HHS Public Access

Author manuscript

Methods Mol Biol. Author manuscript; available in PMC 2020 January 08.

Published in final edited form as:

Methods Mol Biol. 2018 ; 1794: 1–14. doi:10.1007/978-1-4939-7871-7_1.

PROTEIN INTERACTOMICS BY TWO-HYBRID METHODS

Soon Gang Choi[1,2,3,*], **Aaron Richardson**[1,2,3,*], **Luke Lambourne**[1,2,3], **David E Hill**[1,2,3], **Marc Vidal**[1,3]

[1]Center for Cancer Systems Biology (CCSB), Dana-Farber Cancer Institute, Boston, MA 02215, USA

[2]Department of Cancer Biology, Dana-Farber Cancer Institute, Boston, MA 02215, USA

[3]Department of Genetics, Harvard Medical School, Boston, MA 02115, USA

Abstract

Comprehensive identification of direct, physical interactions between biological macromolecules, such as protein-protein, protein-DNA, and protein-RNA interactions, is critical for our understanding of the function of gene products as well as the global organization and interworkings of various molecular machines within the cell. The accurate and comprehensive detection of direct interactions, however, remains a huge challenge due to the inherent structural complexity arising from various post-transcriptional and translational modifications coupled with huge heterogeneity in concentration, affinity, and subcellular location differences existing for any interacting molecules. This has created a need for developing multiple orthogonal and complementary assays for detecting various types of biological interactions. In this introduction, we discuss the methods developed for measuring different types of molecular interactions with an emphasis on direct protein-protein interactions, critical issues for generating high-quality interactome datasets, and the insights into biological networks and human diseases that current interaction mapping efforts provide. Further, we will discuss what future might lie ahead for the continued evolution of two-hybrid methods and the role of interactomics for expanding the advancement of biomedical science.

Keywords

interactome network; interactomics; systems biology; edgetics; protein-protein interaction; two-hybrid

1. INTERACTOMICS AND GENOMICS

The astonishing rate at which next generation sequencing technologies have been implemented over the past decade has ushered in an era of rapid genome sequencing that would have been unthinkable just a short time ago. The initial price tag of three billion dollars and years of global collaboration required to complete the initial draft sequence by the Human Genome Project [1] has since been entirely supplanted by sequencing costs of

*These authors contributed equally to this work.

approximately $1,000 per genome that now require a processing time of no more than a few weeks [2]. A significant outcome of this sequencing technology revolution and adoption has been the ever-expanding ability to use whole genome sequencing to pinpoint genes and mutant alleles associated with a variety of heritable traits and sporadic genetic disorders [3,4]. While such knowledge is critical to understand the genetic basis of many human diseases or dysfunctions in any organism, sequence information alone is limited in its ability to describe how genetic variation arising within the cell can lead to given phenotypic outcomes [5].

Biophysical interactions among and between proteins, DNA, RNA, lipids, and metabolites lie at the heart of all cellular functions. While the genome encodes the blueprint for all macromolecular entities, their biophysical interactions and macromolecular assemblies decode and implement the genetic information into all biochemical reactions, pathways, and structural components in cells. Studying macromolecular interactions, therefore, allows researchers to delineate the precise molecular mechanisms and biological functions that underlie a given genetic trait that are inaccessible by genomics investigations alone. The complete set of macromolecular interactions provides a necessary bridging of genotypes to phenotypes [6], in which gene products are viewed as components of an interaction network that collectively carry out a particular biological function, and genetic variants and disease associated mutations are perturbing agents that alter particular functional interactions within the network to give rise to specific phenotypic outcomes depending on the perturbed interactions [7]. This frequently elusive functional information is critical to address complex human diseases and develop specific and effective therapeutics.

Focusing on the collective behavior of genes, gene products, and their interactions contrasts with the traditional approach that typically seeks to identify the relationship between a particular gene and its function in isolation. In this collective view, biological functions emerge from complex interactions of individual components. A light bulb, by analogy, is made of glass, a filament, electric wires, and a socket, yet none of these components are individually capable of producing its core function of light emission. Only assembled parts with specific interaction partners and with correct interaction orientations can create a functional light bulb. The information found in macromolecular interactomes, therefore, is necessary for identifying synergistic relationships among various gene products and their collective contributions to underlying biological processes.

Much like technological advancements that have led to the expansion of the genomics field, so too has the development of a variety of different methodologies over the past few decades greatly enhanced the study of protein-protein interactions (PPIs). Owing to the inherent complexity of PPIs, such as vast heterogeneities in individual protein abundancies, affinities, subcellular localizations, and post-translational modifications, comprehensively characterizing protein interactions requires a variety of technological innovations and development of different PPI detection methods.

2. DETECTING PROTEIN-PROTEIN INTERACTIONS OR THEIR DISRUPTIONS WITH YEAST TWO-HYBRID

Forward Y2H: One of the most trusted, early-developed, and widely-used systems for detecting direct, biophysical interactions between two proteins is the yeast 2-hybrid (Y2H) assay. Y2H was introduced by Stanley Fields and Ok-Kyu Song in 1989 [8] as a genetic tool for the *in vivo* detection of direct binary PPIs. Y2H is based on the reconstitution of a transcription factor; summarized as DB-X:AD-Y, where DB encompasses a sequence-specific DNA-binding domain, AD encompasses a transcriptional activation domain, and X and Y are the proteins or protein fragments being tested. When selectable markers such as the yeast *HIS3* gene, involved in histidine biosynthesis, are expressed from a promoter containing DB-binding sites, the DB-X:AD-Y interaction confers a selective advantage to cells grown on medium lacking histidine, for example. In principle, any such selectable marker could be employed. An important benefit from having a positive growth selection is that the Y2H assay is scalable. Given the availability of genome-scale clone repositories such as human ORF clone collection [9] or complex cDNA libraries covering most expressed genes, the assay is suitable for conducting systematic and unbiased screens to identify specific interacting protein pairs from a pool of hundreds to thousands of candidates, and has led to the construction of proteome-scale PPI networks, or “protein interactomes” [6].

