reaction may be indicated by a dominant band of 1.9 kb across all wells. Successful LR reactions will give rise to the size distribution of the original ORFs to which $\sim$280 bp of vector sequences are added due to the AD, DB, and Term primer positions. If the PCR results indicate successful LR recombinations, prepare archival stocks by combining 80 μl of bacterial cultures with 80 μl of 40% (w/v) glycerol in a round-bottom 96-well microtiter plate. The rest of the cultures are used for plasmid isolation. Afterwards, ensure successful plasmid isolation by analyzing 2 μl of the DNA preparation on a 1% agarose gel.

To ensure the absence of plate orientation mistakes when processing multiple plates, sequence verify PCR products amplified from one column of each 96-well miniprep plate. Use 1 μl of the DNA preparation as template for PCR. The primers, recipes, and PCR conditions are identical to those presented in Protocol 4. BLASTn of the acquired sequences against a reference database identifies clones and allows verification of their correct locations.

3.2. Yeast transformation

DB-X and AD-Y expression plasmid constructs are individually transformed into competent Y8930 (*MAT*α) and Y8800 (*MAT***a**) strains, respectively.

Protocol 5: Yeast transformation

This protocol requires two solutions that need to be freshly prepared from stock solutions in order to obtain maximum transformation efficiencies. Tris–EDTA–lithium acetate solution (TE/LiAc) is prepared by 10-fold dilution of 10$\times$ TE and 1 *M* LiAc stocks to give 10 m*M* Tris–HCl (pH 8.0), 1 m*M* EDTA (pH 8.0), and 100 m*M* LiAc final concentration. TE/LiAc polyethyleneglycol (PEG) solution is prepared by combining 8 volumes of 44% (w/v) PEG 3350 with 1 volume of 10$\times$ TE and 1 volume

of 1 *M* LiAc. The following volumes and quantities are given for carrying out one 96-well plate of transformations.

1. Streak Y8800 and Y8930 on separate YEPD plates and incubate at 30 °C for 48–72 h to obtain isolated colonies.
2. For each strain, inoculate 20 ml of YEPD with 10 isolated colonies. Incubate at 30 °C on a shaker for 14–18 h.
3. Measure and record the OD_{600}, which should be between 4.0 and 6.0. Dilute cells into YEPD media to obtain a final $OD_{600} = 0.1$. Use 100 ml of YEPD media per 96-well plate of transformations.
4. Incubate at 30 °C on a shaker until OD_{600} reaches 0.6–0.8 (4–6 h).
5. Boil carrier DNA (salmon sperm DNA, Sigma-D9156) for 5 min then place on ice until needed.
6. Harvest cells by centrifugation at 800×*g* for 5 min. Discard the supernatant and resuspend cells gently in 10 ml of sterile water.
7. Centrifuge as described in step 6 and discard the supernatant.
8. Resuspend cells in 10 ml of TE/LiAc solution, centrifuge, and discard the supernatant.
9. Resuspend cells in 2 ml of TE/LiAc solution, then add 10 ml of TE/LiAc/PEG solution supplemented with 200 μl of boiled carrier DNA. Mix the solution by inversion.
10. Dispense 120 μl of this mix into each well of a round-bottom 96-well microtiter plate.
11. Add 10 μl of plasmid DNA to the competent yeast and mix by pipetting up and down several times. Use liquid handling robots to transfer and mix 96 samples at a time. Seal the plate with adhesive foil.
12. Incubate at 30 °C for 30 min.
13. Subject to heat shock in a 42 °C water bath for 15 min.
14. Centrifuge the 96-well plate for 5 min at 800×*g*. Carefully remove the supernatant using a multichannel pipette.
15. To each well add 100 μl of sterile water and resuspend cell pellets by pipetting up and down.
16. Centrifuge the 96-well plate for 5 min at 800×*g*, then carefully remove 90 μl of water from each well using a multichannel pipette.
17. Resuspend cell pellets by vortexing the 96-well plate on a shaker.
18. Spot 5 μl/well of cell suspension onto an appropriate selective plate (Sc-Trp for AD-Y, Sc-Leu for DB-X). For a consistent footprint, use of a liquid handling robot is recommended.
19. Incubate at 30 °C for 72 h.
20. Using sterile flat-end toothpicks, pick transformed yeast colonies into individual wells of a 96-well round-bottom plate containing 160 μl of selective media (Sc-Trp for AD-Y, Sc-Leu for DB-X).
21. Incubate on a shaker at 30 °C for 72 h.

22. Prepare archival stocks by combining 80 μl of the yeast culture with 80 μl of 40% (w/v) autoclaved glycerol in a round-bottom 96-well plate. Store at −80 °C.

3.3. Autoactivator removal and AD-Y pooling

3.3.1. Autoactivator identification and removal

To be as close as possible to the physiology of the cells in which interactions are detected, the identification of autoactivators is achieved in diploid yeast strains obtained by mating DB-X yeast strains with the Y8800 yeast strain transformed with the AD encoding plasmid containing no insert (empty pDEST-AD). All diploid yeast strains showing a growth phenotype stronger than the "no interaction" Y2H control (control 1) are considered autoactivators. Because activation of the *GAL1::HIS3* reporter gene is easier to achieve than that of *GAL7::ADE2*, it is used for autoactivator identification.

Protocol 6: Identification of autoactivating DB-X hybrid proteins

Before starting the experiment, prepare one YEPD plate, one Sc-Leu-Trp plate and one Sc-Leu-Trp-His + 1 m*M* 3AT plate for each 96-well plate of DB-X yeast strains to be tested. The YEPD plates should be prepared at least 1 week in advance to allow them to dry. This allows fast penetration of liquid in the mating step and prevents merging of adjacent spots due to excess liquid.

1. Add 160 μl of fresh liquid Sc-Leu media to each well of a round-bottom 96-well microtiter plate followed by 5 μl from individual glycerol stocks of each DB-X yeast strain to each well.
2. For every plate of DB-X yeast strains to be tested, inoculate a test tube containing 0.55 ml of Sc-Trp media with Y8800 transformed with empty pDEST-AD.
3. Incubate at 30 °C for 72 h on a shaker.
4. Spot 5 μl of DB-X liquid cultures on a YEPD plate using a liquid handling robot.
5. Allow the spots to dry for 30–60 min.
6. Aliquot the pDEST-AD transformed Y8800 yeast culture into a round-bottom 96-well plate.
7. Spot 5 μl of pDEST-AD transformed Y8800 on top of the DB-X spots.
8. Spot Y2H controls at the bottom of the plate.
9. Incubate mating plates at 30 °C for 14–18 h.
10. Replica-plate onto Sc-Leu-Trp media to select for diploid cells.
11. Incubate at 30 °C for 14–18 h.
12. Replica-plate from the Sc-Leu-Trp media onto Sc-Leu-Trp-His + 1 m*M* 3AT media. Nonautoactivating yeast cells are not able to activate the *GAL1::HIS3* reporter gene hence should not grow on this media.

