Journal of proteome research

Reviews

pubs.acs.org/jpr

Assignment of Protein Interactions from Affinity Purification/Mass Spectrometry Data

Mercedes Pardo* and Jyoti S. Choudhary*

Wellcome Trust Sanger Institute, Wellcome Trust Genome Campus, Hinxton, CB10 1SA Cambridgeshire, United Kingdom

ABSTRACT: The combination of affinity purification with mass spectrometry analysis has become the method of choice for protein complex characterization. With the improved performance of mass spectrometry technology, the sensitivity of the analyses is increasing, probing deeper into molecular interactions and yielding longer lists of proteins. These identify not only core complex subunits but also the more inaccessible proteins that interact weakly or transiently. Alongside them, contaminant proteins, which are often abundant proteins in the cell, tend to be recovered in affinity experiments because they bind nonspecifically and with low affinity to matrix, tag, and/or antibody. The challenge now lies in discriminating nonspecific binders from true interactors, particularly at the low level and in a larger scale. This review aims to summarize the variety of methods that have been used to distinguish contaminants from specific interactions in the past few years, ranging from manual elimination using heuristic rules to more sophisticated probabilistic scoring approaches. We aim to give awareness on the processing that takes place before an interaction list is reported and on the different types of list curation approaches suited to the different experiments.

KEYWORDS: *affinity purification, mass spectrometry, protein−protein interaction, background contaminant, specific interactions, review*

INTRODUCTION

Proteins mostly work in association with other proteins, forming macromolecular assemblies that are crucial for many biological functions.[1] On the basis of the principle of "guilt by association", elucidation of interaction partners can provide important information about the function of novel proteins. Such information can also be used to tease out modes of action of known proteins. In recent years, protein interaction studies based on protein complex affinity purification followed by mass spectrometry analysis (AP-MS) have become increasingly common, due to technological advances in mass spectrometry. The combination of affinity purification and mass spectrometry analysis has become the method of choice for the characterization of multiprotein complexes. A protein(s) of interest can be fused to a single tag or a more sophisticated combination of various tags to allow the generic biochemical isolation of the bait, together with its associated proteins, using single step or tandem affinity purification in next to physiological conditions. Tandem affinity purification (TAP) was originally developed to achieve better discrimination of specific associations against unspecific protein background (Figure 1),[2] but it has limitations in identifying very labile or transient interactions due to the extended length of the purification. After elution from the capturing support, interacting proteins are identified by mass spectrometry.[3] Studies of this type reveal not only direct binary associations, like the yeast two-hybrid technique,[4] but also protein neighborhoods or assemblies, yielding useful contextual information. Large-scale interaction studies employing systematic approaches have now been performed in a variety of organisms, ranging from the simpler *Escherichia coli* and yeast to human.[5,6] Interaction data from these large-scale studies can be used to derive interaction networks that are useful for modeling biological processes or used to predict disease genes.[4,7,8]

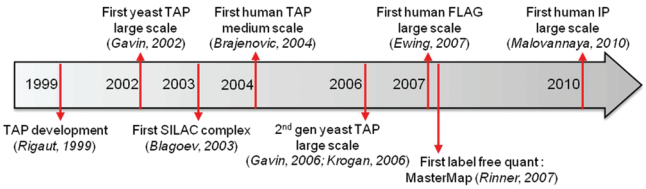

Figure 1. Timeline of major developments and studies on elucidation of protein−protein interactions from affinity purification/mass spectrometry approaches.

Generating a list of high confidence interacting proteins is the first step in the biological interpretation of identified interactions and their functionality (Figure 2). With small scale studies, validation of interactions is traditionally done by purification of the newly identified preys (by immunoprecipitation or reverse-tagging and AP) and identification of the original bait as a prey interactor. In addition, orthogonal methods such as coexpression, colocalization studies and loss-of-function studies can be used to evaluate interactions. This is normally not feasible for large-scale studies, and even medium

Received: November 24, 2011
Published: January 30, 2012

   dx.doi.org/10.1021/pr2011632 | *J. Proteome Res.* 2012, 11, 1462−1474

Figure 2. Workflow of protein−protein interaction studies based on the combination of affinity purification and mass spectrometry analysis. After isolation of a protein complex and mass spectrometry analysis for protein identification, a list of candidate interactors is generated. Discrimination of genuine interactors from background binding is a crucial step for the subsequent elucidation of accurate protein assemblies and networks. Recent comprehensive reviews on MS-based interaction proteomics that best cover each of the topics depicted are listed. This review focuses on the subject of discriminating genuine binding partners, an area only briefly addressed by previous reviews.

scale ones. Benchmarking of big interaction sets can be done by comparison with public protein interaction databases such as STRING, IntAct, BioGRID, HPRD, MINT, and BIND[4,9,10] to gain an idea of the recovery of known interactions. In medium to large scale studies teasing apart contaminants becomes a crucial part of the procedure, so that false positives do not end up populating interactions databases.

Several recent reviews address the basic principles of AP-MS, describing in detail various aspects of the workflow of MS-based interaction studies, namely the different tags available for protein tagging and techniques of affinity purification,[11−13] mass spectrometry analysis and quantification[10,14] and downstream bioinformatic processing (Figure 2).[6,10] However, the subject of discriminating genuine interactors is only briefly discussed by these review papers. Here, we try to bridge a gap in all this information while highlighting an important aspect of management of protein lists derived from interaction studies, the filtering of contaminants or discrimination of true interactors.

■ HEURISTIC MANAGEMENT

Protein abundance is an important factor contributing to protein background in affinity purification experiments. Highly expressed proteins tend to be identified in interaction data sets irrespective of the bait or control used. Such interactions tend to not be very reproducible, which can help in teasing them apart.[15] Whole proteome profiles provide a representative catalog of the major constituents of the protein complement of a given cell type. These protein lists can be used as indicators of abundant cellular proteins to control for contaminants in affinity purification experiments in those cells. Some laboratories routinely use these protein lists to facilitate assessment of experimental data from their biochemical experiments.[12,16] Schirle et al. profiled six human immortalized cell lines generating a large human proteomic data set.[16] Expanding on this, Burkard et al. recently defined a human central proteome, a combined repertoire of the most abundant proteins commonly expressed by several human cell lines.[17] Similar lists exist for *E. coli*,[18] yeast,[19] worm,[20] and mouse embryonic stem cells[21] among others.