Reverse Y2H: Y2H is typically implemented as a positive growth selection for detection of direct, binary interactions. A reverse Y2H system has been developed that employs a positive selection for the disruption of a binary PPI. In this assay, a counter-selectable yeast reporter gene allows the detection of both the presence and absence of an interaction. Under the control of a heterologous promoter sequence that binds DB-X, the *URA3* gene is expressed only when X and Y (expressed as DB-X and AD-Y fusion proteins) interact [10]. The Ura3 protein catalyzes the conversion of 5-fluoroorotic acid (5-FOA) in the media into a toxic compound, 5-florouracil (5-FU), which causes cell death [10]. A positive growth selection ensues when loss of the interaction precludes expression of URA3, resulting in resistance to 5-FOA. This is particularly useful when the goal is to isolate mutant alleles in which a mutation in either X or Y prevent an interaction or when small molecules capable of disrupting an X:Y interaction are present.

3. OTHER METHODS FOR DETECTING PROTEIN-PROTEIN INTERACTIONS

A variety of alternative assays for studying biophysical interactions exist that have their own advantages, such as the ability to measure binding kinetics, localization-specific interactions, interaction visualization, compatibility with live cells, and discovery of PPI inhibitors. Assays designed to detect binary interactions often follow a similar split-protein strategy that characterizes Y2H, whereby a reporter protein function is reconstituted when two of its fragments that are covalently-linked to a pair of proteins are brought in contact through a physical interaction. While the modular nature of the DB and AD fragments in Y2H retain their respective functions independently, many of the split-protein assays require the two

fragments to refold before the reporter's function can be reconstituted. The ubiquitin split-protein sensor system [11] is an example of a yeast-based assay that follows this paradigm. N-terminal and C-terminal fragments of ubiquitin expressed as fusions to test proteins re-associate into a functional ubiquitin protein when the protein pair interacts. This prompts a signal cascade that ends in the activation of a transcriptional reporter gene in the nucleus. Whereas Y2H requires both test proteins to localize to the nucleus for reporter gene activation, a key feature of the split-ubiquitin system is that the bait and prey fusion proteins interact at the cell membrane. This makes it an especially advantageous assay for the study of membrane-bound proteins that are otherwise unable to localize to the nucleus [12]. Along with the split-ubiquitin system, other PPI assays, such as the optimized yeast cytosine deaminase (OyCD) assay [13], the split-dihydrofolate reductase (DHFR) assay [14], split-luciferase assay [15–17], and bimolecular fluorescence complementation assay [18] are classified as protein complementation assays due to the requirement of the reporter protein fragments to refold with each other in order to signal a proper interaction.

Similarly, MAPPIT, KISS, BATCH, and RRS systems measure binary PPIs [19–23] by reconstituting or mimicking the function of a multi-subunit protein complex in a signal transduction pathway. When test bait and prey proteins interact, a signal transduction cascade is activated that leads to reporter gene expression in the nucleus. While these systems have been very useful for studying and validating binary interactions in various host cell settings such as bacteria, yeast, and mammalian cells, a proteome-scale interactome determination has not been implemented.

While Y2H and other split-protein assays detect direct binary interactions, biochemical methods utilizing co-fractionation or affinity purification return information on co-complex associations. Affinity purification frequently involves using antibodies immobilized on a solid substrate to capture a bait protein, either directly or via a particular epitope tag along with all other proteins that inhabit the same protein complex. Mass spectrometry is then employed to determine the identity of the proteins that co-purify with the bait. In contrast to Y2H, these associations of prey proteins with the bait are a mix of both direct and indirect interactions and follow up studies are required to determine the precise contacts between any given pair of interactors. Affinity purification methods can be used with proteins that are endogenously expressed in their native cellular environment rather than having to be expressed in a heterologous host organism, such as yeast. This permits the possibility of capturing proteins containing post-translational modifications that may be critical for mediating an interaction between a particular set of proteins.

The different benefits and drawbacks exhibited by these assays speak to the importance for selecting the most appropriate assay for the task at hand. By the same token, however, the complementarity in the data generated by orthogonal assays underlies a potential benefit for implementing a "toolbox" of different assays, each with its own ability to detect a subset of different interactions, to uncover all possible interactions of a given protein.

The great success of two-hybrid methods has been due largely to their relatively simple, inexpensive, and rapid determination of binary macromolecular interactions in an *in vivo* cellular environment. Further, yeast and bacterial cell-based two-hybrid methods (either as

transcription factor reconstitution or protein complementation assays) are easily scalable to high-throughput, proteome-scale experimental settings often aided by powerful genetic selections with relatively little hands-on time and technical requirements, while such advantages are not readily attainable by pure biochemical interaction determinations.

4. BINARY ASSAYS TO DETECT OTHER MACROMOLECULAR INTERACTIONS

The two-hybrid paradigm has also been adapted for detecting interactions between proteins and other macromolecules. A yeast one-hybrid assay (Y1H) detects DNA-binding proteins that are expressed as AD-fusions in the presence of a DNA target sequence that is cloned immediately upstream of a transcriptional reporter gene. The AD-fusion protein will activate the reporter gene only if it binds to the target sequence in the reporter gene promoter and then recruits the RNAPII transcriptional machinery [10]. While other DNA-binding assays such as chromatin immunoprecipitation (ChIP) are primarily used to identify the specific DNA target sequences that a particular protein binds to, Y1H identifies the DNA-binding proteins among many hundred candidates that interact with a particular DNA target sequence [24]. Some limitations of Y1H method were noted, where Y1H may detect protein-DNA interactions (PDI) that do not occur in native host cell environments or miss PDIs that require specific post-translational modifications [24].