13. Incubate at 30 °C for 14–18 h.
14. "Replica-clean" Sc-Leu-Trp-His + 1 m*M* 3AT plates by placing each plate on a piece of velvet stretched over a replica-plating block and pushing evenly on the plate to remove excess yeast. Replace the cloth and move to process the next plate until all plates have been cleaned.
15. Incubate at 30 °C for 72 h.
16. Score growth phenotypes.

Growth phenotypes are scored by comparison to the "no interaction" Y2H control (control 1). All yeast strains showing a stronger growth phenotype than control 1 are considered autoactivators. To reliably identify autoactivators it is best to score growth twice independently. If a yeast clone is given two different scores, accepting the most stringent one ensures high quality of the starting material for subsequent interactome mapping.

Autoactivators are physically removed from the collection of DB-X transformed yeast clones by robotic rearraying of nonautoactivator yeast clones into new plates. During the rearray step plate positions G12 and H12 are left empty for control purposes. New glycerol stocks are prepared from this consolidated collection of nonautoactivating yeast strains and used for subsequent Y2H screens.

1. From archival glycerol stocks containing all of the individual DB-X yeast clones, cherry pick nonautoactivating DB-X yeast clones into plates containing 160 μl Sc-Leu (DB-X) liquid media.
2. Incubate at 30 °C for 72 h on a 96-well plate shaker.
3. Prepare an archival stock by combining 80 μl of the yeast culture with 80 μl of 40% (w/v) autoclaved glycerol in a round-bottom 96-well plate.

Albeit much less frequent, autoactivating AD-Y can also occur. The previous protocol can easily be adapted for AD-Y autoactivator identification by use of AD specific reagents wherever appropriate. As an alternative, to reduce time and cost, it is possible to test AD-Y autoactivation using pools. For this, each AD pool, described in Protocol 7, is mated with Y8930 transformed with the DB encoding plasmid containing no insert (empty pDEST-DB) then processed (Protocol 6). If a diploid strain shows growth on autoactivator detection plates, the responsible AD-Y yeast clone can be identified by deconvoluting the AD-Y pool. This step is achieved by testing all 188 AD-Y yeast clones constituting the pool for autoactivation. Once identified, the autoactivating AD-Y yeast clone is removed and the affected pool is reassembled without it (Protocol 7).

3.3.2. Efficient screening by AD pooling

The pools used in the Y2H pipeline combine 188 different AD-Y expressing yeast clones. This experimentally defined pool-size provides an optimal compromise between screening efficiency (number of plates to be processed) and screen sensitivity (number of interactors identified).

Protocol 7: Construction of AD pools

This protocol describes the construction of one pool of 188 different AD-Y hybrid constructs transformed into Y8800, starting from two 96-well plates of AD-Y yeast strains.

1. For each of the two plates of 94 AD-Y constructs: inoculate 500 μl per well of Sc-Trp media with 5 μl per well of AD-Y yeast strains.
2. Grow on a shaker at 30 °C for 4 days.
3. Resuspend yeast cell cultures by thoroughly vortexing the culture plates.
4. Measure the OD_{600} to ensure that growth is homogenous throughout each plate, hence that each AD-Y yeast strains will be represented in the same proportion.
5. Transfer the contents of the two culture plates into a sterile trough.
6. Mix thoroughly to ensure equal representation of all AD-Y yeast strains in the pool.
7. On a liquid handling platform, prepare archival stocks by combining 80 μl of the pooled yeast cultures with 80 μl of 40% (w/v) autoclaved glycerol in round-bottom 96-well microtiter plates.

If additional copies of the AD-Y pools are required, these should be made according to the protocol above and not by amplification of existing pools, as amplification can lead to loss of representation within the pool.

Protocol 8: Assessing equal representation of AD-Y clones in pools

Before the AD pools are used for Y2H experiments, equal representation of each of the 188 AD-Y clones in the pools should be confirmed. Biased pools and low representation of some AD-Y yeast cells will decrease if not eliminate the ability to detect protein interactions involving the underrepresented hybrid proteins.

1. For each plate of AD-Y pools streak 5 μl of glycerol stock from two randomly selected wells onto Sc-Trp plates.
2. Grow at 30 °C for 72 h.
3. From each Sc-Trp plate, pick 96 isolated colonies and lyse yeast cells according to Protocol 9.
4. Add 3 μl of the yeast cell lysate as PCR template.
5. Carry out PCR according to Protocol 10.
6. Run 5 μl of PCR product on a 1% agarose gel. If PCR products can be detected in most reactions, proceed to analyze the corresponding plate by end-read sequencing.
7. Identify the obtained sequences by BLASTn. If the created pools are not biased, the frequency at which yeast cells containing identical AD-Y plasmids were picked (and hence the sequence identifications) should follow a normal distribution.

Protocol 9: Yeast cell lysis

1. Prepare lysis buffer by dissolving 2.5 mg/ml zymolase 20T (21,100 U/g, Seikagaku Corp.) in 0.1 *M* sodium phosphate buffer (pH 7.4). Keep on ice.
2. Aliquot 15 μl of lysis buffer into the wells of a 96-well PCR plate. Keep on ice.
3. Pick a small amount of yeast cells (not more than fits on the very end of a standard 200 μl tip) and resuspend in the lysis buffer in the PCR plate.
4. Place the PCR plate on a thermocycler and run the following program:
   Step 1: 37 °C for 15 min
   Step 2: 95 °C for 5 min
   Step 3: Hold at 10 °C
5. Add 100 μl of filter-sterilized water to each well.
6. Centrifuge 10 min at 800×*g*.
7. Store at −20 °C.

Protocol 10: Yeast lysate PCR

For each 96-well plate of PCR reactions, prepare the following reaction mix on ice:

- 330 μl of HiFi Platinum Taq (Invitrogen) polymerase buffer 10×.
- 120 μl of 50 m*M* $MgSO_4$ (final concentration 1.8 m*M*).
- 33 μl of 40 μ*M* dNTPs (final concentration 400 n*M*).
- 3.3 μl of 200 μ*M* AD primer (final concentration 180 n*M*).
- 3.3 μl of 200 μ*M* Term primer (final concentration 180 n*M*).
- 20 μl of HiFi Platinum Taq polymerase (Invitrogen).
- 2.5 ml of filter-sterilized water.

