Next-Generation Sequencing for Binary Protein–Protein Interactions

Bernhard Suter[1]*, Xinmin Zhang[2], C. Gustavo Pesce[1], Andrew R. Mendelsohn[1,3], Savithramma P. Dinesh-Kumar[4] and Jian-Hua Mao[5]

[1] *Next Interactions, Inc., Richmond, CA, USA,* [2] *BioInfoRx, Inc., Madison, WI, USA,* [3] *Regenerative Sciences Institute, Sunnyvale, CA, USA,* [4] *Department of Plant Biology, University of California, Davis, Davis, CA, USA,* [5] *Life Sciences Division, Lawrence Berkeley National Laboratory, Berkeley, CA, USA*

The yeast two-hybrid (Y2H) system exploits host cell genetics in order to display binary protein–protein interactions (PPIs) via defined and selectable phenotypes. Numerous improvements have been made to this method, adapting the screening principle for diverse applications, including drug discovery and the scale-up for proteome wide interaction screens in human and other organisms. Here we discuss a systematic workflow and analysis scheme for screening data generated by Y2H and related assays that includes high-throughput selection procedures, readout of comprehensive results via next-generation sequencing (NGS), and the interpretation of interaction data via quantitative statistics. The novel assays and tools will serve the broader scientific community to harness the power of NGS technology to address PPI networks in health and disease. We discuss examples of how this next-generation platform can be applied to address specific questions in diverse fields of biology and medicine.

Keywords: protein–protein interactions, yeast two-hybrid, interactome mapping, next-generation sequencing, quantitative interaction profiles

INTRODUCTION

Networks of protein–protein interactions (PPIs) govern essentially all biological processes and mechanisms, such as receptor-ligand recognition, immune responses, intracellular and extracellular signaling, growth regulation, and development. Early on, PPI networks or "interactomes" were recognized as the next frontier in biomedicine after the completion of the human genome project (Mendelsohn and Brent, 1999). The role of interaction networks in complex diseases is now a central focus in network biology (Vidal et al., 2011; Sharma et al., 2015). Innovative concepts and technologies are therefore required to satisfy a broad and unmet need for highly reliable and efficient mapping of PPIs.

The charting of interactomes is in many ways more challenging than that of genomes. Proteins are encoded by multiple transcript isoforms and are localized in diverse cellular compartments with distinct milieus. Moreover, variations in amino acids and post-translational modifications affect and determine PPIs. Hence, from a practical perspective, working with proteins is more demanding than working with DNA. For these reasons, to this date, our technical capabilities for systematic approaches toward PPI networks remain limited, when compared with the routine deciphering of genomes, transcriptomes, and exomes at high efficiency and low cost by next-generation sequencing (NGS) technologies (Shendure, 2011; Shendure and Lieberman Aiden, 2012; Mardis, 2013).

Over the last 30 years, diverse technologies have been developed to detect PPIs that are based on different principles with individual strengths and weaknesses. Affinity purification followed

by mass-spectrometry (AP-MS) is the standard method to identify protein complexes (Bensimon et al., 2012; Dunham et al., 2012). On the other hand, a variety of assays, such as yeast two-hybrid (Y2H), as well as protein fragment complementation (PCA) in yeast and various mammalian assays, are currently applied for the *in vivo* screening of binary interactions to identify direct binding partners (Stynen et al., 2012). These assays rely on the reconstitution of PPIs *in vivo* and the direct or indirect activation of reporters for selection and scoring of interactions.

Since its inception (Fields and Song, 1989; Gyuris et al., 1993), Y2H has emerged as a widely applied approach for the exploration of novel PPIs and interactome-wide screens (Vidal and Fields, 2014). The assay relies on the splitting of a transcription factor into its DNA binding and activation domains. In most implementations, the bait protein is fused to the DNA binding domain, whereas the prey or a library of prey cDNAs is fused to the activation domain. A physical interaction between bait and prey reconstitutes the transcription factor and activates one or several reporter genes, allowing selection of yeast cells expressing interacting bait-prey pairs. After selection for growth, only a small minority of cells with interacting proteins is enriched over a large background of cells containing non-interacting proteins. Y2H provides therefore a genetic selection system, in which interaction partners can be identified by sequencing the DNA encoding the prey proteins that interact with a defined bait protein.