The yeast three-hybrid system (Y3H) was developed to study protein interactions with RNAs, small molecules, and natural ligands. In a Y3H system for identifying RNA-binding proteins, a target RNA is expressed in yeast as a hybrid sequence by adding cognate binding sites for a DB-fusion protein that tethers the target RNA to the DB-fusion protein. When an AD-fusion protein interacts with the target RNA, thereby forming a tripartite complex, RNAPII transcriptional machinery is recruited and reporter gene expression ensues [25,26]. Likewise, in a Y3H setting for identifying small molecule binding proteins, a hybrid synthetic molecule is created by adding a chemical spacer to a particular small molecule, and this chemical spacer mediates the attachment of the tested molecule to DB-fusion protein. If an AD-fused protein binds to the target small molecule, the resulting tripartite complex mediates transcriptional activity of a reporter gene [27–29]. Although several limitations are observed for these two methods such as difficulties in detecting weak ligand-protein interactions and targeting multimeric protein complexes, and the loss of activities by the chemical modification of ligands [30], they allow a high-throughput screen of candidate proteins that interact with target RNA or small molecules.

5. EMPIRICAL FRAMEWORK FOR INTERACTOME DATA GENERATION AND QUALITY ASSESSMENT

Various two hybrid methods and other interaction assays allow the systematic determination of macromolecular interactions such as protein-protein, protein-DNA, and protein-RNA interactions. Yet, several important considerations remain [31]. The first issue is how to ensure that an interaction dataset is of high-quality, and the second is how to assess the completeness of an interactome. To achieve the highest quality maps and assess their

completeness, it is important to implement a rigorous experimental and analytical framework. Four parameters that have been suggested for gauging the quality and completeness of an interaction dataset are assay sensitivity, assay precision, sampling sensitivity, and search space [31].

All interaction assays have a limited assay sensitivity such that each assay can detect only a fraction of all true-positive interactions. Assay sensitivity is affected by various experimental parameters, such as the presence and orientation of various tags or linkers fused to the proteins being tested, the choice of reporters, expression levels of the two tested proteins, and various cellular environments or *in vitro* assay conditions, which collectively affect the detectability of true-positive interactions. For binary PPI assays of all types performed under different cellular environments, all individual assays were shown to detect only 10~30% of a positive reference set of well-established binary interactions. Therefore, each assay exhibits a false negative detection rate of 70–90%, and different assays display limited overlap among detected PPIs [32].

Assay precision refers to the fraction of observed pairs in an interactome dataset that are true positives. The precision of an interactome dataset can be estimated with a positive reference set (PRS), a random reference set (RRS), and a random subset of pairs of the interactome dataset in question. Using another interaction assay orthogonal to the mapping method, the random subset, PRS, and RRS pairs are tested for interactions. The quality of the dataset can be inferred by comparing the recovery rate of the sample from the interactome to the PRS, at a threshold with a very low recovery rate for the RRS. A comparable recovery between the interactome pairs and the PRS indicates a high quality, low false positive rate. If there is a statistically significant non-zero correlation between testing positive in the interactome mapping assay and in the validation assay, then this should be accounted for. This is achieved by using only the subset of the PRS that is found with the same experimental method as the interactome dataset. A full quantitative estimation of the precision value and its uncertainty can be estimated by modeling the results as two binomial distributions.

$$P_{PRS} \sim \mathrm{Bin}(n_{PRS}, \alpha + \beta),\ P_{DATA} \sim \mathrm{Bin}\ (n_{DATA}, \rho\alpha + \beta)$$

Where P_{PRS} and P_{DATA} are the number of pairs scored positive in the orthogonal assay out of the total number of tested pairs for PRS and the dataset (n_{PRS} and n_{DATA} respectively). a is the probability of true positive interactions to score positive with the orthogonal assay, estimated by testing PRS. β is the false positive rate of the orthogonal assay, estimated by RRS. ρ is the parameter of interest, which is the precision of the dataset. The estimate of the precision parameter and its uncertainty are calculated using a Monte Carlo method. Two underlying assumptions for this method that the PRS is of very high precision and that it is representative of the interactions found in the interactome dataset. High-quality interactome datasets have been reported to show ~80% estimated precision using this method [31].

‘Sampling sensitivity’ is the percentage of all identifiable true-positive interactions detected by a single run of a given assay performed under a specific set of experimental conditions. Since a single screen never detects 100% of all detectable true positive interactions,

experiments are usually repeated multiple times to approach a detection saturation. In a high-throughput experiment setting, several parameters can contribute to the sampling sensitivity. In Y2H, for example, the mating efficiency of yeast cells and the pooling density of AD proteins limit the ability to systematically test all possible combinations of protein pairs in any one screen (e.g. particular yeast cells fail to mate with yeast strains of their interaction partner in the pool of many candidate yeast strains). Typically, as an example, there is a fraction of positive interactions repeatedly detected in multiple experiments while some interactions are detected in only one experiment among many repeats. If everything is properly set up, the number of uniquely identified pairs should approach a plateau after a number of repeated screens with decreasing numbers of newly identified positive interactors in each subsequent screen.

'Search space' indicates the fraction of all possible pairwise protein interactions that are tested. Depending on the list of bait and prey constructs or the expressed host cell proteome for mass spectrometry-based assays, the resulting dataset is derived from only a fraction of the entire proteome, which consequently limits the maximal detection of true-positive interactions.