Aliquot 30 μl into every well of a 96-well PCR plate. Keep on ice. To each well, add 3 μl of the yeast cell lysate (Protocol 9) as DNA template. Seal plate with adhesive aluminum foil.

Place the PCR plate on a thermocycler and run the following program:

Step 1: Denaturation at 94 °C for 4 min.
Step 2: Denaturation at 94 °C for 30 s.
Step 3: Annealing at 58 °C for 30 s.
Step 4: Elongation at 68 °C for 3 min.
Repeat Step 2-3-4, 34 times.
Step 5: 68 °C for 10 min.
Step 6: Hold at 10 °C.

3.4. Screening and phenotyping

The Y2H pipeline consists of three essential stages, which together yield highly reliable interactions: primary screening, secondary phenotyping, and verification (Fig. 12.4). The high-throughput Y2H pipeline presented here

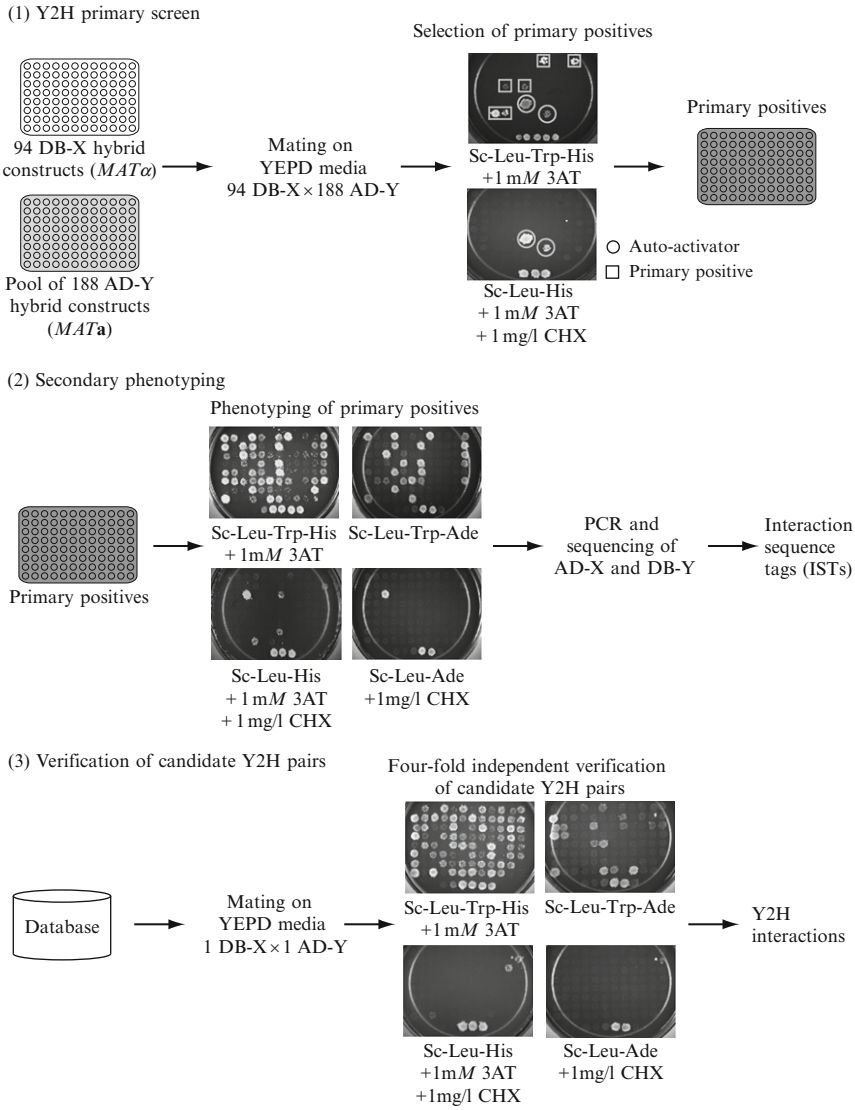

Figure 12.4 Y2H screening pipeline. Three steps, primary screening, phenotyping, and retesting, ensure high-throughput and reliable removal of artifacts. For primary screens, 94 distinct DB-X constructs are mated against a minipool containing 188 AD-Y hybrids. Positive colonies are picked from selective plates and in ''secondary phenotyping'' are evaluated on two types of selective plates and respective autoactivation control plates. Protein pairs considered as ''candidate Y2H interactions'' are identified by DNA sequencing of PCR products amplified from positive colonies. All identified pairs are verified using fresh archival yeast stocks. DB-X/AD-Y pairs that score positive on at least three out of four independent plate sets are considered high-quality Y2H interactions (see text for details).

has been used to produce several high-quality proteome-scale binary interactome maps (Rual *et al.*, 2005; Simonis *et al.*, 2009; Yu *et al.*, 2008).

Protocol 11: Y2H primary screening

Pour all required agar plates at least 1 week before starting the experiments and store them without wrapping at room temperature. Storage ensures that the plates are sufficiently dry, which in turn will prevent merging of spotted yeast cultures in the mating step which can otherwise occur due to excess liquid and slow absorption into the agar.

[Day 0: Inoculation]

1. Thaw glycerol stocks of the DB-X yeast strains and AD pools to be tested. One person can easily handle a batch of 100 mating plates, for example, 10 96-well plates of DB-X yeast clones tested against 10 96-well plates with AD-Y pools.
2. Inoculate 96-well plates that contain 160 μl selective media in every well (Sc-Leu for DB plates, Sc-Trp media for AD pool plates), with 5 μl/well of the thawed glycerol stock plates.
3. Seal all plates with adhesive tape and return glycerol stocks to -80°C.
4. Incubate the inoculated cultures at 30 °C on a shaker for 72 h.

[Day 3: Mating]

1. For each combination [AD-Y pool plate×96 DB-X plate] spot 5 μl/well of the respective AD-Y pool liquid culture onto a mating plate (YEPD) using a liquid handling robot.
2. Allow spots to dry for 30–60 min.
3. Spot 5 μl/well of each DB-X on top of the AD pool spots.
4. Spot Y2H controls onto every plate.
5. Incubate mating plates at 30 °C for 14–18 h.

[Day 4: Replica-plating]

1. Replica-plate mated yeast cells from mating plates onto screening plates (Sc-Leu-Trp-His + 1 m*M* 3AT).
2. To detect *de novo* autoactivators, for each distinct plate of DB-X yeast clones, replica-plate yeast from three mating plates (with three different AD pools) onto Sc-Leu-His + 1 m*M* 3AT + 1 mg/l CHX plates.
3. Incubate at 30 °C for 14–18 h.