A variety of other existing *in vivo* assays for screening binary PPIs can be considered alternative implementation of Y2H principles, such as split ubiquitin system for membrane proteins (Obrdlik et al., 2004; Jones et al., 2014), the reverse Y2H screening system and the two-bait interaction trap to explore the effect of allelic variants on PPIs (Vidal et al., 1996; Xu et al., 1997). The yeast one-hybrid technique is a variant for the identification of proteins that bind to DNA motifs and transcription factor binding sites (Fuxman Bass et al., 2015). With yeast three-hybrid (Y3H) the goal is identification of proteins binding to small molecule drugs (protein-drug interactions; PDIs; Moser and Johnsson, 2013).

In this article, we give an overview on existing methods that present different solutions to use NGS as readout for Y2H data. We also present our own experimental and bioinformatics platform that we developed for this purpose and discuss how NGS can overcome the existing limitations of Y2H and diverse other binary interaction assays.

YEAST TWO-HYBRID TECHNOLOGIES AND MAPPING OF INTERACTOMES

High-throughput Y2H assays have been instrumental in proteome-wide screens for the mapping of PPIs that were so far undertaken in human and various model organisms (Uetz et al., 2000; Rual et al., 2005; Stelzl et al., 2005; Yu et al., 2008; Simonis et al., 2009; Rolland et al., 2014). A recent focus for high-throughput Y2H is on differential PPIs of normal and disease-associated alleles occurring in the human population (Dittmer et al., 2014; Sahni et al., 2015). In matrix-based Y2H

procedures, comprehensive collections of bait and prey strains are combined in high-throughput, using robotic infrastructure (Uetz et al., 2000, 2004; Stelzl et al., 2005). Yeast clones are arrayed on defined matrix positions, therefore PPIs are scored as visual readouts, eliminating the need to do DNA sequencing for identification. Moreover, the use of annotated full-length open reading frames (ORFs) also circumvents potential artifacts that are associated with cDNA libraries. On the other hand, the requirement for preassembled and defined libraries restricts this method to human and well-defined model organisms for which ORF collections have been made. Moreover, the automated setup that is required for this approach is expensive and not readily available for many researchers.

Despite the importance of Y2H as a discovery system, most Y2H results, also those generated in high-throughput experiments, are not based on truly quantitative measurements. This contrasts with gene expression and protein–DNA interactions which have been systematically explored with DNA microarrays and NGS. Notably, the use of DNA microarrays for parallel identification of Y2H screening results was recognized early on (Cho et al., 1998). More recently, a microarray-Y2H screening and scoring system was introduced and applied to identify interaction partners of huntingtin and ataxin-1, two important determinants for neurodegenerative diseases (Suter et al., 2013). Using the Qi-Sampler repeat sampling tool (Fontaine et al., 2011), microarray-Y2H results were benchmarked against sets of known positives (golden sets) and other gene sets for statistical enrichments. High-confidence microarray-Y2H interactions correlated with positives from the literature and PPIs that were confirmed with luminescence-based mammalian interactome mapping as an alternative assay. Moreover, the quantitative scoring of interaction data and comparison to background controls allowed the elimination of many non-specific binders or sticky prey proteins.