Small/Medium Scale Studies

Management of small data sets has traditionally been approached by subtracting from the list of proteins identified in the experimental sample a list of contaminant proteins obtained from a nonrelevant control. With increased familiarity in complex analysis authors often choose to also manually remove certain protein groups, like ribosomal and/or cytoskeletal proteins, as they are normally predominant groups in the contaminant lists, even if individual proteins not always specifically appear in them.[22,23] Several approaches have been used to classify frequently observed proteins. The first medium scale TAP study conducted in human cells (Figure 1) addressing the Par pathway with nine baits defined a set of nonspecific binders (so-called "sticky" proteins) based on the frequency of their appearance in the majority of TAP purifications conducted, as judged by filtering against a control data set consisting of more than 400 TAP purifications, and eliminated them from the final data set.[24] These included heat shock proteins, ribosomal proteins, keratins, actin, myosin, and α- and β-tubulins. Bouwmeester et al. also used data sets of control purifications to filter out contaminants in their analysis of the TNF pathway based on interactions around 32 bait proteins. Here, frequencies of specific interactors were tested against the background frequency of the control data set using a binomial test and threshold *P*-values placed below a discontinuity jump in the ratio of accepted interactions to recovered known interactions.[25] Such data parsing suffers from the possibility of losing genuine high abundance interactors. In a more recent study of fission yeast deubiquitinating enzyme interaction networks involving 20 targets, Kouranti et al. wisely chose to show all identified interacting proteins, though they marked those appearing in control experiments or in more than 50% of unrelated affinity purification experiments as background.[26]

Reproducibility is a major acceptance criterion in small and medium scale studies. Some works report interactors observed in at least two replicates,[27] but with increased sensitivity yielding longer lists, a higher number of replicates is desirable, especially in cells lines from higher organisms.[12,28,29] A related strategy applied to immunoprecipitations of endogenous proteins from worm suggested that performing purifications

dx.doi.org/10.1021/pr2011632 | *J. Proteome Res.* 2012, 11, 1462−1474

with two different antibodies to the same target or two subunits of the protein complex under study was a useful discriminating approach. Proteins identified in both pull-downs are considered more likely to be true interactors.[30]

A compilation of several of the above-described approaches, termed iPAC (for Interactome by Parallel Affinity Capture), was used by Rees and colleagues in a pilot study for a large-scale project analyzing the interaction networks in *Drosophila melanogaster*.[31] Single step affinity purifications of 6 baits were performed in parallel using two different epitope tags (FLAG and StrepII tag). After elimination of proteins present in the 25 negative controls, interaction lists were generated by focusing on proteins reproducibly pulled down with the two different tags. Interaction confidence scores were assigned based on the number of times an interactor was seen in multiple pull-downs, giving greater weight to proteins detected in different affinity experiments than to those detected several times with only one affinity tag. The authors reported a BEADome, a concept first coined by Trinkle-Mulcahy et al.,[23] listing the set of contaminants common to both FLAG and StrepII resins. In addition, they also described a "TRAPome", the set of proteins not present in controls but appearing in multiple interaction lists. Prudently in this instance the TRAPome is not automatically eliminated from the final interaction lists, given the possibility that it could contain genuine interactors.

Large-scale Interaction Studies

However, the reproducibility criterion is not as easily applicable to large-scale studies, where typically replicates are not produced. Other strategies have been developed specifically for this type of analyses. Initial large-scale interaction studies performed in yeast[32,33] relied on the large number of bait experiments to address the specificity of the associated proteins identified, the reasoning being that preys copurifying frequently with unrelated baits were likely to be nonspecific binders (Figure 1). Even medium scale studies have used this frequency filter. For instance, in an analysis of 22 yeast baits by TAP followed by MudPIT MS, proteins recovered in greater than 20% of the experiments in the reference data set (containing one representative experiment for every tagged open reading frame) were automatically considered contaminants,[34] but the authors acknowledged the limitations of this approach due to all the baits being functionally related. The remaining subset still included abundant "sticky" proteins, which were further filtered out by eliminating proteins with maximum codon adaptation index >0.6.

The blanket frequency filtering schemes evolved into the "socio-affinity index"[35] and the interaction probability,[36] calculating the tendency of two proteins to appear in a complex together, be it in direct reciprocal bait-prey interactions or by co-occurrence of the pair in other purifications, or be true interactors, respectively (Figure 1). The socio-affinity index has been further refined with the integration of the relative abundance of a prey across different purifications (based on spectral counting) in a large-scale analysis of the interactome of *Mycoplasma pneumonia* using TAP/MS,[37] bypassing the need for sets of negative controls.[6] All of the identified proteins (preys) were ranked by peptide mass fingerprinting scores, which are based on spectral counts and normalized for length of the peptides and number of possible tryptic peptides per protein. The rank divided by the total number of times the prey was identified was then used to derive a modified socio-affinity index. However, it has become apparent that the filters applied in these studies are useful largely for reciprocal data sets,[38] but their application to smaller data sets is less discriminatory.[6,39]

The first published large-scale human interaction proteomic study mapped protein interactions associated to 338 FLAG-tagged baits by single affinity purification followed by MS (Figure 1).[40] An empirical filter was used to eliminate contaminant proteins and nonspecific binders. After building a database based on 202 control (vector only) immunoprecipitation experiments, all protein identifications associated with more than 2.5% of control experiments (spurious) or more than 5% of bait experiments (frequent binders) were removed, including tubulins, ribosomal and heat shock proteins. Subsequently, in an attempt to rank the interactions for subsequent data mining, an interaction confidence score was calculated for each bait-prey pair based on prey reproducibility using parameters such as total Mascot score, total number of peptides, protein rank within the experiment, maximum Mascot score across bands and best rank across bands in the experiment. Interaction confidence scores were significantly higher in the subset of known interactions compared to the novel interactions, and in the set of reciprocally observed interactions, demonstrating its validity. The authors set a threshold interaction confidence score of 0.3 for further in-depth interpretation of interaction results, because most associations between components of well-known complexes represented in the data set had scores >0.3, and also because most preys with confidence score >0.3 were identified by two or more unique peptides. Employing this threshold approximately a third of detected bait-prey interactions were selected for further investigation. In their study of 100 human protein complexes by TAP-MS, Hutchins et al. removed all identifications from experiments where no bait was identified and those recovered in 4.5% of purifications.[41] Data from purification of five related proteins (APC/C subunits) were omitted from this analysis to avoid classifying their associated proteins as contaminants.

Lambert et al. developed a method termed mCHIP for immunoaffinity isolation of protein−DNA complexes and their subsequent analysis by MS. This method identifies both direct and indirect (through chromatin) protein interactions. They then reported the first large scale mCHIP characterization of the chromatin interactome in budding yeast, based on single affinity purification of 102 TAP-tagged proteins involved in chromatin biology.[42] The first curation step removed common contaminants based on a list of identifications from control mCHIP experiments on untagged cells[43] and ribosomal proteins, common contaminants in affinity purification experiments,[11] resulting in over 5000 bait-prey associations. Next they eliminated frequent preys (present in more than 3 experiments) that were not relevant to chromatin biology by removing protein identifications associated with certain molecular function and localization terms, namely protein folding, mRNA export, fatty acid biosynthesis, ribosome biogenesis and RNA processing, mitochondria and preribosome. Other large scale interaction studies have undertaken the elimination of frequent preys in a more systematic manner, like the socio-affinity index or frequency filters as described above. However in this study the baits were functionally related, therefore often associated with common proteins, and so to indiscriminately remove all frequent binders was not appropriate.