[Day 5: Replica-clean]

1. Replica-clean all plates by placing each plate on a piece of velvet stretched over a replica-plating block and pushing evenly on the plate to remove excess yeast cells. Replace the cloth and move to process the next plate until all plates have been cleaned. The plates need to be cleaned enough to reduce background, but excessive cleaning can also lead to accidental removal of positives.

2. Incubate at 30 °C for 5 days.

[Day 10: Score and pick colonies]

Pick primary positive colonies from screening plates and resuspend in a 96-well plate containing liquid media (Sc-Leu-Trp). Only consider colonies that grew better than background as indicated by control 1 of the six Y2H controls (Fig. 12.2). Only pick primary positives where the corresponding spots on the CHX plates are negative. Consider all three CHX plates as controls. Since every individual DB-X construct is mated against a pool of 188 AD-Y constructs, it is possible to obtain multiple interactions per spot. To account for this infrequent yet possible event we pick at most three colonies per spot.

1. Pick positive yeast colonies into a 96-well plate containing 160 μl/well Sc-Leu-Trp media. Leave positions G12 and H12 empty for subsequent controls.
2. Incubate the culture plate at 30 °C for 72 h.
3. The cultures can be used directly for phenotyping (Protocol 12—start at Step 2). It is also recommended to prepare an archival glycerol stock by combining 80 μl of the yeast culture with 80 μl of 40% (w/v) autoclaved glycerol in a 96-well plate, sealing the plates with adhesive tape and storing at −80°C.

Protocol 12: Phenotyping

[Day 0: Inoculation]

1. Thaw glycerol stocks of primary positives.
2. Spot 5 μl/well onto Sc-Leu-Trp plates using a 96-well liquid handling robot.
3. Seal all glycerol stock plates with adhesive tape and return to −80°C.
4. Add Y2H controls.
5. Incubate the Sc-Leu-Trp plates at 30 °C for 48 h.

[Day 2: Replica-plating]

1. Replica-plate from Sc-Leu-Trp plates onto four phenotyping plates:
   - Sc-Leu-Trp-His + 1 m*M* 3AT
   - Sc-Leu-Trp-Ade
   - Sc-Leu-His + 1 m*M* 3AT + CHX (1 mg/l)
   - Sc-Leu-Ade + CHX (1 mg/l)

The first two plates are used to assess Y2H reporter activity; the two CHX plates enable detection of autoactivators.

2. Clean the plates immediately after replica-plating. This step will minimize background growth.
3. Incubate the phenotyping plates at 30 °C for 72 h.

[Day 5: Scoring]

1. Identify autoactivators by inspecting CHX plates. Any yeast spot showing growth on these plates should not be considered for further processing.
2. Identify candidate interactions (secondary positives). It is useful to differentiate positives activating one reporter gene (most often *GAL1::HIS3*) from those activating both reporter genes. An example of a Sc-Leu-Trp plate and the four assay plates along with proper scoring are shown (Fig. 12.5).
3. Patch all secondary positives on fresh Sc-Leu-Trp plates.
4. Incubate the Sc-Leu-Trp plates at 30 °C for 48 h.
5. Lyse cells according to Protocol 9.
6. Amplify the inserts of the DB-X and AD-Y inserts of positive colonies by yeast colony PCR according to Protocol 10 for subsequent ORF identification by end-read sequencing. At this stage the matched PCR products coding for putatively interacting proteins are physically separated. It is critical to track the matching AD-Y and DB-X PCR products so that interacting pairs can be identified after sequencing.

Once sequencing data have been received and the candidate protein pairs have been identified, a list of unique candidate interaction pairs can be compiled.

3.5. Verification

Protocol 13: Verification of candidate Y2H interaction pairs

While the CHX control at every step identifies spontaneous autoactivators arising from mutations in DB-X, this last verification step protects against other potential artifacts, for example, from mutations elsewhere in the yeast genome, and ensures robust high data quality. To reach maximum reproducibility, robustness, and reliability of Y2H interactions, this critical step is carried out a total of four times independently (16 plates corresponding to four sets of four assay plates), ideally by four different experimenters. Only interactions that score positive at least three out of four plate sets, and do not once score as autoactivators, are accepted as verified Y2H interactions.

Before the verification experiment can be done, it is necessary to rearray yeast clones corresponding to candidate Y2H interacting pairs into new plates. During the rearray step, plate positions G12 and H12 should be left empty for subsequent controls.

1. From archival glycerol stocks of the individual AD and DB transformed yeast clones, rearray the (candidate) interaction partner clones into matching positions of plates containing 160 μl Sc-Trp (AD-Y) and Sc-Leu (DB-X) liquid media.
2. Incubate at 30 °C on a 96-well plate shaker for 72 h.
3. Prepare an archival stock by combining 80 μl of the yeast culture with 80 μl of 40% (v/v) autoclaved glycerol in a round-bottom 96-well plate.