The first adaptation of NGS technology for Y2H came from the lab of Marc Vidal (Yu et al., 2011). In the Stitch-Seq method, the sequences of putatively interacting bait and prey proteins are concatenated so that they comprise a single amplicon for a massive and parallel NGS readout. The method was successfully used to generate high-throughput Y2H datasets (Rolland et al., 2014). The Y2H-Seq approach by the group of Ulrich Stelzl relies on the combination of NGS with matrix Y2H (Weimann et al., 2013). It demonstrated the advantages of the NGS readout for scalability by sequencing the results of hundreds of separate screens through barcode indexing in a single Illumina run. A higher interaction coverage in the screened interactome space was achieved by increasing the sensitivity for detection of PPIs. The Y2H-Seq screens resulted in a network of 523 interactions involving 22 methyltransferases or demethylases for previously undiscovered cellular roles in non-histone protein methylation. However, while Y2H-Seq and Stitch-Seq are powerful tools and pioneering implementations of NGS for Y2H, they are intended for interactome screenings with ORF libraries and aim primarily at increasing scale and sensitivity but do not fully exploit the quantitative potential of NGS.
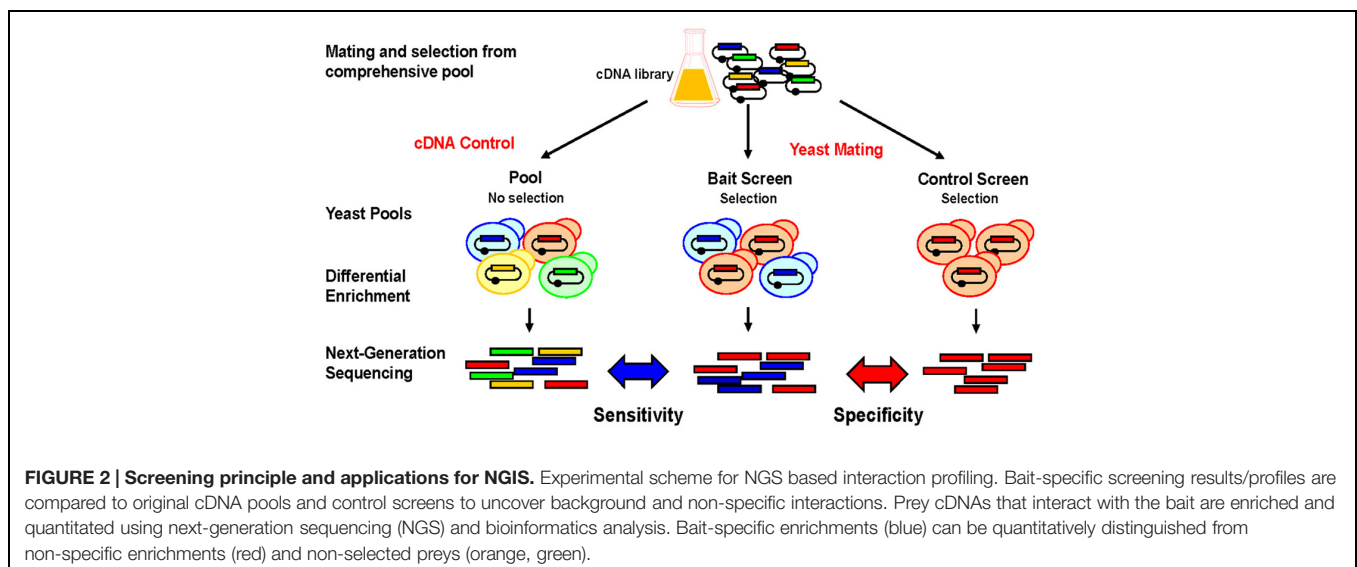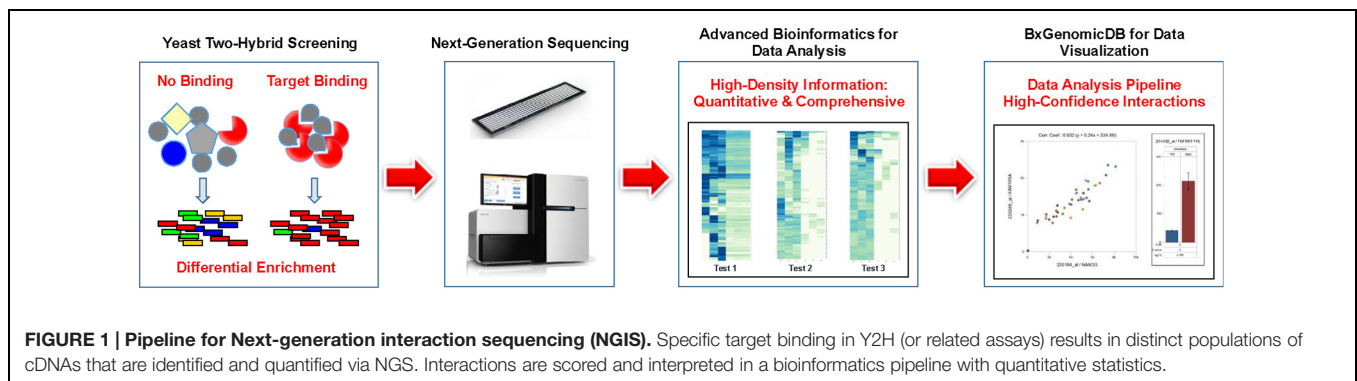
A NEXT GENERATION SOLUTION FOR Y2H SCREENS