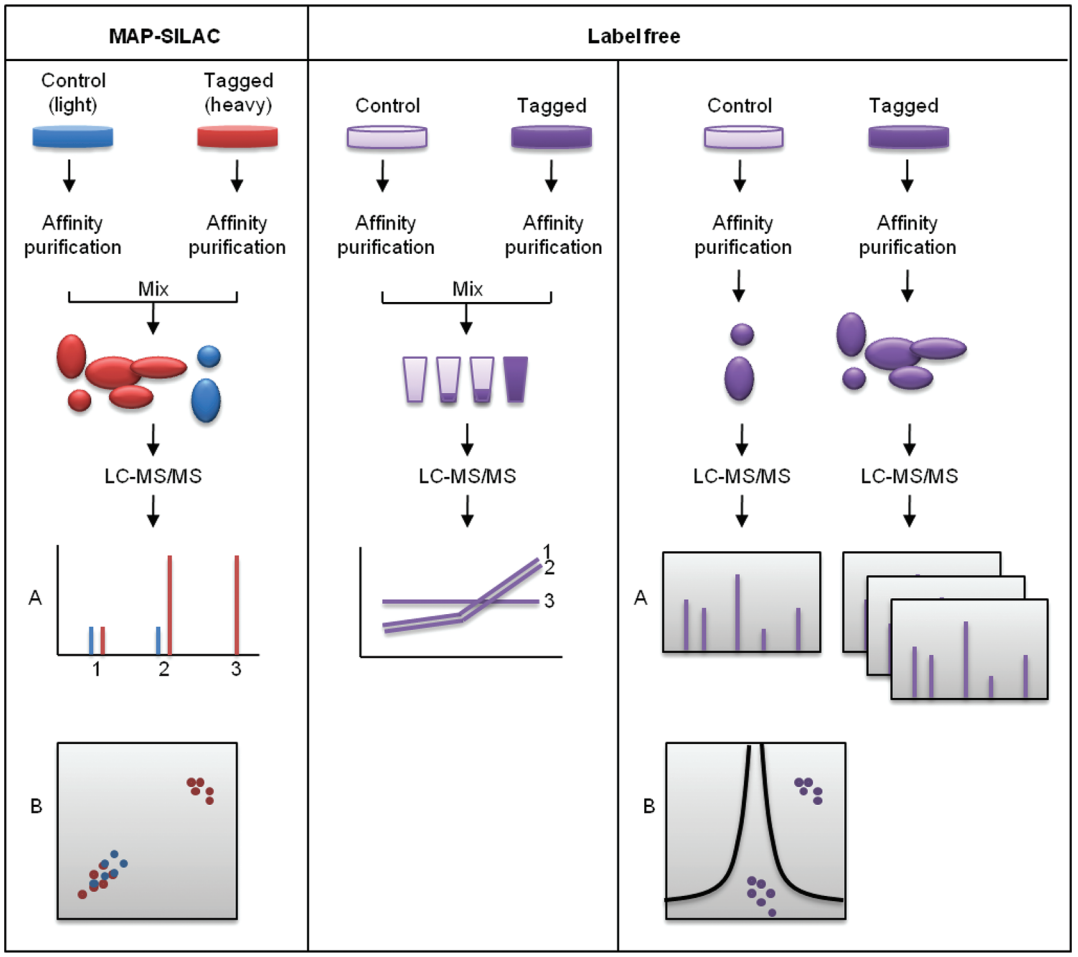

Figure 3. Quantitative approaches for the identification of true protein interactions. In the MAP-SILAC approach, proteins are labeled in culture and mixed after affinity purification prior to MS analysis. (A) Three categories of proteins are often detected: 1, contaminant proteins (ratio H/L ≈ 1); 2, interacting proteins, also present in control in small amounts (ratio H/L > 1); 3, protein complex members (not detected in control). (B) In a further refinement, experiments are done switching the labeling, and log2(ratio H/L) vs log2(ratio L/H) plotted. Specific interactors appear in the upper right corner. In the label-free quantification approach, two different strategies have been used. Control and tagged samples can be mixed in serial dilutions, and analyzed together by MS. The characteristically different quantification profiles allow separation of true interactors (2), which closely resemble the bait profile (1). Contaminants concentration does not vary with dilution (3). Alternatively, samples are not mixed but analyzed separately by MS/MS. (A) Quantification based on spectral counts, which can be subject to further mathematical/probabilistic analysis. (B) Quantification based on MS1 peptide peak height using MaxQuant. The "volcano plot" separates true interactors in the upper right corner, while background proteins remain in bottom center of the plot.

There are drawbacks to this type of heuristic management of the data. First, frequency filters are only really applicable to large-scale studies, but not to studies of individual proteins or small numbers of baits. The frequency filters are often chosen in a context or data-dependent manner, and are hard to apply if baits are functionally related and copurify with similar sets of proteins. In addition, these filtering methods are only based on binary data, and ignore quantitative features of the mass spectrometry results. Genuine abundant interactors of any given bait are often missed since they are occasionally identified as low abundance interactors with several other baits or detected in negative control experiments. This results in false negatives, as is nicely exemplified by Skarra et al.[44] The authors analyzed multiple FLAG affinity purifications of wild type phosphatase 5, a mutated form of this enzyme and a tag-only control. They first used a frequency filter of 5% to eliminate nonspecific binders. However, this resulted in the removal of all putative interactors, including the previously described Hsp90. Further processing of results using quantification and the SAINT algorithm described below uncovered 7 interaction partners and recapitulated the Hsp90 interaction. Quantitative measures of protein abundance in interaction data sets are therefore important to resolve whether the interaction bait-prey is specific. Even semiquantitative data can be helpful in teasing apart contaminants, like taking into account peptide numbers or Mascot protein scores before indiscriminate removal of any protein that appears in the negative control lists, as showed respectively in two Oct4 interactome studies.[28,45]

■ LABELING QUANTITATIVE APPROACHES

Several isotope labeling methods have been applied to discriminate true specific protein interactions in affinity purification experiments using quantitative mass spectrometry approaches, by differentially marking specific and control purifications using different heavy stable isotopes.[10,14] This strategy is based on the assumption that nonspecific binders will associate similarly to the affinity matrix in bait or control experiments, and therefore should be detected in similar

dx.doi.org/10.1021/pr2011632 | J. Proteome Res. 2012, 11, 1462–1474

amounts by mass spectrometry. Specific interactors are readily distinguishable from contaminants by their enrichment in bait compared to control experiments, and therefore will show a quantity ratio different from 1 (Figure 3). Quantitative strategies like this can be applied to single step purifications, bypassing the need to do lengthier tandem purifications that can result in loss of weaker or transient interactions.[46]

Modification of proteins with stable isotopes can be achieved chemically after trypsin digestion (ICAT, iTRAQ, dimethyl labeling) or metabolically during cell culture (SILAC), with both approaches having been applied to interaction studies.[14,47−49]

Chemical Labeling

Ranish et al. used ICAT-based quantification to characterize the yeast Pol II preinitiation complexes from a single step isolation using a DNA affinity procedure. Most ratios of specific purification:control purification followed a normal distribution and the outliers with ratios >1.2 were deemed *bona fide* complex components.[50] The same cutoff was determined empirically in another ICAT study based on immunoprecipitation of endogenous transcription factor complexes from murine cells by comparison of abundance ratios of true and nonspecific (identified in mock IPs) interactors.[51] For quantitative analysis of human androgen receptor interactions via single step AP, Jasavala et al. used gene ontology annotation to establish an arbitrary but biologically determined ICAT protein enrichment cutoff ratio of 1.5 to define true binding proteins.[52] More recently, iTRAQ was used by Pflieger et al. in the study of insulin receptor substrate interactions after single step purification from insulin-stimulated and untreated cells. Here, to discriminate real interactors, *p*-values of each ratio were calculated by fitting a log-normal distribution to the ratio values.[53] For these calculations, the distributions of ratios from enriched and nonenriched peptides have to be sufficiently different to allow statistical detection, requiring that the number of background proteins is much higher than that of the complex constituents. An inherent drawback of these approaches is the limitation of the quantification methods; i.e., iTRAQ has a specificity issue.[54]

A precursor intensity-based chemical labeling quantification strategy uses dimethyl labeling for the incorporation of stable isotopes in peptides.[55] This approach was used to identify novel targets of phosphatidylinositol 3-kinase signaling after purification of 14-3-3 binding proteins in insulin-stimulated and unstimulated human cells by affinity chromatography using the yeast counterparts.[56] Rather than a distinct boundary to separate high and low ratio (i.e., insulin dependent binding or not), a range of ratio levels was obtained, as is often the case in this type of analyses, making it difficult to establish a ratio threshold for calling an interaction. The advantage of this method is that it is cheap, quick and can be applied to any sample.