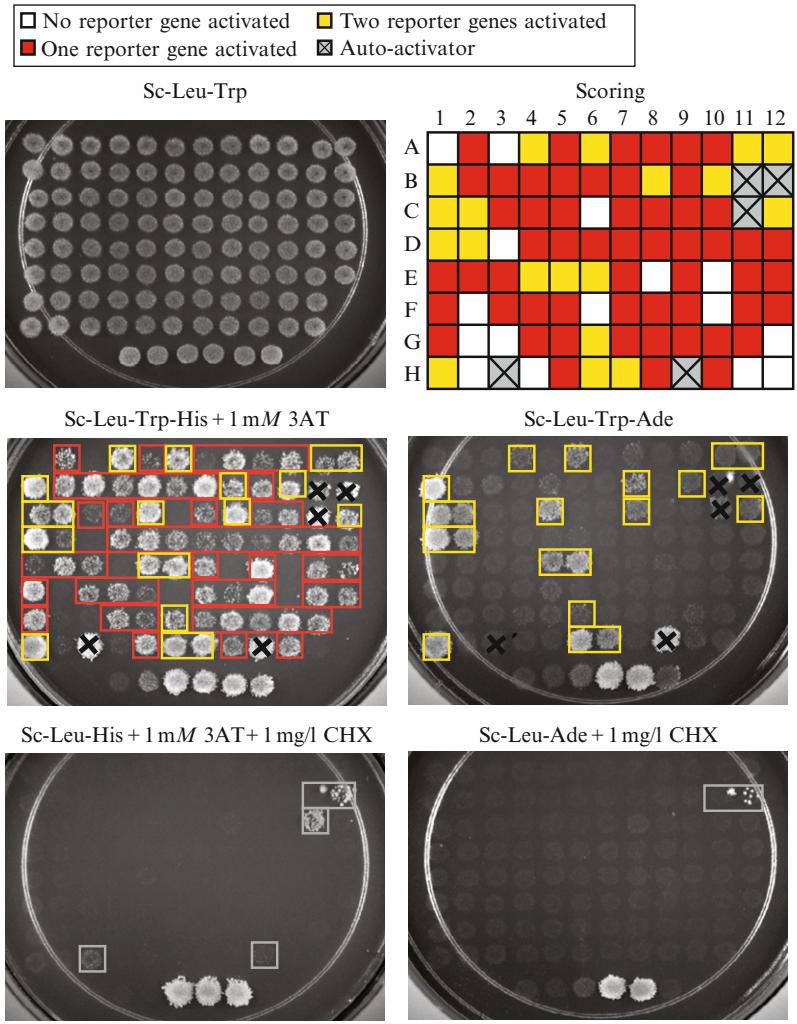

Figure 12.5 Phenotyping plates and scoring. First, autoactivators are identified and crossed out. The stringency of autoactivator detection is high such that even slight growth on the CHX control plates leads to elimination of the respective candidate. Subsequently, growth is evaluated on the selective -His and -Ade plates using the six controls (Fig. 12.1) as reference.

[Day 0: Inoculation]

1. Thaw glycerol stocks of rearrayed Y2H candidate pairs completely.
2. With 5 μl of glycerol stock, inoculate 160 μl of fresh Sc-Leu (DB-X) and Sc-Trp (AD-Y) liquid media dispensed in round-bottom 96-well culture plates.
3. Incubate at 30 °C for 72 h.

[Day 3: Mating]

1. Dispense 5 μl/well of AD-Y liquid culture onto a YEPD mating plate.
2. From the matching DB-X plate, dispense 5 μl/well of DB-X on top of the AD-Y spots.
3. Add Y2H controls.
4. Incubate at 30 °C for 14–18 h.

[Day 4: Selection of diploids]

1. Replica-plate mated yeast cells onto Sc-Leu-Trp diploid selection plates.
2. Incubate at 30 °C for 14–18 h.

[Day 5: Phenotyping of diploids]

1. Replica-plate diploid yeast cells onto the four phenotyping plates and autoactivator identification plates.
2. Immediately after, replica-clean all plates thoroughly by placing each plate on a piece of velvet stretched over a replica-plating block and pushing evenly on the plate to remove excess yeast. Replace the cloth and move to process the next plate until all plates have been cleaned.
3. Incubate at 30 °C for 3 days.

[Day 10: Scoring]

The scoring of each of the four plate sets is done independently in the same way as for secondary phenotyping. We consider as verified only those Y2H pairs that scored positive in at least three out of four plate sets and are never scored as an autoactivator.

3.6. Media and plates

3.6.1. Nonselective rich yeast medium (YEPD)

The Y8800 and Y8930 yeast strains are propagated on solid agar YEPD plates or in liquid YEPD medium.

YEPD media

1. Mix 20 g of yeast extract, 40 g of bacto-peptone and 1900 ml of water.
2. Autoclave for 45 min.
3. Store at room temperature.
4. Before use add 50 ml of 40% (w/v) autoclaved glucose and 15 ml of 65 m*M* adenine solution per liter of media.

YEPD agar plates

1. Mix 20 g of yeast extract and 40 g of bacto-peptone with 950 ml of water in a 2 l flask.

2. Add a stir bar.
3. Mix 40 g of agar and 950 ml of water in a second 2 l flask and shake well.
4. Autoclave the two flasks for 45 min.
5. Transfer the contents of each agar flask to one media flask and mix well.
6. Cool to 55 °C and keep in a water bath until ready to pour.
7. Before pouring, add 100 ml of autoclaved 40% (w/v) glucose and 30 ml of 65 m*M* adenine solution.
8. Pour 15 cm agar plates.
9. Dry for 5–7 days at room temperature and store at room temperature. If the plates need to be used earlier, they can be dried for 30 min in a sterile hood with the ventilation on.

3.6.2. Selective yeast media

Selective media are used for maintaining the AD-Y and DB-X plasmids and detection of reporter activity. Prototrophic markers are used for selection on plates lacking the appropriate amino acid or nucleotide. In our system the DB-expressing plasmid contains the selectable marker *LEU2* which enables growth of the Y8800/Y8930 yeast strains on plates lacking leucine (-Leu), while the AD-expressing plasmid contains the *TRP1* marker which enables growth on plates lacking tryptophan (-Trp). The other two prototrophic markers (*HIS3* and *ADE2*) are used as reporter genes in our experiments. Expression of these markers is selected on plates lacking histidine (-His) (supplemented with 1 m*M* 3-amino-1,2,4-triazole, 3AT) or lacking adenine (-Ade). Autoactivator detection plates are supplemented with 1 mg/l of CHX and contain tryptophan to allow growth of yeast cells without the AD-Y plasmid.

Synthetic complete (Sc) media The different selective media are based on the same Sc drop-out media recipe, but then supplemented with different amino acids to prepare the media appropriate for the various applications.

- *Sc media*

1. Mix 5.2 g of amino acid powder lacking leucine, tryptophan, histidine, and adenine, 6.8 g of yeast nitrogen base (without ammonium sulfate and amino acids), and 20 g of ammonium sulfate.
2. Dissolve in 1900 ml water and add a stir bar.
3. Adjust the pH to 5.9 by adding a few drops of 10 *M* NaOH.
4. Autoclave the flasks for 45 min.
5. Add 8 ml of each stock solution as needed. Store at room temperature.

- *Sc agar plates*

For a 4 l preparation of 15 cm agar Petri plates containing Sc medium lacking particular amino acids or nucleotides:

1. Place a magnetic stir bar into two 2 l flasks and label as the "media flasks."
2. Mix 5.2 g of amino acid powder lacking leucine, tryptophan, histidine, and adenine, 6.8 g of yeast nitrogen base (without ammonium sulfate and amino acids), and 20 g of ammonium sulfate.
3. Dissolve in 1900 ml water and add a stir bar.
4. Adjust the pH to 5.9 by adding a few drops of 10 *M* NaOH.
5. Add 40 g of agar and 900 ml of water to two 2 l flasks.
6. Autoclave the four flasks for 45 min.
7. Transfer the contents of each agar flask to one media flask and mix well.
8. Cool to 55 °C and keep in a water bath until ready to pour.
9. Add 100 ml of autoclaved 40% glucose (w/v).
10. Add the required concentrated stock solutions, and 3AT or CHX as needed.
11. Pour approximately 100 ml in 15-cm sterile Petri plates.
12. Dry for 5–7 days at room temperature then store indefinitely at 4 °C. If the plates need to be used earlier, they can be dried for 30 min in a sterile hood with ventilation on.