We believe that the perceived shortcomings of Y2H such as inconsistent or non-reproducible results, lack of quantitation, laborious procedures, and above all, high rates of false positive results can be traced to the lack of an adequate readout system. With next-generation interaction screening (NGIS), we developed an innovative concept and methodology to harness the power of NGS technologies for the exploration of PPIs. The application of NGS removes the main restrictions on Y2H imposed by the cost of DNA sequencing. Replacing conventional Sanger sequencing with NGS leads to a massively increased throughput while reducing the cost of sequencing per screen to a small fraction of the conventional readouts (1,000–10,000-fold or more). Currently we are providing screening services for clients that include experimental work, data analysis, and the use of a cloud-based platform (**Figure 1**). NGIS procedures can be applied to every available Y2H and Y2H variant setup for binary interaction screens.

The technical principle of NGIS is shown in **Figure 2**. Tissue- or organ derived cDNA libraries that were cloned into Y2H prey vectors are combined with individual Y2H bait strains via cDNA transformation and mating procedures, and grown on selective medium. Selected prey cDNA clones are then amplified and products are fragmented and sequenced at their entire lengths with Illumina MiSeq or HiSeq. Most important, entire pools are sequenced after unbiased selection without the need to isolate individual clones. Another benefit of the NGIS protocol is that multiple repeat screens can now be undertaken to screen at maximum sensitivity, such that a weak enrichment corresponding to a single clone can be detected in a larger overall population and maximum coverage of the interaction space is achieved. With bioinformatics tools and algorithms adapted from RNA-Seq analytical methods, NGIS data can be processed to assign fold change and false discovery rate for every cDNA clone being sequenced in the assay. By comparing replicated bait results with controls (unrelated baits), the maximum information can be extracted out of the assays, scoring false positives and also taking into account the occurrence of false negatives and the reproducibility of the screening results.

With substantial cost reduction for screening and sequencing, it is worthwhile to generate large repeat datasets only for the purpose of screening the background of non-specific interactions. Indeed, non-specific Y2H activation by a subset of prey cDNAs (sticky preys) often makes up a majority of all hits in a

FIGURE 1 | Pipeline for Next-generation interaction sequencing (NGIS). Specific target binding in Y2H (or related assays) results in distinct populations of cDNAs that are identified and quantified via NGS. Interactions are scored and interpreted in a bioinformatics pipeline with quantitative statistics.

FIGURE 2 | Screening principle and applications for NGIS. Experimental scheme for NGS based interaction profiling. Bait-specific screening results/profiles are compared to original cDNA pools and control screens to uncover background and non-specific interactions. Prey cDNAs that interact with the bait are enriched and quantitated using next-generation sequencing (NGS) and bioinformatics analysis. Bait-specific enrichments (blue) can be quantitatively distinguished from non-specific enrichments (red) and non-selected preys (orange, green).

Y2H screen (Uetz, 2002). Hence, without prior knowledge, conventional Y2H requires specificity tests to confirm each identified PPI after the screening procedure is done, usually by isolation of cDNAs and retests with control strains (Vidalain et al., 2004). Using the NGIS screening scheme, bait specific DNA enrichments can be scored for specificity and non-specific interactions can be excluded *a priori*. This closes an existing gap to other technologies, such as AP-MS for which control datasets for background contaminants are routinely applied to distinguish bona fide interactors from non-specific contaminants (Lavallée-Adam et al., 2011; Mellacheruvu et al., 2013). Importantly, Y2H screening data can be viewed and interpreted as interaction profiles, comparable to transcription profiles in RNA sequencing (Trapnell et al., 2012; Law et al., 2014). Quantitative comparisons between different screen sets allow data mining and predictions for gene function that are impossible to do with the conventional Y2H readout by Sanger sequencing (Suter et al., 2013).

With NGIS interaction and interactome profiles, binary interaction screens can be adapted in several ways and toward different goals (**Table 1**). The primary goal in most Y2H screens is to define the function of proteins by identifying their molecular neighborhoods and to find specific targets that are relevant in diseases, e.g., proteins with functions in cancer or host receptors for pathogen effector proteins in microbial pathogenesis. NGIS interaction profiles and gene enrichment analysis help to understand the function of proteins of interest and the search for relevant interaction targets. An approach related to ours, Quantitative Interactor Sequencing (Qi-Seq), applied the split-ubiquitin system and Illumina NGS to screen for plant host targets for the HopZ2 effector protein that is secreted by the Gram-negative bacterial pathogen *Pseudomonas syringae*, and identified the *Arabidopsis thaliana* MLO2 protein as a target (Lewis et al., 2012).