Metabolic Labeling

The most widely used labeling technique for interaction proteomics employs stable isotope labeling in culture (SILAC). Stable isotope-labeled aminoacids are metabolically incorporated into proteins during cell culture, and several versions can be used to differently label different samples, i.e., bait and control. Although the quantification used is the same, approaches for the discrimination of true interactors vary between different groups and has also evolved with time, becoming more sophisticated in recent years. Blagoev et al. first used SILAC in the identification of post-translational modification-dependent EGF-signaling interactions based on affinity purification using a GST-SH2 domain fusion as bait (Figure 1).[57] The authors set a 1.3-fold change as the threshold for defining specific interactions since this value was above the analytical error for replicate runs. A different discriminating criterion was applied in ErbB receptor family interaction screens based on peptide pull-downs with phosphorylated and nonmodified peptides, where Schulze et al. considered as significant those ratios that were significantly (2σ) different from the average ratio of the majority (mean $\pm$ σ) of identified proteins.[58] In their study of MyD88-interacting proteins, Wang et al. combined FLAG single affinity purification with amino acid coded mass tagging.[46] After calculating the relative standard deviation (RSD) of the quantitative measurements to be around 7%, they set the threshold for selection of specific interactions at 5 times the RSD, identifying 16 specific interactors with over 40% enrichment. This approach was much more sensitive than the TAP approach, which only identified the bait and one other partner with high confidence.

Trinkle-Mulcahy used the SILAC quantitative approach to distinguish specific interaction partners in single pull-down experiments using GFP-binder tagged proteins.[23] But even using this approach, discrimination of true binding partners is still problematic, particularly for the low abundance proteins where the signal-to-noise ratio is close to the background. To address this issue, Trinkle-Mulcahy opted for reduction of noise by characterizing a set of proteins that routinely associate with commonly used matrices, such as Sepharose, agarose and magnetic beads, which they termed "bead proteome". These proteins include histones, hnRNP proteins, heat shock proteins, ribosomal proteins, translation and initiation factors, DEAD box proteins and multiple cytoskeletal proteins. These should be carefully dealt with whenever they appear in protein interaction studies as they could be potential nonspecific binders.

More recently, in an approach termed BioCAT (He, 2009),[59] SILAC was combined with single step affinity purification of human 14-3-3ε using the acceptor peptide tag that can be site-specifically biotinylated *in vivo* by BirA ligase first used by others.[60,61] They used the lowest ratio of the known interactors as a cutoff threshold for discriminating specific interactors, which was considerably higher than near-to-one ratios of housekeeping proteins such as tubulins and other known contaminants.

In the QUBIC (Quantitative BAC Interactomics) approach suggested by Hubner et al. baits were labeled with GFP.[62] Two SILAC single pull-down experiments (sample + control) with GFP antibody are performed, reversing the isotope label (thus providing useful biological replicates), and then the log2 values of the ratios are plotted (i.e., log2(H/L) vs log2(L/H)). Nonspecific interactors appear in the center of the graph and specific binders in the upper right corner due to high ratios (Figure 3). A statistical measure of significance is determined by the outlier probability calculated on protein subsets obtained by intensity binning.[63] True interactors are defined by a significance <0.01 in both SILAC purifications. The same approach was used in identifying binders of specific histone marks and interactors of novel chromatin readers,[64] showing that the method was discriminatory enough to allow detection of post-translational modification-specific interactions.

Tomecki et al. combined SILAC with FLAG affinity purification of the human nuclear exosome complex and put

dx.doi.org/10.1021/pr2011632 | J. Proteome Res. 2012, 11, 1462−1474

the threshold at a SILAC ratio of 1 plus 2 times the average standard deviation of all SILAC ratios in the experiment.[65] In a similar follow-up study, the group plotted log SILAC ratios versus an abundance measure calculated as total intensity/$M_w \times 10^6$ to distinguish highly specific interactors.[66] In a much more conservative manner than the previous study, here all interactors with log SILAC ratios >0.5 were considered highly specific, with highly specific and abundant interactors appearing in the upper right-hand corner of the plot.

In a temporal analysis of cyclin/cdk complexes, Pagliuca et al. combined SILAC with single-step affinity purification of FLAG-tagged cyclins from human cells to address cell cycle stage-specific protein complex composition.[67] To distinguish true associations from nonspecific binders, the median and the median absolute deviation of the M/L and H/L ratios were calculated for each replicate, since the distribution did not follow a normal distribution. Proteins were considered to be true interactors if they were quantified at least 3 times and scored above the median plus median absolute deviation of the total distribution. On average, this resulted in a cutoff ratio of 1.8. If a protein was not quantified (due to not being present in one of the conditions), the number of unique peptide sequences identified in each SILAC state was calculated and the protein considered a true binder if the number of heavy or medium peptides identified was more than five times the number of light peptides identified. In addition, true interactors were also required to be identified in the majority of pull-downs of the considered cyclin.

The combination of SILAC-based quantification with parallel independent purifications can increase the specificity of the results. In a study of the proteins that interact with PTEN, PAP (parallel affinity purification)-SILAC used two different tags and one antibody to the endogenous protein to pull down bait and binding proteins in single affinity purifications, and discarded proteins with SILAC ratios <2 and those that only copurify in one of the pull downs, on the assumption that true interactors should be captured by two independent affinity experiments.[68] In this way they were able to remove some artifactual false positives. However, some epitope tags might affect protein interaction and preclude identification of true interactors,[69] and this method could potentially miss them.

A novel strategy has been used in the field of organelle proteomics that could also be applied to the domain of interaction proteomics. Organelle protein characterization poses a great challenge as traditional purification procedures never yield a contaminant-free sample. MCCP (multiclassifier combinatorial proteomics) combined quantitative proteomics using SILAC with bioinformatic analysis to characterize the proteome of mitotic chromosomes.[70] Ohta et al. performed a series of SILAC experiments to quantify the chromosomal association of proteins. Each experiment constitutes a so-called "classifier", which is an independent gauge of a protein's association with mitotic chromosomes (including MS data such as protein abundance in sample, enrichment vs control, and more sophisticated functionally relevant MS data such as SILAC-interaction experiments in mutant lines). Combination of classifiers' data revealed groups of proteins behaving similarly, which can then be ascribed to protein complexes or functional relationships. Integration of the data from all the classifiers using random forest analysis allowed good discrimination of chromosomal from nonchromosomal proteins. There are two important advantages to this method: it allows the use of data sets with missing proteins, that is, not every single protein has to be identified by every single experiment. And second, nonproteomic attributes can be used as classifiers and integrated in the overall analysis.