When all these parameters are considered and optimized, a well-defined, high-quality interactome map can be generated.

6. SYSTEMS BIOLOGY ANALYSIS OF MACROMOLECULAR INTERACTOMES

With high-quality interactome maps, interesting biological questions can be addressed, often utilizing graph theory analysis of the network. From a high-quality, systematically derived and unbiased human interactome map, the vast majority of proteins associated with human diseases are distributed uniformly throughout the entire protein interactome space, while the interactions reported in the literature tend to be highly focused on only a very small fraction of the space [6]. This observation indicated a tendency that individual studies have been heavily influenced by the popularity of genes and their pathways, rather than the actual occurrence of diseases, which leaves many disease-associated genes and their interaction partners significantly understudied. As an interactome map is meant to bridge the gap between various genetic events and underlying biological and phenotypic responses, it has been used to identify how diseases are manifested in the network of interacting proteins. When all known phenotypes and disease-associated genes were linked as a network, Goh et al observed that multiple diseases share common genetic origins across a wide range of diseases [33]. Conversely, proteins encoded by genes associated with similar disorders exhibited a tendency to physically interact with each other more than random proteins and form clusters in the same neighborhood of the interactome network [33,34]. These clusters are called disease modules [35], where diseases can be viewed as localized perturbations within a certain interactome neighborhood [36]. A high similarity of phenotype and disease comorbidity was observed when disease modules overlap in the interactome network [37,38], and applied to identify new drug targets [38]. Similarly, interactome maps have helped to identify genes that have not been previously associated with a given disease

[39,40], for example, candidate cancer gene products in GWAS loci exhibit significant interconnectivity to known cancer gene products encoded by Sanger Cancer Census genes [6]. Likewise, a high-quality interactome map is often combined with other comparable quality interaction datasets to generate more comprehensive interactome maps [41] or integrated with other omics datasets to address various biological questions [42,40].

The tsunami of human population genetic variation data from modern genetic studies [43] highlights the need to be able to discern phenotypes conferred by disease-associated mutations versus natural variants. Nearly half of all reported inherited disease mutations in nuclear genes are missense mutations in a protein coding region leading to single amino acid substitutions [44]. From 4,222 structurally resolved PPIs, it was observed that such mutations are significantly enriched in amino acid sequences located in or near the interface with interacting partner proteins represented by 56% of all cases [45], which often results in the perturbation of interaction edges (‘edgetic’) rather than a ‘nodal’ perturbation [46,45]. Interestingly, when these mutations occur at different positions of a single protein, different phenotypes could be conferred as a reflection of differently perturbed interactions in the interactome network [46,45]. The question of whether disease-associated mutations specifically perturb the protein interactome has been systematically explored in the context of 2,890 disease-associated human mutations with non-synonymous missense changes [47]. After filtering WT genes for displaying multiple Y2H positive PPIs found in a screen against 7200 human proteins, the interactome of 197 disease associated mutant proteins were determined and subsequently compared to the interactome of WT proteins and that of natural variants. In the study, two-thirds of disease-associated mutants exhibited a perturbation in their wild-type PPIs, while the interactions among natural variants were largely unaffected. A similar observation was made for the DNA-binding capability of transcription factors that harbor disease mutations in that they showed a tendency of perturbed DNA-binding profile compared to those of natural variants and wild-type proteins. These findings presented the concept that interactomics can be applied to discern “wheat from chaff” such that functionally defective mutations can be prioritized over benign natural variants by testing their perturbation signature within a reference interactome network. Utilizing this concept, interactomics have been applied to interpret genetic variants associated with a particular disease such as cancer [48,49], neurodegeneration [50], and others, all of which help to prioritize candidate genes based on the global view of macromolecular interactome.

7. FUTURE PERSPECTIVE OF TWO HYBRID METHODS

In the three decades since the invention of the first two-hybrid method, we have witnessed the expansion and various adaptations of this general method [15,19,20,51,21,10], which has been instrumental in broadening our understanding of complex biological systems [33]. In spite of these advances, challenges remain for two hybrid methods. The cost, time, and resources for determining an interactome are several orders of magnitude higher than the cost of genome sequencing and gene expression profiling. However, genome-sequencing alone cannot unambiguously establish function and phenotype. Limited assay sensitivity and the need for combining multiple complementary two-hybrid methods to maximize the identification of true-positive interactions all make building a complete interactome very

challenging. Adherence to an empirical framework approach will ensure that high quality datasets that approach a complete reference interactome are produced. A comprehensive characterization of macromolecular interactomes in normal tissue or under disease conditions in human or other model organisms has not yet been attained, nor is yet understood how interactomes might change over time, under various physiological conditions or under disease states, which will be a very important goal to address in the years to come. How two-hybrid methods can and will evolve remains to be seen. Systematic, unbiased two-hybrid methods are still the most effective way to generate high quality binary interaction maps, and, when the networks are integrated with high quality data from other technologies, are best suited to provide the most proximal, mechanistic information to elucidate the molecular and genetic events that increase or decrease susceptibility to a wide range of human disease.