Besides the discovery of novel PPIs, NGIS also provides a systematic approach to address changes in interaction profiles introduced by variants and polymorphisms in proteins that underlie phenotypes in complex and inheritable diseases. A number of studies have shown that Y2H assay is well-suited to detect changes in PPIs that are introduced by disease-specific alleles or random-generated amino acid mutations (Vidal et al., 1996; Xu et al., 1997; Dreze et al., 2009; Rolland et al., 2014).

A recent study profiled the interactions of several thousand missense mutations across a spectrum of Mendelian disorders (Sahni et al., 2015). The analysis indicated that two-thirds of disease-associated alleles perturb PPIs, while common variants from healthy individuals rarely affect interactions. Our NGIS platform provides a rapid way to compare PPI patterns from wild-type and mutant versions of the same protein. Quantitative Y2H data will not only show presence or absence of individual PPIs, but also shift in overall interaction patterns, which may cause gain or loss of protein function.

PERSPECTIVES AND FUTURE CHALLENGES

An immediate use of NGS based interaction screens with Y2H or Y2H variant techniques can be seen in the extraction of valuable and specific leads from quantitative and comprehensive interaction profiles. PPI profiles can be from wild-type and mutant proteins, as well as from isoforms of the same proteins, and also from full length proteins and their individual domains. Often, researchers are not interested in the complete set PPIs exhibited by a target of interest, but rather in a set of PPIs that are altered in disease. By providing an effective way to discover differential or regulated PPIs, NGIS could therefore constitute an important application to explore biological pathways and disease mechanisms.

Other areas in which NGIS could have an impact are protein engineering and target discovery for small molecule drugs (see **Table 1**). Considering that protein domains rather than full-length proteins are at the basic level of proteome organization, screening for protein fragments often reveals specific interaction sites and also PPIs that are masked in full length-proteins by steric hindrance. The value of fragment-based Y2H approaches was demonstrated previously (Boxem et al., 2008; Waaijers et al., 2013). NGS with complex cDNA libraries for high-resolution mapping of interaction sites could therefore be instrumental to achieve a full coverage of the protein interaction space. Reducing the lengths of interaction motifs further down to peptides, NGIS can also be applied for peptide aptamers for which Y2H has been instrumental (Bickle et al., 2006; Hamdi and Colas, 2012). We can also envision a role for NGIS procedures for the selection and optimization of scaffolds for aptamer displays. For example, libraries of novel aptamer scaffolds could be selected that can be targeted to diseased tissue and used both extra- and intracellularly. Scaffolds could then be optimized for functional interactions with proteins of interest.

Within the proper framework, NGIS could also, in principle, be applied for Y2H-based protein-drug interactions, such as Y3H to screen for novel protein targets that bind to known drugs (Moser and Johnsson, 2013), or to address the disruption of PPIs and protein complexes by small molecule binding (Flusin et al., 2012). The screening and selection in small volumes of liquid culture as opposed to large volumes of agar plates is a prerequisite for efficient screens in the presence of drugs. Quantitative analysis of NGIS data could be used to effectively distinguish drug-specific from non-specific interactions.

TABLE 1 | Solutions provided by NGIS for diverse problems and applications.

Area	Problem	Solution
Biological pathways	Mechanism of diverse diseases	Comparative interaction profiling
Microbial pathogenesis	Host virulence determinants	Comparative interaction profiling
Complex and inheritable diseases	Variants of unknown significance	Parallel interaction fingerprints
Protein engineering	Determinants of protein and peptide binding	Complete interaction landscapes
Drug discovery	Search for drug targets	Three-hybrid target discovery

By providing quantitative measurements, reproducibility by repeat assays, background controls for false positives, streamlined scoring and statistical analysis, NGIS overcomes existing bottlenecks of Y2H, thus providing a valuable technology and service platform. In addition, reconstruction of the components for Y2H fusion expression and reporter selection could increase accuracy, speed, automation, and cost-effectiveness for Y2H screens. A wide repertoire of sequence elements and well-characterized parts is now available for this purpose, although less attention had been paid to PPI and interaction affinities than to transcription parameters (Galdzicki et al., 2011). Regulated promoters that could compensate for differential expression of individual bait proteins could allow a better comparison between different interaction profiles. Another area for improvements is the use of new reporter assays to score interactions. For example, fluorescence measurements by cytometry for Y2H were already recognized as an alternative to the existing reporter systems (Chen et al., 2008). We expect that improved Y2H and Y2H-like assays will unlock the full potential of interaction screening and therefore provide a great benefit for biological and biomedical sciences.