When applied to the discrimination of true interactions, SILAC poses a technical challenge whereby heavy-light exchange of dynamic interaction partners can occur in the lysates when the mixing is done before the purification, which results in heavy-light ratio equalization, and incorrect discrimination of transient interactors as contaminants.[71−74] To avoid this, sample mixing should be carried out after the purification step (so-called MAP-SILAC).[10] Taking this into account, SILAC-based quantification can be very useful in the dissection of dynamic interactions, PTM-dependent interactions, and changes in interactions between different conditions due to its high accuracy. However, quantitative approaches using stable isotope labeling are expensive, and they are not always feasible in all instances. Label-free quantification approaches have been recently exploited that overcome these drawbacks.

LABEL-FREE QUANTIFICATION: STATISTICAL/PROBABILISTIC ANALYSIS

The total number of peptides (tandem mass spectra) identifying a protein strongly correlates with the abundance of that protein.[19,75,76] The realization of this fact meant that in addition to quantification methods that rely on labeling, the relative amount of protein between different samples can also be determined by label-free quantification. Label-free quantification has mainly been used in profiling experiments aiming to uncover variations in protein expression between different states and in biomarker discovery. It has also been applied to organelle analysis, using it to assess localization of proteins to specific subcellular compartments.[77] In interaction studies, the Washburn group has used it extensively to address complex stoichiometry.[78]

Several additional methods have been developed that use the label-free quantitative information inherent to the mass spectrometry data, such as spectral counts or MS1 ion currents, to filter out nonspecific binders in AP-MS data sets. Because label-free quantification can be inherently less accurate than labeling strategies such as SILAC, studies relying on it for discrimination of true interactors from contaminating proteins require a higher number of replicates and normally carry out some form of statistical or probabilistic analysis in order to assign confidence to the results.

Precursor Intensity/Peak Area-based Quantification

Label-free quantification based on MS1 ion currents or peak area calculation was used by Rinner et al. in an innovative approach to discriminate real interactors (Figure 1).[79] The quantitative analysis of serial dilutions of HA-tag purified complexes with control purifications is based on precursor ion alignment and integration into a MasterMap. After clustering of the quantification profiles, true interactors showed increased enrichment along the dilution series, as did the bait, while concentration of contaminants remained constant in all dilutions (Figure 3). Using this strategy the authors determined FoxO3A-HA partners and could distinguish a set of 6 proteins whose profile matched almost precisely the target cluster profile (i.e., the bait). Profiles of the next 8 best candidates were not so well matched; they showed only very slight enrichment and were also present in the control sample. This approach was also used to monitor changes in complex composition under growth

dx.doi.org/10.1021/pr2011632 | J. Proteome Res. 2012, 11, 1462−1474

inhibition conditions. The MasterMap concept is technically amenable to large-scale interaction studies and could be applied to systematically monitor protein complex composition changes upon cell perturbation,[80] although in practice this has not been shown yet.

The QUBIC method described in the previous section has also been adapted as a label-free approach based on integrating total intensity from all peptides as the sum of all peak height.[62,81] Identifications and quantifications were obtained from MS data using the MaxQuant algorithm and analyzed with a standard "equal group variance" *t*-test of the observed fold change of protein intensities between samples and controls. Three replicates of sample and control experiments at least are required. The changes are plotted against the negative log *P*-value of the *t*-test, and then a volcano plot is derived where significant interactors appear in the upper right corner defined by high ratio and high *P*-value (Figure 3). True interactors are delimited by a significance line corresponding to the FDR calculated using a permutation-based method with the accompanying software Perseus, in a manner similar to that applied to microarray data.[81] There is no universal FDR as this threshold is experiment dependent, and it should be chosen empirically, where no real interactors appear to the left of the line because they should not be present in control pulldowns. After comparing with SILAC-based quantification, the authors found that label-free quantification of high resolution MS data using MaxQuant was most efficient in discriminating specific interactors from background binders.[62]

Spectral Counts-based Quantification

Spectral counts, that is, the total number of spectra that match to each identified protein in the search, can be used as a measure for quantification of proteins after single AP-MS. Normalized spectral abundance factors (NSAF) are calculated based on the total number of spectra identifying each protein, normalized by the protein's length and the total number of identified spectra in the sample, to estimate the relative protein abundance, allowing comparison of protein levels across different experiments.[78,82,83] The normalization of spectral counts can also be applied to complex components to estimate the relative levels of each subunit within the complex.[78,84,85] Sardiu et al. used NSAF values to discriminate true interactors from contaminant binders in a medium scale one-step AP-MS proteomic experiment of 27 related FLAG-tagged baits and 35 untagged controls.[86] Each prey was defined by two vectors consisting of NSAF values in specific and control purifications. A protein was considered a contaminant if the vector ratio magnitude between the two sets was >1. When analyzing protein complexes, sometimes a distinction can be made between core components and accessory proteins. To discriminate between these, a matrix was then created representing all NSAF values for all preys vs all baits, and proteins enriched from the immunoprecipitations, that is, contributing most to the matrix, were extracted by a rank estimated method after single value decomposition. Of the 945 proteins initially classified as specific binders, only 125 were found to be the most enriched in the IPs. Hierarchical clustering was performed on this set to define core complexes. The NSAF values were clustered using Pearson correlation and unweighted paired group average linkage, resulting on good separation of core components of complexes. Finally, they generated a probabilistic measure of the preference of proteins to interact with one another using a Bayesian approach based on the NSAF alone, rather than on reciprocity of bait-prey interactions or copurification with a third bait. This probability represents the preference of a protein to interact with any given bait. This work signifies a very detailed dissection of protein interaction and organization from AP-MS data.

In a recent refinement of the NSAF method, Mosley et al. calculated distributed normalized spectral abundance factors (dNSAF) to correct for differences in protein length and to avoid redundant spectral assignment.[87] In a small scale analysis of TAP-purified protein complexes from four related baits in yeast, preys with a dNSAF value in the purification experiment < two times the value in the control were discarded as nonspecific binders. Mak et al. also used NSAF to enhance the specificity of their interaction study. They performed a two-tailed *t* test on the preys associated to 19 human tagged baits versus a tagged-eGFP and a no tag control to filter out background contaminants.[29] Out of 1916 total preys identified, only 222 had significant p-values and were marked as confident interactors. For further stringency, the authors only listed preys reproducibly identified in all three replicate experiments.

Very recently, Guruharsha et al. developed a scoring system to generate a large-scale *Drosophila* interaction map based on single AP-MS.[88] The HGSCore (HyperGeometric Spectral Counts score) method scores the probability of each protein interacting with another based on a hypergeometric distribution error model modified to take into account the total spectral counts. Normalized spectral abundance factors (NSAF) were transformed to compress the range of values, and the new values converted into a matrix model representation capturing the quantitative aspect of the interactions. The compressed values were used to calculate the hypergeometric probability of observing an interaction between a pair of proteins.