Amino acid powder mix and stock solutions All amino acids that are never used as prototrophic markers are combined in a amino acid mix that is added to all Sc plates.

To prepare the amino acid powders:

1. Mix 6 g of each of the following amino acids: alanine, arginine, aspartic acid, asparagine, cysteine, glutamic acid, glutamine, glycine, isoleucine, lysine, methionine, phenylalanine, proline, serine, threonine, tyrosine, and valine.
2. For the amino acid powder containing adenine, add 6 g of adenine sulfate.

Tryptophan, histidine, leucine, uracil, and adenine are omitted so they can be added to batches of plates as needed. The concentrated stock solutions are used at 8 ml/l of media, except for adenine which is used at 15 ml/l of media. The different stock solutions are prepared at the following concentrations: 100 m*M* histidine–HCl (store light protected), 100 m*M* leucine, 65 m*M* adenine sulfate, and 40 m*M* tryptophan. These stock solutions are stored at room temperature, except for tryptophan, which should be stored in the dark at 4 °C.

4. Validation Using Orthogonal Binary Interaction Assays

Complementary assays are essential to assess the precision of a dataset against PRS and RRS (see 2.2.1.). The following complementary assays can be used to determine the precision of a dataset by testing a random sample, and as

part of an interaction assay tool-kit for confidence scoring of individual interactions (Braun *et al.*, 2009). We describe the yellow fluorescent protein (YFP) based protein complementation assay (Nyfeler *et al.*, 2005) and the sandwich ELISA-like well-NAPPA protein interaction assay (Braun *et al.*, 2009). All expression constructs for these methods can be assembled using Gateway recombinational cloning or other high-throughput cloning methods.

Protocol 14: Yellow fluorescent protein complementation assay (YFP-PCA)

In YFP–PCA, two nonfluorescent fragments of YFP (F1 and F2) are genetically attached to ORFs coding for the two proteins that are to be tested in this assay. If the two proteins interact functional YFP can be reconstituted and detected by fluorescence-activated cell sorting (FACS). In this protocol a cyan fluorescent protein (CFP) coding plasmid is cotransfected as a transfection control.

[Day 0: Seed cells, measure DNA]

1. In a 96-well tissue culture plate, seed CHO-K1 cells at 6×10^4 cells/well in 100 μl Ham's F12 media + 10% fetal calf serum. After 24 h, confluence should reach 70%.
2. Determine the concentration of the expression plasmids with PicoGreen assay (Invitrogen) or related assay.

[Day 1: Transfection]

1. Replace growth media on cells with 100 μl Opti-MEM media (Invitrogen) equilibrated to 37 °C.
2. Combine 30 ng of each PCA construct with 140 ng CFP plasmid for a total of 200 ng DNA in 25 μl Opti-MEM media per well to obtain the DNA mix.
3. Combine 0.5 μl Lipofectamine 2000 reagent (Invitrogen) with 25 μl Opti-MEM media per well to obtain the transfection reagent mix.
4. Incubate 5–25 min at room temperature.
5. Combine the DNA and transfection reagent mixes to yield 50 μl transfection mix.
6. Incubate for at least 20 min (not longer than 6 h).
7. Add transfection mix to the cells.
8. Incubate for 18 h.

[Day 3: FACS Analysis]

1. Wash cells three times gently with phosphate-buffered saline (PBS).
2. Add 20 μl trypsin.
3. Incubate $\sim$10 min at room temperature until cells are detached.
4. Resuspend in 100 μl PBS.
5. Analyze cells by FACS.

Count a minimum of 10,000 events. Gate for CFP positive cells and analyze YFP fluorescence only in this subpopulation. Discard any result that is supported by less than 200 cells or if the CFP transfection rate is unacceptably low (< 5%). On every FACS instrument the voltages and gates need to be calibrated using YFP and CFP controls. The best criteria for scoring positive interactions should be identified using a large enough set of controls (at least one plate worth of each PRS and RRS). After such a calibration, score a pair positive if at least 30% of CFP positive cells are YFP positive and if the average YFP signal is above background and if the YFP/CFP ratio was at least twice as high as the ratio of the average YFP signal over the average CFP ratio on that plate. Calibrate gating of the instrument by using full-length YFP and CFP constructs. Scoring parameters must be recalibrated on PRS/RRS data for each implementation.

Protocol 15: Well-nucleic acid programmable protein array (wNAPPA)

In well-NAPPA, the two proteins are genetically fused to a glutathione-*S*-transferase tag and to an HA epitope tag respectively and expressed in a coupled transcription/translation reticulocyte lysate. The GST-tagged protein (GST-X) is captured using an anti-GST antibody that is immobilized at the bottom of a 96-well microtiter plate. If the two proteins are interacting, this interaction can be detected with an anti-HA antibody. Like all assays, this biochemical pull-down assay from *in vitro* coupled transcription–translation needs to be calibrated against PRS and RRS datasets to evaluate performance.

[Day 0: Blocking]

1. Add 200 μl/well blocking buffer (5% (w/v) fat-free dry milk powder dissolved in PBS prepared according to standard protocols) to a microtiter plated coated with rabbit anti-GST antibody (GST 96-well Detection Module, GE Healthcare).
2. Block at 4 °C for 14–24 h.

[Day 1: wNAPPA assay]

1. Determine the DNA concentration of expression plasmids using PicoGreen or a similar assay.
2. Add 0.5–1 μg of each of the two plasmids to complete reticulocyte lysate reaction mix (25 μl) (TnT Coupled Transcription/Translation System, Promega).
3. Incubate for 1.5 h at 30 °C on a shaker.
4. Dilute the reaction mix with 100 μl/well blocking solution.
5. Transfer the diluted reaction mix to the prepared anti-GST coated plate.

6. Incubate at 15 °C on a shaker for 2 h.
7. Discard reaction mix and wash three times with 200 μl blocking buffer for 5 min.
8. Add 150 μl anti-HA monoclonal antibody (Cell Signaling Technologies) 1:5000 in blocking buffer.
9. Wash three times with 150 μl with blocking buffer for 5 min each.
10. Add horseradish peroxidase (HRP) coupled goat anti-mouse antibody (Amersham) 1:1000–1:2000 in blocking buffer.
11. Wash three times with 150 μl PBS for 5 min.
12. Develop with 100 μl enhanced chemiluminescence (ECL) reagent like Pierce PicoWest ECL reagent. Alternatively a colorimetric HRP substrate will give similar results.
13. Chemiluminescence is measured with a Biorad molecular imager gel doc system, but measurement could also be done with a 96-well plate spectrophotometer reader.