Continued development of versatile, ultra-high-throughput, robust, cost-effective, and highly sensitive two-hybrid methods is expected to come. Recent adaptations of DNA-barcoding technologies to protein interactomics are quite promising, ongoing developments. There, individual plasmids or proteins are labeled with unique DNA barcodes, and the protein interactions are determined by next-generation DNA sequencing of the genetic fusion of DNA barcodes in yeast [52] or their co-localization on a polyacrylamide thin film [53] by library-by-library format multiplex assays. Similarly, a massive multiplex Y2H method that employs the use of a Cre recombinase as a reporter gene was developed [54] whereby the gene fusion of interacting bait and prey proteins is induced by the expression of Cre. By converting the problem of detecting PPIs into a much simpler problem of DNA sequencing for detecting DNA-barcoded proteins [53], fused DNA-barcodes [52], or fused genes [54], these methods have the potential to drastically increase the throughput capacity of protein interactomics by several orders of magnitudes. With such advances, two hybrid methods can be applied to new frontiers such as building precision medicine tools for a patient diagnosis by profiling the interactome of patient-specific variants [55] or predicting their drug responses. The development of therapeutic drugs against specific PPIs is beginning to emerge [56], which are characteristically distinctive from traditional drug targets. Such efforts will be further facilitated by technological advances in two-hybrid methods, since one of major technological bottlenecks involves testing large numbers of therapeutically relevant human PPIs against vast numbers of small molecules in a combinatorial fashion to discover highly specific PPI modulatory compounds.

Altogether, the field of interactomics and its continued evolution holds great promise for advancing biology and medicine and understanding and precisely controlling complex biological networks.

Acknowledgements

We thank Yang Wang, Hong Yue, Julien Olivet, and Michael Calderwood for their critical reading of the manuscript and valuable comments. This work was supported by a Claudia Adams Barr Program for Innovative Cancer Research Award to S.G.C., and NHGRI grants U41HG001715, awarded to D.E.H. and M.V., and P50HG004233, awarded to M.V.

Bibliography

1. Venter JC (2010) Multiple personal genomes await. Nature 464 (7289):676–677. 10.1038/464676a [PubMed: 20360717]
2. Henson J, Tischler G, Ning Z (2012) Next-generation sequencing and large genome assemblies. Pharmacogenomics 13 (8):901–915. 10.2217/pgs.12.72 [PubMed: 22676195]
3. Altshuler D, Daly MJ, Lander ES (2008) Genetic mapping in human disease. Science 322 (5903): 881–888. 10.1126/science.1156409 [PubMed: 18988837]
4. Amberger JS, Bocchini CA, Schiettecatte F, Scott AF, Hamosh A (2015) OMIM.org: Online Mendelian Inheritance in Man (OMIM(R)), an online catalog of human genes and genetic disorders. Nucleic Acids Res 43 (Database issue):D789–798. 10.1093/nar/gku1205 [PubMed: 25428349]
5. Vidal M, Cusick ME, Barabasi AL (2011) Interactome networks and human disease. Cell 144 (6): 986–998. 10.1016/j.cell.2011.02.016 [PubMed: 21414488]
6. Rolland T, Taşan M, Charloteaux B, Pevzner SJ, Zhong Q, Sahni N, Yi S, Lemmens I, Fontanillo C, Mosca R, Kamburov A, Ghiassian SD, Yang X, Ghamsari L, Balcha D, Begg BE, Braun P, Brehme M, Broly MP, Carvunis A-R, Convery-Zupan D, Corominas R, Coulombe-Huntington J, Dann E, Dreze M, Dricot A, Fan C, Franzosa E, Gebreab F, Gutierrez BJ, Hardy MF, Jin M, Kang S, Kiros R, Lin GN, Luck K, MacWilliams A, Menche J, Murray RR, Palagi A, Poulin MM, Rambout X, Rasla J, Reichert P, Romero V, Ruyssinck E, Sahalie JM, Scholz A, Shah AA, Sharma A, Shen Y, Spirohn K, Tam S, Tejeda AO, Trigg SA, Twizere J-C, Vega K, Walsh J, Cusick ME, Xia Y, Barabási A-L, Iakoucheva LM, Aloy P, De Las Rivas J, Tavernier J, Calderwood MA, Hill DE, Hao T, Roth FP, Vidal M (2014) A Proteome-Scale Map of the Human Interactome Network. Cell 159 (5):1212–1226. 10.1016/j.cell.2014.10.050 [PubMed: 25416956]
7. Zhong Q, Simonis N, Li QR, Charloteaux B, Heuze F, Klitgord N, Tam S, Yu H, Venkatesan K, Mou D, Swearingen V, Yildirim MA, Yan H, Dricot A, Szeto D, Lin C, Hao T, Fan C, Milstein S, Dupuy D, Brasseur R, Hill DE, Cusick ME, Vidal M (2009) Edgetic perturbation models of human inherited disorders. Mol Syst Biol 5:321 10.1038/msb.2009.80 [PubMed: 19888216]
8. Fields S, Song O (1989) A novel genetic system to detect protein-protein interactions. Nature 340 (6230):245–246. 10.1038/340245a0 [PubMed: 2547163]
9. Wiemann S PC, Hu Y, Hunter P, Harbers M, Amiet A, Bethel G, Busse M, Carninci P, Dunham I, Hao T, Harper JW, Hayashizaki Y, Heil O, Hennig S, Hotz-Wagenblatt A, Jang W, Jöcker A, Kawai J, Koenig C, Korn B, Lambert C, LeBeau A, Lu S, Maurer J, Moore T, Ohara O, Park J, Rolfs A, Salehi-Ashtiani K, Seiler C, Simmons B, van Brabant Smith A, Steel J, Wagner L, Weaver T, Wellenreuther R, Yang S, Vidal M, Gerhard DS, LaBaer J, Temple G, Hill DE. (2016) The ORFeome Collaboration: a genome-scale human ORF-clone resource. Nat Methods 13 (3):191–192. 10.1038/nmeth.3776 [PubMed: 26914201]
10. Vidal M, Brachmann RK, Fattaey A, Harlow E, Boeke JD (1996) Reverse two-hybrid and one-hybrid systems to detect dissociation of protein-protein and DNA-protein interactions. Proc Natl Acad Sci U S A 93 (19):10315–10320 [PubMed: 8816797]
11. Johnsson N, Varshavsky A (1994) Split ubiquitin as a sensor of protein interactions in vivo. Proc Natl Acad Sci U S A 91 (22):10340–10344 [PubMed: 7937952]
12. Thaminy S, Miller J, Stagljar I (2004) The split-ubiquitin membrane-based yeast two-hybrid system. Methods Mol Biol 261:297–312. 10.1385/1-59259-762-9:297 [PubMed: 15064465]
13. Ear PH, Kowarzyk J, Booth MJ, Abd-Rabbo D, Shulist K, Hall C, Vogel J, Michnick SW (2016) Combining the Optimized Yeast Cytosine Deaminase Protein Fragment Complementation Assay and an In Vitro Cdk1 Targeting Assay to Study the Regulation of the gamma-Tubulin Complex. Methods Mol Biol 1342:237–257. 10.1007/978-1-4939-2957-3_14 [PubMed: 26254928]
14. Remy I, Campbell-Valois FX, Michnick SW (2007) Detection of protein-protein interactions using a simple survival protein-fragment complementation assay based on the enzyme dihydrofolate reductase. Nat Protoc 2 (9):2120–2125. 10.1038/nprot.2007.266 [PubMed: 17853867]
15. Cassonnet P, Rolloy C, Neveu G, Vidalain PO, Chantier T, Pellet J, Jones L, Muller M, Demeret C, Gaud G, Vuillier F, Lotteau V, Tangy F, Favre M, Jacob Y (2011) Benchmarking a luciferase complementation assay for detecting protein complexes. Nat Methods 8 (12):990–992. 10.1038/nmeth.1773 [PubMed: 22127214]