Sowa et al. developed an automated software platform called CompPASS to identify true binding partners in single affinity interaction proteomics experiments based on label-free quantification data also using total spectral counts.[38] The data processing module of CompPASS assigns two scores, the Z-score and the D-score, to proteins identified in parallel proteomic data sets using an unbiased comparative method. A conventional Z-score is calculated for each protein based on the total spectral counts of the protein across all experiments. The Z-score is useful for proteins present in high amount in a subset of experiments. The D-score also takes into account the uniqueness of the prey, its abundance and the reproducibility of the interaction, and is assigned to each protein in each pull-down experiment. All the raw D-scores (D^R) are normalized to a global D-score threshold (D^T), calculated from mock random data sets simulating the actual data collected in their experiments, yielding the D^N-scores. True interactors in each immunoprecipitation are those with D^N-score ≥ 1. This score is useful in identifying true interactors that are relatively unique. For proteins with D^N-score ≤ 1, examination of the Z-score is needed to establish if the protein is a candidate true interactor. *P*-values calculated from the mean and standard deviation of the TSCs of the simulated data representing the probability that a given TSC for a specific interactor would occur by chance can be used to further select candidate interacting proteins. Behrends et al. have recently updated the CompPASS platform to make it applicable to reciprocate experiments by calculating a weighted WD^N score that can discriminate high confidence interacting proteins associated with multiple baits in a network.[89]

dx.doi.org/10.1021/pr2011632 | *J. Proteome Res.* 2012, 11, 1462–1474

Table 1. Details of Interaction Proteomics Studies Mentioned in This Review

study reference	no. baits	organism	no. interactors (true/total identified)	purification method	contaminant analysis
Gavin, 2002	589	Yeast	1374/1440	TAP	Frequency filtering + untagged control
Ho, 2002	600	Yeast	1578/3617	Single AP (FLAG)	Frequency filtering + untagged control + ribosome
Bouwmeester, 2004	32	Human	131/680	TAP	Unrelated control data sets (p-value of interactor frequency tested against background frequency)
Brajenovic, 2004	9	Human		TAP	Sticky set based on 400 TAPs
Gavin, 2006	1993	Yeast		TAP	Socio-affinity index
Krogan, 2006	2357	Yeast	2708/4087	TAP	Interaction probability
Ewing, 2007	338	Human	2235/2826	Single AP (FLAG)	Frequency filter
Kühner, 2009	212	*Mycoplasma*	1058/10083[a]	TAP	Socio-affinity index + normalized spectral count
Kouranti, 2010	20	Yeast		TAP	Frequency filter (50%) + control purifications
Lambert, 2010	102	Yeast		Single AP (TAP)	Control proteins Biased frequency filter
Hutchins, 2010	239	Human	936/1353	TAP	Controls + Frequency filter (4.5%)
Pardo, 2010	1	Mouse	92/692	TAP	Controls (weighed, peptide numbers)
Rees, 2011	6	Fly		Single AP (FLAG and StrepII)	Control untagged + BEADome
Ranish, 2003	1	Yeast		Single step DNA affinity	ICAT ratio >1.2
Jasavala, 2007	1	Human	181/421	Single AP (IgG binding domain)	ICAT ratio >1.5
Pflieger, 2008	1	Fly	3/23	Single AP (HA)	Significant iTRAQ ratio based on ratios distribution
Blagoev, 2003	1	Human	28/228	Single GST pull down	SILAC ratio >1.3
Schulze, 2005	94	Human		Peptide AP	SILAC ratio > average ratio $\pm 2\sigma$
Wang, 2005	1	Mouse	16/42	Single AP (FLAG)	SILAC enrichment >5 x RSD
Trinkle-Mulcahy, 2008	1	Human		Single AP (GFP-binder)	SILAC + beadome
Hubner, 2010	7	Human		Single AP (GFP)	SILAC (QUBIC)/MaxQuant
Tomecki, 2010	3	Human		Single AP (FLAG)	SILAC ratio >1 ± 2s.d.
Gunaratne, 2011	1	Human		Single AP (FLAG and GFP trap), immunoprecipitation	SILAC ratio >2 Pull down with two independent matrices required
Lubas, 2011	2	Human		Single AP (FLAG)	Log SILAC ratio >0.5
Pagliuca, 2011	3	Human		Single AP (FLAG)	SILAC ratio > median ratio ± median absolute deviation or no. peptides H > 5xL
Rinner, 2007	1	Human		Single AP (HA)	Label-free quantification of serial dilutions with control (MasterMap)
Sardiu, 2008	27	Human	945/1278	Single AP (FLAG)	Label-free NSAF
Sowa, 2009	75	Human	774/2458	Single AP (HA)	Label-free/CompPASS
Behrends, 2010	65	Human	409/2553	Single AP (HA)	Label-free/Modified CompPASS
Breitkreutz, 2010	276	Yeast		Single AP (HA or FLAG)	Label-free/SAINT
Hubner, 2010	3	Human		Single AP (GFP)	Label-free/MaxQuant QUBIC*validator*
Mak, 2010	19	Human	222/1916	TAP (FLAG-His or Strep-FLAG)	Label-free/Two-tailed *t* test of NSAFs of bait vs untagged and eGFP controls
Malovannaya, 2010		Human		Immunoprecipitation	Label-free enrichment (spectral count based abundance) in IP vs enrichment in total cell extract
Skarra, 2011	2	Human	7/503	Single AP (FLAG)	Label-free/SAINT
Jäger, 2011	n.a.	HIV	n.a.	Single AP (FLAG and StrepII)	Label-free/SIN
Mosley, 2011	4	Yeast		TAP	Label-free dNSAF(bait expt)/dNSAF(control) > 2

[a]Interactions, not interactors.

The SAINT (significance analysis of interactome) approach was originally designed to analyze the yeast kinase and phosphatase interactome after single AP,[69] and has been generalized to make it applicable to a variety of different scale data sets, including or not negative controls, and even to the analysis of a single bait.[44] This statistical method computes an interaction probability based on the spectral counts of a particular prey protein identified in the purification of a bait, taking into account all interactions that involve said prey and bait. The spectral counts are normalized to the length of the proteins and to the total number of spectra in the purification. The method then models distributions for true and false interactions to calculate the probability of a genuine protein−protein interaction. When negative controls exist, they are used to model the spectral count distribution for false interactions. The algorithm can also work without negative controls, provided there is sufficient number of independent baits that are not densely interconnected. After the probability scores are calculated, interactions can be ordered by decreasing probability. The average of SAINT probabilities can be used as a threshold considered as the false discovery rate. Exceptionally, the authors compared the performance of their algorithm with ComPASS[38] and PP-NSAF[86] by assessing the presence of selected interactions by each method in interaction databases BioGrid and iRefWeb, and the co-occurrence of GO Biological Process terms in interaction partners, and concluded that SAINT interactions were better overlapping with databases and showed higher rates of GO coannotation when using negative controls. When probing large data sets without

dx.doi.org/10.1021/pr2011632 | *J. Proteome Res.* 2012, 11, 1462−1474

negative controls, SAINT and ComPASS performed similarly, albeit SAINT behaved slightly more conservatively.