AUTHOR CONTRIBUTIONS

BS developed NGS for Y2H screens is responsible for content and wrote the article. J-HM and SD-K are scientific collaborators and advisors for the NGIS technology, XZ does the bioinformatics and codeveloped the concept. GP and AM, helped develop the Y2H procedures. All authors read and commented the manuscript.

FUNDING

REFERENCES

Bensimon, A., Heck, A. J., and Aebersold, R. (2012). Mass spectrometry-based proteomics and network biology. *Annu. Rev. Biochem.* 81, 379–405. doi: 10.1146/annurev-biochem-072909-100424

Bickle, M. B., Dusserre, E., Moncorgé, O., Bottin, H., and Colas, P. (2006). Selection and characterization of large collections of peptide aptamers through optimized yeast two-hybrid procedures. *Nat. Protoc.* 1, 1066–1091. doi: 10.1038/nprot.2006.32

Boxem, M., Maliga, Z., Klitgord, N., Li, N., Lemmens, I., Mana, M., et al. (2008). A protein domain-based interactome network for *C. elegans* early embryogenesis. *Cell* 134, 534–545. doi: 10.1016/j.cell.2008.07.009

Chen, J., Zhou, J., Bae, W., Sanders, C. K., Nolan, J. P., and Cai, H. (2008). A yEGFP-based reporter system for high-throughput yeast two-hybrid assay by flow cytometry. *Cytometry A* 73, 312–320. doi: 10.1002/cyto.a.20525

Cho, R. J., Fromont-Racine, M., Wodicka, L., Feierbach, B., Stearns, T., Legrain, P., et al. (1998). Parallel analysis of genetic selections using whole genome oligonucleotide arrays. *Proc. Natl. Acad. Sci. U.S.A.* 95, 3752–3757. doi: 10.1073/pnas.95.7.3752

Dittmer, T. A., Sahni, N., Kubben, N., Hill, D. E., Vidal, M., Burgess, R. C., et al. (2014). Systematic identification of pathological lamin A interactors. *Mol. Biol. Cell* 25, 1493–1510. doi: 10.1091/mbc.E14-02-0733

Dreze, M., Charloteaux, B., Milstein, S., Vidalain, P. O., Yildirim, M. A., Zhong, Q., et al. (2009). 'Edgetic' perturbation of a *C. elegans* BCL2 ortholog. *Nat. Methods* 6, 843–849. doi: 10.1038/nmeth.1394

Dunham, W. H., Mullin, M., and Gingras, A. C. (2012). Affinity-purification coupled to mass spectrometry: basic principles and strategies. *Proteomics* 12, 1576–1590. doi: 10.1002/pmic.201100523

Fields, S., and Song, O. (1989). A novel genetic system to detect protein-protein interactions. *Nature* 340, 245–246. doi: 10.1038/340245a0

Flusin, O., Saccucci, L., Contesto-Richefeu, C., Hamdi, A., Bardou, C., Poyot, T., et al. (2012). A small molecule screen in yeast identifies inhibitors targeting protein-protein interactions within the vaccinia virus replication complex. *Antiviral Res.* 96, 187–195. doi: 10.1016/j.antiviral.2012.07.010

Fontaine, J. F., Suter, B., and Andrade-Navarro, M. A. (2011). QiSampler: evaluation of scoring schemes for high-throughput datasets using a repetitive sampling strategy on gold standards. *BMC Res. Notes* 4:57. doi: 10.1186/1756-0500-4-57

Fuxman Bass, J. I., Sahni, N., Shrestha, S., Garcia-Gonzalez, A., Mori, A., Bhat, N., et al. (2015). Human gene-centered transcription factor networks for enhancers and disease variants. *Cell* 161, 661–673. doi: 10.1016/j.cell.2015.03.003

Galdzicki, M., Rodriguez, C., Chandran, D., Sauro, H. M., and Gennari, J. H. (2011). Standard biological parts knowledgebase. *PLoS ONE* 6:e17005. doi: 10.1371/journal.pone.0017005