5. Conclusion

Information on interactome networks constitutes a critical element of systems biology. We have spelled out a general approach to high-quality interactome mapping in which a reliable high-throughput assay is used as a primary screening platform. Subsequently, alternative validation assays are used to demonstrate data quality in a way unprejudiced by preconceived ideas and biases about what protein interactions are supposed to look like. To produce high-quality data, appropriate controls need to be implemented at every stage of a binary interactome mapping pipeline, including thorough controls for technical artifacts and subsequent experimental determination of the quality of interactome network maps. Experimental validation of primary screening data ensures data quality unbiased by current scientific perceptions and hence of greatest utility for exploring biology.

Use of this general framework of interactome mapping, the main features of which are stringent removal of technical artifacts and experimental control of data quality, will enable production of high-quality datasets.

ACKNOWLEDGMENTS

We thank past and current members of the Vidal Lab and the Center for Cancer Systems Biology (CCSB) for their help and constructive discussions over the course of developing our binary interaction mapping strategies, framework, and protocols. This work was supported by National Human Genome Research Institute grants R01-HG001715 awarded to M.V. and D.E.H and P50-HG004233 awarded to M.V., grant DBI-0703905 from the National Science Foundation to M. V. and D. E. H., and by Institute Sponsored Research

funds from the Dana-Farber Cancer Institute Strategic Initiative awarded to M. V. and CCSB. D. M. is supported by grant LSHG-CT-2006-037704 from the 6th Framework Program of the European Commission. M. V. is a "Chercheur Qualifié Honoraire" from the Fonds de la Recherche Scientifique (FRS-FNRS, French Community of Belgium).

REFERENCES

Bader, J. S., Chaudhuri, A., Rothberg, J. M., and Chant, J. (2004). Gaining confidence in high-throughput protein interaction networks. *Nat. Biotechnol.* **22,** 78–85.

Braun, P., Tasan, M., Dreze, M., Barrios-Rodiles, M., Lemmens, I., Yu, H., Sahalie, J. M., Murray, R. R., Roncari, L., De Smet, A.-S., Venkatesan, K., Rual, J.-F., *et al.* (2009). An experimentally derived confidence score for binary protein–protein interactions. *Nat. Methods* **6,** 91–97.

Collins, S. R., Kemmeren, P., Zhao, X. C., Greenblatt, J. F., Spencer, F., Holstege, F. C., Weissman, J. S., and Krogan, N. J. (2007a). Towards a comprehensive atlas of the physical interactome of *Saccharomyces cerevisiae*. *Mol. Cell Proteomics* **6,** 439–450.

Collins, S. R., Miller, K. M., Maas, N. L., Roguev, A., Fillingham, J., Chu, C. S., Schuldiner, M., Gebbia, M., Recht, J., Shales, M., Ding, H., Xu, H., *et al.* (2007b). Functional dissection of protein complexes involved in yeast chromosome biology using a genetic interaction map. *Nature* **446,** 806–810.

Cusick, M. E., Yu, H., Smolyar, A., Venkatesan, K., Carvunis, A. R., Simonis, N., Rual, J.-F., Borick, H., Braun, P., Dreze, M., Vandenhaute, J., Galli, M., *et al.* (2009). Literature-curated protein interaction datasets. *Nat. Methods* **6,** 39–46.

Deane, C. M., Salwinski, L., Xenarios, I., and Eisenberg, D. (2002). Protein interactions: Two methods for assessment of the reliability of high throughput observations. *Mol. Cell Proteomics* **1,** 349–356.

De Nicolo, A., Parisini, E., Zhong, Q., Dalla Palma, M., Stoeckert, K. A., Domchek, S. M., Nathanson, K. L., Caligo, M. A., Vidal, M., Cusick, M. E., and Garber, J. E. (2009). Multimodal assessment of protein functional deficiency supports pathogenicity of BRCA1 p.V1688del. *Cancer Res.* **69,** 7030–7037.

Dreze, M., Charloteaux, B., Milstein, S., Vidalain, P.-O., Yildirim, M. A., Zhong, Q., Svrzikapa, N., Romero, V., Laloux, G., Brasseur, R., Vandenhaute, J., Boxem, M., *et al.* (2009). 'Edgetic' perturbation of a *C. elegans* BCL-2 ortholog. *Nat. Methods* **7,** 843–849.

Durfee, T., Becherer, K., Chen, P. L., Yeh, S. H., Yang, Y., Kilburn, A. E., Lee, W. H., and Elledge, S. J. (1993). The retinoblastoma protein associates with the protein phosphatase type 1 catalytic subunit. *Genes Dev.* **7,** 555–569.

Endoh, H., Vincent, S., Jacob, Y., Real, E., Walhout, A. J., and Vidal, M. (2002). Integrated version of reverse two-hybrid system for the postproteomic era. *Methods Enzymol.* **350,** 525–545.

Fields, S., and Song, O. (1989). A novel genetic system to detect protein–protein interactions. *Nature* **340,** 245–246.

Ge, H., Liu, Z., Church, G. M., and Vidal, M. (2001). Correlation between transcriptome and interactome mapping data from *Saccharomyces cerevisiae*. *Nat. Genet.* **29,** 482–486.

Ge, H., Walhout, A. J., and Vidal, M. (2003). Integrating 'omic' information: A bridge between genomics and systems biology. *Trends Genet.* **19,** 551–560.

Gunsalus, K. C., Ge, H., Schetter, A. J., Goldberg, D. S., Han, J. D., Hao, T., Berriz, G. F., Bertin, N., Huang, J., Chuang, L. S., Li, N., Mani, R., *et al.* (2005). Predictive models of molecular machines involved in *Caenorhabditis elegans* early embryogenesis. *Nature* **436,** 861–865.

Gyuris, J., Golemis, E., Chertkov, H., and Brent, R. (1993). Cdi1, a human G1 and S phase protein phosphatase that associates with Cdk2. *Cell* **75,** 791–803.

Han, J. D., Bertin, N., Hao, T., Goldberg, D. S., Berriz, G. F., Zhang, L. V., Dupuy, D., Walhout, A. J., Cusick, M. E., Roth, F. P., and Vidal, M. (2004). Evidence for dynamically organized modularity in the yeast protein–protein interaction network. *Nature* **430,** 88–93.

Ito, T., Chiba, T., Ozawa, R., Yoshida, M., Hattori, M., and Sakaki, Y. (2001). A comprehensive two-hybrid analysis to explore the yeast protein interactome. *Proc. Natl. Acad. Sci. USA* **98,** 4569–4574.

James, P., Halladay, J., and Craig, E. A. (1996). Genomic libraries and a host strain designed for highly efficient two-hybrid selection in yeast. *Genetics* **144,** 1425–1436.