A different strategy based on label-free quantification information was applied to the analysis of >1000 immunoprecipitation experiments on endogenous human proteins (Figure 1).[90] First, the input, the loose precipitate and the packed precipitate were analyzed by mass spectrometry to semi-quantitatively compare their protein composition to that of the immunoprecipitated complex. Fractional contributions of proteins to their corresponding list were calculated based on spectral counts and protein size, and then enriched identifications defined as those who fractional contribution in the IP is bigger than a multiple of the fractional contributions in one of the other fractions (in this particular case 5 was chosen as the adequate value). They then established a cutoff threshold to discriminate enriched proteins from background. After observing that likely nonspecific binders are frequently identified with characteristically different spectral counts distributions, they carried out a statistical quartile analysis of spectral counts distributions. The extreme upper hand outlier threshold was identified as a suitable cutoff filter, and proteins with lower spectral counts were eliminated as contaminants. What this implies in practice is that proteins were always judged specific if present in <25% of experiments, and more recurrent proteins were eliminated if present at background levels.

Building on this analysis, the same authors developed a semiquantitative approach that they have termed Near Neighbor Analysis to assign proteins to complexes. For this analysis, a group of ribosomal, cytoskeletal and heat shock proteins were manually eliminated, due to the ambiguous identification of the different isoforms and them being ubiquitous. The basis of the 3N analysis is the co-occurrence of pairs of proteins, and the approach is bait-independent. In this strategy, a protein-centered top IP subset is selected, that is, the subset of IPs were a certain protein (called "seed", not necessarily the bait) is present at higher spectral counts. Then all proteins from the top IPs are sorted by their co-occurrence with the initially selected protein. Specific interactors must co-occur at least 3 times with the selected protein in independent experiments, and only two replicate IPs are allowed. A statistical distance-based interaction proximity cutoff is then calculated, based on the assumption that ratios between complex components in different experiments should be similar, to discriminate near-neighbors of the selected protein. Comparison of iterative related 3N analyses on different seed proteins reveals associations between complexes and can distinguish between core components of the complex and frequently interacting proteins. This strategy was applied to the elucidation of complexes and in vivo associations of human nuclear receptors.[91]

In a methods paper, Jäger et al. have proposed to use the label-free SI_N (normalized spectral index) normalization method[92] to address the specificity and reproducibility of interactors identified in single affinity purification studies of HIV-host complexes carrying out parallel purifications using two different tags.[15] The SI_N method incorporates fragment-ion (ms/ms) intensity as a measure of protein abundance, combining it with peptide and spectral counts, and claims to be more accurate than other current methods.[92] It will be interesting to see how this method performs when applied to interaction data.

CONCLUSION

In interaction studies using AP-MS, as with any experiment really, it is of the utmost importance to define what the aim of the analysis is in order to choose the best experimental strategy. The variety of techniques and data processing procedures available can yield very different results. For instance, is the objective to isolate/characterize a core complex or cellular machinery, or in contrast, to capture transient interactions? This will also impact in the validation approach for the discrimination of true interactors, since the multitude of AP-MS strategies result in data of very variable quality, therefore requiring either dedicated threshold criteria to yield a confident set or a careful assessment of the cutoff criteria used. Table 1 summarizes details of the interaction studies mentioned here. We hope this review has successfully conveyed that the approach to follow in the discrimination of true interactors is very much sample and experiment dependent. A comparison of the most common approaches and methods is shown in Table 2.

Table 2. Summary of Most Common Contaminant Discrimination Approaches and Details of Their Application

filtering approach	quantitative	study scale	AP method	interaction type
Control purification	No	Small Medium	TAP	Stable, core complex
Frequency filtering	No	Medium Large	TAP Single AP	Stable, core complex
MAP-SILAC ratio (MaxQuant)	Yes	Small	Single AP	Weak, transient, complex dynamics
PAM-SILAC ratio (MaxQuant)	Yes	Small	Single AP	Stable, core complex
Label-free (MasterMap)	Yes	Small Medium? Large?	Single AP	Varied
Label-free quant (MaxQuant/NSAF/SAINT/CompPASS)	Yes	Small Medium Large	Single AP	Varied

Recent advances in quantitative mass spectrometry have resulted in quantitative analysis gaining a bigger role in the discrimination of false positives from AP-MS interaction data sets. Only 10 years ago, the development of the TAP-tag was a major advancement in the elucidation of true interactors (Figure 1). But with the development of SILAC, the popularity of the methodology for protein complex isolation shifted from the tandem affinity purification to quicker but dirtier single step purifications that maintain transient or weaker interactions but obviously carry more background proteins. Due to the lower cost of the analysis compared to stable-isotope labeling, label-free quantification of interaction data sets and the use of this data for discrimination of contaminants has seen a substantial increase in use in the past few years.

We have observed that quantification of label-free samples by MaxQuant analysis is not ideal for the analysis of experimental conditions where controls are very clean, due to the normalization algorithms used. Similarly, the SAINT algorithm is also better suited for dirtier purifications, where controls contain a large number of proteins, and therefore bait and control data are more similar, that is, data sets derived from

single affinity purifications (Nesvizhskii, personal communication). Label-free quantification based on precursor ion intensity (peak height) relies on accurate peak alignment and so the analytical conditions must be very reproducible, ideally with all samples analyzed on the same chromatographic and mass spectrometry run. Carried out in this careful manner, this method of quantification is more accurate than methods based on spectral counting.[93] However, when these conditions are not met, spectral counting might be more appropriate.

The presence of contaminants in the resulting interaction lists is almost a "must-have" for the mathematical management of the data, helping in the discrimination of true interactors by statistical analysis. Several laboratories have developed a plethora of probabilistic approaches to address this important challenge. However, although many of the algorithms described above are freely accessible and implementable (i.e., MaxQuant/Perseus, SAINT, CompPASS, dNSAF), they are being very rarely used by other laboratories. All the probabilistic analyses described above were developed by individual laboratories for a particular application, and the driver for the application, that is, the exact characteristics of the data and the logic for the algorithm, are rarely reported. The lack of documentation makes it difficult for other laboratories to apply them adequately to their own data and prevents their wider usage. This is a common problem in science and the recently launched Science Code Manifesto (http://sciencecodemanifesto.org/) aims to rectify it by promoting full and continued availability of source code and platform description from software creators, together with proper recognition from software users, institutions and funding bodies.

In conclusion, it is important to address the issue carefully and disclose in detail the criteria used in methods' sections. With full MS data now required to be freely accessible by many journals the full extent of interaction lists filtering is also becoming more transparent. The difficulty resides in striking a balance between being conservative in order to avoid populating databases with false positive interactions, and also avoiding bias in the elimination of contaminants, the ever-difficult balance of sensitivity vs specificity.

AUTHOR INFORMATION

Corresponding Author

*Tel: +44 1223 494843. Fax: +44 1223 494919. E-mail: mp3@sanger.ac.uk, jc4@sanger.ac.uk.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

We thank Walter Blackstock and Mark Collins for critical reading of the manuscript. The authors were supported by the Wellcome Trust [079643/Z/06/Z].