16. Ozawa T, Kaihara A, Sato M, Tachihara K, Umezawa Y (2001) Split luciferase as an optical probe for detecting protein-protein interactions in mammalian cells based on protein splicing. Anal Chem 73 (11):2516–2521 [PubMed: 11403293]
17. Paulmurugan R, Gambhir SS (2003) Monitoring protein-protein interactions using split synthetic renilla luciferase protein-fragment-assisted complementation. Anal Chem 75 (7):1584–1589 [PubMed: 12705589]
18. Hu CD, Chinenov Y, Kerppola TK (2002) Visualization of interactions among bZIP and Rel family proteins in living cells using bimolecular fluorescence complementation. Mol Cell 9 (4):789–798 [PubMed: 11983170]
19. Eyckerman S, Verhee A, der Heyden JV, Lemmens I, Ostade XV, Vandekerckhove J, Tavernier J (2001) Design and application of a cytokine-receptor-based interaction trap. Nat Cell Biol 3 (12): 1114–1119. 10.1038/ncb1201-1114 [PubMed: 11781573]
20. Lievens S, Gerlo S, Lemmens I, De Clercq DJ, Risseeuw MD, Vanderroost N, De Smet AS, Ruyssinck E, Chevet E, Van Calenbergh S, Tavernier J (2014) Kinase Substrate Sensor (KISS), a mammalian in situ protein interaction sensor. Mol Cell Proteomics 13 (12):3332–3342. 10.1074/mcp.M114.041087 [PubMed: 25154561]
21. Aronheim A (2004) Ras signaling pathway for analysis of protein-protein interactions in yeast and mammalian cells. Methods Mol Biol 250:251–262. 10.1385/1-59259-671-1:251 [PubMed: 14755093]
22. Karimova G, Gauliard E, Davi M, Ouellette SP, Ladant D (2017) Protein-Protein Interaction: Bacterial Two-Hybrid. Methods Mol Biol 1615:159–176. 10.1007/978-1-4939-7033-9_13 [PubMed: 28667611]
23. Karimova G, Ullmann A, Ladant D (2001) Protein-protein interaction between Bacillus stearothermophilus tyrosyl-tRNA synthetase subdomains revealed by a bacterial two-hybrid system. J Mol Microbiol Biotechnol 3 (1):73–82 [PubMed: 11200232]
24. Fuxman Bass JI, Reece-Hoyes JS, Walhout AJ (2016) Gene-Centered Yeast One-Hybrid Assays. Cold Spring Harb Protoc 2016 (12):pdb top077669 10.1101/pdb.top077669
25. SenGupta DJ, Zhang B, Kraemer B, Pochart P, Fields S, Wickens M (1996) A three-hybrid system to detect RNA-protein interactions in vivo. Proc Natl Acad Sci U S A 93 (16):8496–8501 [PubMed: 8710898]
26. Cho SK, Kang IH, Carr T, Hannapel DJ (2012) Using the Yeast Three-Hybrid System to Identify Proteins that Interact with a Phloem-Mobile mRNA. Front Plant Sci 3:189 10.3389/fpls.2012.00189 [PubMed: 22969782]
27. Wang Y, Letham DS, John PC, Zhang R (2016) Synthesis of a Cytokinin Linked by a Spacer to Dexamethasone and Biotin: Conjugates to Detect Cytokinin-Binding Proteins. Molecules 21 (5). 10.3390/molecules21050576
28. Licitra EJ, Liu JO (1996) A three-hybrid system for detecting small ligand-protein receptor interactions. Proc Natl Acad Sci U S A 93 (23):12817–12821 [PubMed: 8917502]
29. Becker F, Murthi K, Smith C, Come J, Costa-Roldan N, Kaufmann C, Hanke U, Degenhart C, Baumann S, Wallner W, Huber A, Dedier S, Dill S, Kinsman D, Hediger M, Bockovich N, Meier-Ewert S, Kluge AF, Kley N (2004) A three-hybrid approach to scanning the proteome for targets of small molecule kinase inhibitors. Chem Biol 11 (2):211–223. 10.1016/j.chembiol.2004.02.001 [PubMed: 15123283]
30. Cottier S, Monig T, Wang Z, Svoboda J, Boland W, Kaiser M, Kombrink E (2011) The yeast three-hybrid system as an experimental platform to identify proteins interacting with small signaling molecules in plant cells: potential and limitations. Front Plant Sci 2:101 10.3389/fpls.2011.00101 [PubMed: 22639623]
31. Venkatesan K, Rual J-F, Vazquez A, Stelzl U, Lemmens I, Hirozane-Kishikawa T, Hao T, Zenkner M, Xin X, Goh KI, Yildirim MA, Simonis N, Heinzmann K, Gebreab F, Sahalie JM, Cevik S, Simon C, de Smet AS, Dann E, Smolyar A, Vinayagam A, Yu H, Szeto D, Borick H, Dricot A, Klitgord N, Murray RR, Lin C, Lalowski M, Timm J, Rau K, Boone C, Braun P, Cusick ME, Roth FP, Hill DE, Tavernier J, Wanker EE, Barabási A-L, Vidal M (2009) An empirical framework for binary interactome mapping. Nat Methods 6 (1):83–90. 10.1038/nmeth.1280 [PubMed: 19060904]