Jeong, H., Mason, S. P., Barabasi, A. L., and Oltvai, Z. N. (2001). Lethality and centrality in protein networks. *Nature* **411,** 41–42.

Lamesch, P., Milstein, S., Hao, T., Rosenberg, J., Li, N., Sequerra, R., Bosak, S., Doucette-Stamm, L., Vandenhaute, J., Hill, D., and Vidal, M. (2004). *C. elegans* ORFeome version 3.1: Increasing the coverage of ORFeome resources with improved gene predictions. *Genome Res.* **14,** 2064–2069.

Lamesch, P., Li, N., Milstein, S., Fan, C., Hao, T., Szabo, G., Hu, Z., Venkatesan, K., Bethel, G., Martin, P., Rogers, J., Lawlor, S., *et al.* (2007). hORFeome v3.1: A resource of human open reading frames representing over 10, 000 human genes. *Genomics* **89,** 307–315.

Miller, J. P., Lo, R. S., Ben-Hur, A., Desmarais, C., Stagljar, I., Noble, W. S., and Fields, S. (2005). Large-scale identification of yeast integral membrane protein interactions. *Proc. Natl. Acad. Sci. USA* **102,** 12123–12128.

Milo, R., Shen-Orr, S., Itzkovitz, S., Kashtan, N., Chklovskii, D., and Alon, U. (2002). Network motifs: Simple building blocks of complex networks. *Science* **298,** 824–827.

Nyfeler, B., Michnick, S. W., and Hauri, H. P. (2005). Capturing protein interactions in the secretory pathway of living cells. *Proc. Natl. Acad. Sci. USA* **102,** 6350–6355.

Oliver, S. (2000). Guilt-by-association goes global. *Nature* **403,** 601–603.

Pujana, M. A., Han, J. D., Starita, L. M., Stevens, K. N., Tewari, M., Ahn, J. S., Rennert, G., Moreno, V., Kirchhoff, T., Gold, B., Assmann, V., ElShamy, W. M., *et al.* (2007). Network modeling links breast cancer susceptibility and centrosome dysfunction. *Nat. Genet.* **39,** 1338–1349.

Reboul, J., Vaglio, P., Rual, J.-F., Lamesch, P., Martinez, M., Armstrong, C. M., Li, S., Jacotot, L., Bertin, N., Janky, R., Moore, T., Hudson, J. R. Jr., *et al.* (2003). *C. elegans* ORFeome version 1.1: Experimental verification of the genome annotation and resource for proteome-scale protein expression. *Nat. Genet.* **34,** 35–41.

Rual, J.-F., Hirozane-Kishikawa, T., Hao, T., Bertin, N., Li, S., Dricot, A., Li, N., Rosenberg, J., Lamesch, P., Vidalain, P. O., Clingingsmith, T. R., Hartley, J. L., *et al.* (2004). Human ORFeome version 1.1: A platform for reverse proteomics. *Genome Res.* **14,** 2128–2135.

Rual, J.-F., Venkatesan, K., Hao, T., Hirozane-Kishikawa, T., Dricot, A., Li, N., Berriz, G. F., Gibbons, F. D., Dreze, M., Ayivi-Guedehoussou, N., Klitgord, N., Simon, C., *et al.* (2005). Towards a proteome-scale map of the human protein–protein interaction network. *Nature* **437,** 1173–1178.

Simonis, N., Rual, J.-F., Carvunis, A.-R., Tasan, M., Lemmons, I., Hirozane-Kishikawa, T., Hao, T., Sahalie, J. M., Venkatesan, K., Gebreab, F., Cevik, S., Klitgord, N., *et al.* (2009). Empirically controlled mapping of the *Caenorhabditis elegans* protein–protein interactome network. *Nat. Methods* **6,** 47–54.

Tarassov, K., Messier, V., Landry, C. R., Radinovic, S., Molina, M. M., Shames, I., Malitskaya, Y., Vogel, J., Bussey, H., and Michnick, S. W. (2008). An *in vivo* map of the yeast protein interactome. *Science* **320,** 1465–1470.

Venkatesan, K., Rual, J.-F., Vazquez, A., Stelzl, U., Lemmens, I., Hirozane-Kishikawa, T., Hao, T., Zenkner, M., Xin, X., Goh, K. I., Yildirim, M. A., Simonis, N., *et al.* (2009). An empirical framework for binary interactome mapping. *Nat. Methods* **6,** 83–90.

Vidal, M. (2001). A biological atlas of functional maps. *Cell* **104,** 333–339.

Vidal, M. (2005). Interactome modeling. *FEBS Lett.* **579,** 1834–1838.

Vidal, M., and Endoh, H. (1999). Prospects for drug screening using the reverse two-hybrid system. *Trends Biotechnol.* **17,** 374–381.

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

Vidalain, P. O., Boxem, M., Ge, H., Li, S., and Vidal, M. (2004). Increasing specificity in high-throughput yeast two-hybrid screens. *Methods* **32,** 363–370.

Walhout, A. J., and Vidal, M. (1999). A genetic strategy to eliminate self-activator baits prior to high-throughput yeast two-hybrid screens. *Genome Res.* **9,** 1128–1134.

Walhout, A. J., and Vidal, M. (2001). High-throughput yeast two-hybrid assays for large-scale protein interaction mapping. *Methods* **24,** 297–306.

Walhout, A. J., Temple, G. F., Brasch, M. A., Hartley, J. L., Lorson, M. A., van den Heuvel, S., and Vidal, M. (2000). GATEWAY recombinational cloning: Application to the cloning of large numbers of open reading frames or ORFeomes. *Methods Enzymol.* **328,** 575–592.

Yu, H., Braun, P., Yildirim, M. A., Lemmens, I., Venkatesan, K., Sahalie, J., Hirozane-Kishikawa, T., Gebreab, F., Li, N., Simonis, N., Hao, T., Rual, J.-F., *et al.* (2008). High-quality binary protein interaction map of the yeast interactome network. *Science* **322,** 104–110.

Zhong, Q., Simonis, N., Li, Q.-R., Charloteaux, B., Heuze, F., Klitgord, N., Tam, S., Yu, H., Venkatesan, K., Mou, D., Swearingen, V., Yildirim, M. A., *et al.* (2009). Edgetic perturbation models of human genetic disorders. *Mol. Syst. Biol.* **5,** 321.

Zhu, H., Bilgin, M., Bangham, R., Hall, D., Casamayor, A., Bertone, P., Lan, N., Jansen, R., Bidlingmaier, S., Houfek, T., Mitchell, T., Miller, P., *et al.* (2001). Global analysis of protein activities using proteome chips. *Science* **293,** 2101–2105.