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. doi: 10.1016/0092-8674(93)90498-F

Hamdi, A., and Colas, P. (2012). Yeast two-hybrid methods and their applications in drug discovery. *Trends Pharmacol. Sci.* 33, 109–118. doi: 10.1016/j.tips.2011.10.008

Jones, A. M., Xuan, Y., Xu, M., Wang, R. S., Ho, C. H., Lalonde, S., et al. (2014). Border control–a membrane-linked interactome of *Arabidopsis*. *Science* 344, 711–716. doi: 10.1126/science.1251358

Lavallée-Adam, M., Cloutier, P., Coulombe, B., and Blanchette, M. (2011). Modeling contaminants in AP-MS/MS experiments. *J. Proteome Res.* 10, 886–895. doi: 10.1021/pr100795z

Law, C. W., Chen, Y., Shi, W., and Smyth, G. K. (2014). Voom: precision weights unlock linear model analysis tools for RNA-seq read counts. *Genome Biol.* 15, R29. doi: 10.1186/gb-2014-15-2-r29

Lewis, J. D., Wan, J., Ford, R., Gong, Y., Fung, P., Nahal, H., et al. (2012). Quantitative Interactor Screening with next-generation Sequencing (QIS-Seq) identifies *Arabidopsis thaliana* MLO2 as a target of the *Pseudomonas syringae* type III effector HopZ2. *BMC Genomics* 13:8. doi: 10.1186/1471-2164-13-8

Mardis, E. R. (2013). Next-generation sequencing platforms. *Annu. Rev. Anal. Chem.* (Palo Alto Calif.) 6, 287–303. doi: 10.1146/annurev-anchem-062012-092628

Mellacheruvu, D., Wright, Z., Couzens, A. L., Lambert, J. P., St-Denis, N. A., Li, T., et al. (2013). The CRAPome: a contaminant repository for affinity purification-mass spectrometry data. *Nat. Methods* 10, 730–736. doi: 10.1038/nmeth.2557

Mendelsohn, A. R., and Brent, R. (1999). Protein interaction methods–toward an endgame. *Science* 284, 1948–1950. doi: 10.1126/science.284.5422.1948

Moser, S., and Johnsson, K. (2013). Yeast three-hybrid screening for identifying anti-tuberculosis drug targets. *Chembiochem* 14, 2239–2242. doi: 10.1002/cbic.201300472

Obrdlik, P., El-Bakkoury, M., Hamacher, T., Cappellaro, C., Vilarino, C., Fleischer, C., et al. (2004). K+ channel interactions detected by a genetic system optimized for systematic studies of membrane protein interactions. *Proc. Natl. Acad. Sci. U.S.A.* 101, 12242–12247. doi: 10.1073/pnas.0404467101

Rolland, T., Taşan, M., Charloteaux, B., Pevzner, S. J., Zhong, Q., Sahni, N., et al. (2014). A proteome-scale map of the human interactome network. *Cell* 159, 1212–1226. doi: 10.1016/j.cell.2014.10.050

Rual, J. F., Venkatesan, K., Hao, T., Hirozane-Kishikawa, T., Dricot, A., Li, N., et al. (2005). Towards a proteome-scale map of the human protein-protein interaction network. *Nature* 437, 1173–1178. doi: 10.1038/nature04209

Sahni, N., Yi, S., Taipale, M., Fuxman Bass, J. I., Coulombe-Huntington, J., Yang, F., et al. (2015). Widespread macromolecular interaction perturbations in human genetic disorders. *Cell* 161, 647–660. doi: 10.1016/j.cell.2015.04.013

Sharma, A., Menche, J., Huang, C. C., Ort, T., Zhou, X., Kitsak, M., et al. (2015). A disease module in the interactome explains disease heterogeneity, drug response and captures novel pathways and genes in asthma. *Hum. Mol. Genet.* 24, 3005–3020. doi: 10.1093/hmg/ddv001

Shendure, J. (2011). Next-generation human genetics. *Genome Biol.* 12, 408. doi: 10.1186/gb-2011-12-9-408

Shendure, J., and Lieberman Aiden, E. (2012). The expanding scope of DNA sequencing. *Nat. Biotechnol.* 30, 1084–1094. doi: 10.1038/nbt.2421