32. Braun P, Taşan M, Dreze M, Barrios-Rodiles M, Lemmens I, Yu H, Sahalie JM, Murray RR, Roncari L, de Smet AS, Venkatesan K, Rual J-F, Vandenhaute J, Cusick ME, Pawson T, Hill DE, Tavernier J, Wrana JL, Roth FP, Vidal M (2009) An experimentally derived confidence score for binary protein-protein interactions. Nat Methods 6 (1):91–97. 10.1038/nmeth.1281 [PubMed: 19060903]
33. Goh KI, Cusick ME, Valle D, Childs B, Vidal M, Barabasi AL (2007) The human disease network. Proc Natl Acad Sci U S A 104 (21):8685–8690. 10.1073/pnas.0701361104 [PubMed: 17502601]
34. Feldman I, Rzhetsky A, Vitkup D (2008) Network properties of genes harboring inherited disease mutations. Proc Natl Acad Sci U S A 105 (11):4323–4328. 10.1073/pnas.0701722105 [PubMed: 18326631]
35. Barabasi AL, Gulbahce N, Loscalzo J (2011) Network medicine: a network-based approach to human disease. Nat Rev Genet 12 (1):56–68. 10.1038/nrg2918 [PubMed: 21164525]
36. Ghiassian SD, Menche J, Barabasi AL (2015) A DIseAse MOdule Detection (DIAMOnD) algorithm derived from a systematic analysis of connectivity patterns of disease proteins in the human interactome. PLoS Comput Biol 11 (4):e1004120 10.1371/journal.pcbi.1004120 [PubMed: 25853560]
37. Zhou X, Menche J, Barabasi AL, Sharma A (2014) Human symptoms-disease network. Nat Commun 5:4212 10.1038/ncomms5212 [PubMed: 24967666]
38. Menche J, Sharma A, Kitsak M, Ghiassian SD, Vidal M, Loscalzo J, Barabasi AL (2015) Disease networks. Uncovering disease-disease relationships through the incomplete interactome. Science 347 (6224):1257601 10.1126/science.1257601 [PubMed: 25700523]
39. Lim J, Hao T, Shaw C, Patel AJ, Szabo G, Rual JF, Fisk CJ, Li N, Smolyar A, Hill DE, Barabasi AL, Vidal M, Zoghbi HY (2006) A protein-protein interaction network for human inherited ataxias and disorders of Purkinje cell degeneration. Cell 125 (4):801–814. 10.1016/j.cell.2006.03.032 [PubMed: 16713569]
40. Pujana MA, Han JD, Starita LM, Stevens KN, Tewari M, Ahn JS, Rennert G, Moreno V, Kirchhoff T, Gold B, Assmann V, Elshamy WM, Rual JF, Levine D, Rozek LS, Gelman RS, Gunsalus KC, Greenberg RA, Sobhian B, Bertin N, Venkatesan K, Ayivi-Guedehoussou N, Sole X, Hernandez P, Lazaro C, Nathanson KL, Weber BL, Cusick ME, Hill DE, Offit K, Livingston DM, Gruber SB, Parvin JD, Vidal M (2007) Network modeling links breast cancer susceptibility and centrosome dysfunction. Nat Genet 39 (11):1338–1349. 10.1038/ng.2007.2 [PubMed: 17922014]
41. Luck K, Sheynkman GM, Zhang I, Vidal M (2017) Proteome-Scale Human Interactomics. Trends Biochem Sci 42 (5):342–354. 10.1016/j.tibs.2017.02.006 [PubMed: 28284537]
42. Rozenblatt-Rosen O, Deo RC, Padi M, Adelmant G, Calderwood MA, Rolland T, Grace M, Dricot A, Askenazi M, Tavares M, Pevzner SJ, Abderazzaq F, Byrdsong D, Carvunis AR, Chen AA, Cheng J, Correll M, Duarte M, Fan C, Feltkamp MC, Ficarro SB, Franchi R, Garg BK, Gulbahce N, Hao T, Holthaus AM, James R, Korkhin A, Litovchick L, Mar JC, Pak TR, Rabello S, Rubio R, Shen Y, Singh S, Spangle JM, Tasan M, Wanamaker S, Webber JT, Roecklein-Canfield J, Johannsen E, Barabasi AL, Beroukhim R, Kieff E, Cusick ME, Hill DE, Munger K, Marto JA, Quackenbush J, Roth FP, DeCaprio JA, Vidal M (2012) Interpreting cancer genomes using systematic host network perturbations by tumour virus proteins. Nature 487 (7408):491–495. 10.1038/nature11288 [PubMed: 22810586]
43. Genomes Project C, Auton A, Brooks LD, Durbin RM, Garrison EP, Kang HM, Korbel JO, Marchini JL, McCarthy S, McVean GA, Abecasis GR (2015) A global reference for human genetic variation. Nature 526 (7571):68–74. 10.1038/nature15393 [PubMed: 26432245]
44. Stenson PD, Ball EV, Mort M, Phillips AD, Shaw K, Cooper DN (2012) The Human Gene Mutation Database (HGMD) and its exploitation in the fields of personalized genomics and molecular evolution. Curr Protoc Bioinformatics Chapter 1:Unit1 13 10.1002/0471250953.bi0113s39 [PubMed: 22948725]
45. Wang X, Wei X, Thijssen B, Das J, Lipkin SM, Yu H (2012) Three-dimensional reconstruction of protein networks provides insight into human genetic disease. Nat Biotechnol 30 (2):159–164. 10.1038/nbt.2106 [PubMed: 22252508]
46. Zhong Q, Simonis N, Li QR, Charloteaux B, Heuze F, Klitgord N, Tam S, Yu HY, Venkatesan K, Mou D, Swearingen V, Yildirim MA, Yan H, Dricot A, Szeto D, Lin CW, Hao T, Fan CY, Milstein

S, Dupuy D, Brasseur R, Hill DE, Cusick ME, Vidal M (2009) Edgetic perturbation models of human inherited disorders. Mol Syst Biol 5:10 10.1038/msb.2009.80

47. Sahni N, Yi S, Taipale M, Bass JIF, Coulombe-Huntington J, Yang F, Peng J, Weile J, Karras GI, Wang Y, Kovacs IA, Kamburov A, Krykbaeva I, Lam MH, Tucker G, Khurana V, Sharma A, Liu YY, Yachie N, Zhong Q, Shen Y, Palagi A, San-Miguel A, Fan CY, Balcha D, Dricot A, Jordan DM, Walsh JM, Shah AA, Yang XP, Stoyanova AK, Leighton A, Calderwood MA, Jacob Y, Cusick ME, Salehi-Ashtiani K, Whitesell LJ, Sunyaev S, Berger B, Barabási A-L, Charloteaux B, Hill DE, Hao T, Roth FP, Xia Y, Walhout AJM, Lindquist S, Vidal M (2015) Widespread Macromolecular Interaction Perturbations in Human Genetic Disorders. Cell 161 (3):647–660. 10.1016/j.cell.2015.04.013 [PubMed: 25910212]
48. Krogan NJ, Lippman S, Agard DA, Ashworth A, Ideker T (2015) The cancer cell map initiative: defining the hallmark networks of cancer. Mol Cell 58 (4):690–698. 10.1016/j.molcel.2015.05.008 [PubMed: 26000852]
49. Engin HB, Kreisberg JF, Carter H (2016) Structure-Based Analysis Reveals Cancer Missense Mutations Target Protein Interaction Interfaces. PLoS One 11 (4):e0152929 10.1371/journal.pone.0152929 [PubMed: 27043210]
50. Hosp F, Vossfeldt H, Heinig M, Vasiljevic D, Arumughan A, Wyler E, Genetic, Environmental Risk for Alzheimer's Disease GC, Landthaler M, Hubner N, Wanker EE, Lannfelt L, Ingelsson M, Lalowski M, Voigt A, Selbach M (2015) Quantitative interaction proteomics of neurodegenerative disease proteins. Cell Rep 11 (7):1134–1146. 10.1016/j.celrep.2015.04.030 [PubMed: 25959826]
51. Grefen C, Obrdlik P, Harter K (2009) The determination of protein-protein interactions by the mating-based split-ubiquitin system (mbSUS). Methods Mol Biol 479:217–233. 10.1007/978-1-59745-289-2_14 [PubMed: 19083185]
52. Yachie N, Petsalaki E, Mellor JC, Weile J, Jacob Y, Verby M, Ozturk SB, Li S, Cote AG, Mosca R, Knapp JJ, Ko M, Yu A, Gebbia M, Sahni N, Yi S, Tyagi T, Sheykhkarimli D, Roth JF, Wong C, Musa L, Snider J, Liu YC, Yu H, Braun P, Stagljar I, Hao T, Calderwood MA, Pelletier L, Aloy P, Hill DE, Vidal M, Roth FP (2016) Pooled-matrix protein interaction screens using Barcode Fusion Genetics. Mol Syst Biol 12 (4):863 10.15252/msb.20156660 [PubMed: 27107012]
53. Gu L, Li C, Aach J, Hill DE, Vidal M, Church GM (2014) Multiplex single-molecule interaction profiling of DNA-barcoded proteins. Nature 515 (7528):554–557. 10.1038/nature13761 [PubMed: 25252978]
54. Trigg SA, Garza RM, MacWilliams A, Nery JR, Bartlett A, Castanon R, Goubil A, Feeney J, O'Malley R, Huang SC, Zhang ZZ, Galli M, Ecker JR (2017) CrY2H-seq: a massively multiplexed assay for deep-coverage interactome mapping. Nat Methods 14 (8):819–825. 10.1038/nmeth.4343 [PubMed: 28650476]
55. Wang Y, Sahni N, Vidal M (2015) Global Edgetic Rewiring in Cancer Networks. Cell Syst 1 (4):251–253. 10.1016/j.cels.2015.10.006 [PubMed: 27136053]
56. Scott DE, Bayly AR, Abell C, Skidmore J (2016) Small molecules, big targets: drug discovery faces the protein-protein interaction challenge. Nat Rev Drug Discov 15 (8):533–550. 10.1038/nrd.2016.29 [PubMed: 27050677]