Simonis, N., Rual, J. F., Carvunis, A. R., Tasan, M., Lemmens, I., Hirozane-Kishikawa, T., et al. (2009). Empirically controlled mapping of the *Caenorhabditis elegans* protein-protein interactome network. *Nat. Methods* 6, 47–54. doi: 10.1038/nmeth.1279

Stelzl, U., Worm, U., Lalowski, M., Haenig, C., Brembeck, F. H., Goehler, H., et al. (2005). A human protein-protein interaction network: a resource for annotating the proteome. *Cell* 122, 957–968. doi: 10.1016/j.cell.2005.08.029

Stynen, B., Tournu, H., Tavernier, J., and Van Dijck, P. (2012). Diversity in genetic in vivo methods for protein-protein interaction studies: from the yeast two-hybrid system to the mammalian split-luciferase system. *Microbiol. Mol. Biol. Rev.* 76, 331–382. doi: 10.1128/MMBR.05021-11

Suter, B., Fontaine, J. F., Yildirimman, R., Raskó, T., Schaefer, M. H., Rasche, A., et al. (2013). Development and application of a DNA microarray-based yeast two-hybrid system. *Nucleic Acids Res.* 41, 1496–1507. doi: 10.1093/nar/gks1329

Trapnell, C., Roberts, A., Goff, L., Pertea, G., Kim, D., Kelley, D. R., et al. (2012). Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. *Nat. Protoc.* 7, 562–578. doi: 10.1038/nprot.2012.016

Uetz, P. (2002). Two-hybrid arrays. *Curr. Opin. Chem. Biol.* 6, 57–62. doi: 10.1016/S1367-5931(01)00288-5

Uetz, P., Giot, L., Cagney, G., Mansfield, T. A., Judson, R. S., Knight, J. R., et al. (2000). A comprehensive analysis of protein-protein interactions in *Saccharomyces cerevisiae*. *Nature* 403, 623–627. doi: 10.1038/35001009

Uetz, P., Rajagopala, S. V., Dong, Y. A., and Haas, J. (2004). From ORFeomes to protein interaction maps in viruses. *Genome Res.* 14, 2029–2033. doi: 10.1101/gr.2583304

Vidal, M., Braun, P., Chen, E., Boeke, J. D., and Harlow, E. (1996). Genetic characterization of a mammalian protein-protein interaction domain by using a yeast reverse two-hybrid system. *Proc. Natl. Acad. Sci. U.S.A.* 93, 10321–10326. doi: 10.1073/pnas.93.19.10321

Vidal, M., Cusick, M. E., and Barabasi, A. L. (2011). Interactome networks and human disease. *Cell* 144, 986–998. doi: 10.1016/j.cell.2011.02.016

Vidal, M., and Fields, S. (2014). The yeast two-hybrid assay: still finding connections after 25 years. *Nat. Methods* 11, 1203–1206. doi: 10.1038/nmeth.3182

Vidalain, P. O., Boxem, M., Ge, H., Li, S., and Vidal, M. (2004). Increasing specificity in high-throughput yeast two-hybrid experiments. *Methods* 32, 363–370. doi: 10.1016/j.ymeth.2003.10.001

Waaijers, S., Koorman, T., Kerver, J., and Boxem, M. (2013). Identification of human protein interaction domains using an ORFeome-based yeast two-hybrid fragment library. *J. Proteome Res.* 12, 3181–3192. doi: 10.1021/pr400047p

Weimann, M., Grossmann, A., Woodsmith, J., Özkan, Z., Birth, P., Meierhofer, D., et al. (2013). A Y2H-seq approach defines the human protein methyltransferase interactome. *Nat. Methods* 10, 339–342. doi: 10.1038/nmeth.2397

Xu, C. W., Mendelsohn, A. R., and Brent, R. (1997). Cells that register logical relationships among proteins. *Proc. Natl. Acad. Sci. U.S.A.* 94, 12473–12478. doi: 10.1073/pnas.94.23.12473

Yu, H., Braun, P., Yildirim, M. A., Lemmens, I., Venkatesan, K., Sahalie, J., et al. (2008). High-quality binary protein interaction map of the yeast interactome network. *Science* 322, 104–110. doi: 10.1126/science.1158684

Yu, H., Tardivo, L., Tam, S., Weiner, E., Gebreab, F., Fan, C., et al. (2011). Next-generation sequencing to generate interactome datasets. *Nat. Methods* 8, 478–480. doi: 10.1038/nmeth.1597