REFERENCES

(1) Alberts, B. The cell as a collection of protein machines: preparing the next generation of molecular biologists. *Cell* **1998**, *92* (3), 291–4.
(2) Rigaut, G.; Shevchenko, A.; Rutz, B.; Wilm, M.; Mann, M.; Seraphin, B. A generic protein purification method for protein complex characterization and proteome exploration. *Nat. Biotechnol.* **1999**, *17* (10), 1030–2.
(3) Walther, T. C.; Mann, M. Mass spectrometry-based proteomics in cell biology. *J. Cell Biol.* **2010**, *190* (4), 491–500.
(4) Charbonnier, S.; Gallego, O.; Gavin, A. C. The social network of a cell: recent advances in interactome mapping. *Biotechnol. Annu. Rev.* **2008**, *14*, 1–28.
(5) Figeys, D. Mapping the human protein interactome. *Cell Res.* **2008**, *18* (7), 716–24.
(6) Gavin, A. C.; Maeda, K.; Kuhner, S. Recent advances in charting protein-protein interaction: mass spectrometry-based approaches. *Curr Opin Biotechnol* **2011**, *22* (1), 42–9.
(7) Ruffner, H.; Bauer, A.; Bouwmeester, T. Human protein-protein interaction networks and the value for drug discovery. *Drug Discov. Today* **2007**, *12* (17–18), 709–16.
(8) Sang, L.; Miller, J. J.; Corbit, K. C.; Giles, R. H.; Brauer, M. J.; Otto, E. A.; Baye, L. M.; Wen, X.; Scales, S. J.; Kwong, M.; Huntzicker, E. G.; Sfakianos, M. K.; Sandoval, W.; Bazan, J. F.; Kulkarni, P.; Garcia-Gonzalo, F. R.; Seol, A. D.; O'Toole, J. F.; Held, S.; Reutter, H. M.; Lane, W. S.; Rafiq, M. A.; Noor, A.; Ansar, M.; Devi, A. R.; Sheffield, V. C.; Slusarski, D. C.; Vincent, J. B.; Doherty, D. A.; Hildebrandt, F.; Reiter, J. F.; Jackson, P. K. Mapping the NPHP-JBTS-MKS protein network reveals ciliopathy disease genes and pathways. *Cell* **2011**, *145* (4), 513–28.
(9) Ethier, M.; Lambert, J. P.; Vasilescu, J.; Figeys, D. Analysis of protein interaction networks using mass spectrometry compatible techniques. *Anal. Chim. Acta* **2006**, *564* (1), 10–8.
(10) Kaake, R. M.; Wang, X.; Huang, L. Profiling of protein interaction networks of protein complexes using affinity purification and quantitative mass spectrometry. *Mol. Cell. Proteomics* **2010**, *9* (8), 1650–65.
(11) Gingras, A. C.; Gstaiger, M.; Raught, B.; Aebersold, R. Analysis of protein complexes using mass spectrometry. *Nat. Rev. Mol. Cell. Biol.* **2007**, *8* (8), 645–54.
(12) Kocher, T.; Superti-Furga, G. Mass spectrometry-based functional proteomics: from molecular machines to protein networks. *Nat. Methods* **2007**, *4* (10), 807–15.
(13) Collins, M. O.; Choudhary, J. S. Mapping multiprotein complexes by affinity purification and mass spectrometry. *Curr. Opin. Biotechnol.* **2008**, *19* (4), 324–30.
(14) Vermeulen, M.; Hubner, N. C.; Mann, M. High confidence determination of specific protein-protein interactions using quantitative mass spectrometry. *Curr. Opin. Biotechnol.* **2008**, *19* (4), 331–7.
(15) Jager, S.; Gulbahce, N.; Cimermancic, P.; Kane, J.; He, N.; Chou, S.; D'Orso, I.; Fernandes, J.; Jang, G.; Frankel, A. D.; Alber, T.; Zhou, Q.; Krogan, N. J. Purification and characterization of HIV-human protein complexes. *Methods* **2011**, *53* (1), 13–9.
(16) Schirle, M.; Heurtier, M. A.; Kuster, B. Profiling core proteomes of human cell lines by one-dimensional PAGE and liquid chromatography-tandem mass spectrometry. *Mol. Cell. Proteomics* **2003**, *2* (12), 1297–305.
(17) Burkard, T. R.; Planyavsky, M.; Kaupe, I.; Breitwieser, F. P.; Burckstummer, T.; Bennett, K. L.; Superti-Furga, G.; Colinge, J. Initial characterization of the human central proteome. *BMC Syst. Biol.* **2011**, *5*, 17.
(18) Franzel, B.; Wolters, D. A. Advanced MudPIT as a next step toward high proteome coverage. *Proteomics* **2011**, *11* (18), 3651–6.
(19) Washburn, M. P.; Wolters, D.; Yates, J. R. 3rd Large-scale analysis of the yeast proteome by multidimensional protein identification technology. *Nat. Biotechnol.* **2001**, *19* (3), 242–7.
(20) Mawuenyega, K. G.; Kaji, H.; Yamuchi, Y.; Shinkawa, T.; Saito, H.; Taoka, M.; Takahashi, N.; Isobe, T. Large-scale identification of Caenorhabditis elegans proteins by multidimensional liquid chromatography-tandem mass spectrometry. *J. Proteome Res.* **2003**, *2* (1), 23–35.
(21) Nagano, K.; Taoka, M.; Yamauchi, Y.; Itagaki, C.; Shinkawa, T.; Nunomura, K.; Okamura, N.; Takahashi, N.; Izumi, T.; Isobe, T. Large-scale identification of proteins expressed in mouse embryonic stem cells. *Proteomics* **2005**, *5* (5), 1346–61.
(22) Glatter, T.; Wepf, A.; Aebersold, R.; Gstaiger, M. An integrated workflow for charting the human interaction proteome: insights into the PP2A system. *Mol. Syst. Biol.* **2009**, *5*, 237.

(23) Trinkle-Mulcahy, L.; Boulon, S.; Lam, Y. W.; Urcia, R.; Boisvert, F. M.; Vandermoere, F.; Morrice, N. A.; Swift, S.; Rothbauer, U.; Leonhardt, H.; Lamond, A. Identifying specific protein interaction partners using quantitative mass spectrometry and bead proteomes. *J. Cell Biol.* **2008**, *183* (2), 223−39.
(24) Brajenovic, M.; Joberty, G.; Kuster, B.; Bouwmeester, T.; Drewes, G. Comprehensive proteomic analysis of human Par protein complexes reveals an interconnected protein network. *J. Biol. Chem.* **2004**, *279* (13), 12804−11.
(25) Bouwmeester, T.; Bauch, A.; Ruffner, H.; Angrand, P. O.; Bergamini, G.; Croughton, K.; Cruciat, C.; Eberhard, D.; Gagneur, J.; Ghidelli, S.; Hopf, C.; Huhse, B.; Mangano, R.; Michon, A. M.; Schirle, M.; Schlegl, J.; Schwab, M.; Stein, M. A.; Bauer, A.; Casari, G.; Drewes, G.; Gavin, A. C.; Jackson, D. B.; Joberty, G.; Neubauer, G.; Rick, J.; Kuster, B.; Superti-Furga, G. A physical and functional map of the human TNF-alpha/NF-kappa B signal transduction pathway. *Nat. Cell Biol.* **2004**, *6* (2), 97−105.
(26) Kouranti, I.; McLean, J. R.; Feoktistova, A.; Liang, P.; Johnson, A. E.; Roberts-Galbraith, R. H.; Gould, K. L. A global census of fission yeast deubiquitinating enzyme localization and interaction networks reveals distinct compartmentalization profiles and overlapping functions in endocytosis and polarity. *PLoS Biol.* **2010**, *8* (9), e1000471.
(27) Goudreault, M.; D'Ambrosio, L. M.; Kean, M. J.; Mullin, M. J.; Larsen, B. G.; Sanchez, A.; Chaudhry, S.; Chen, G. I.; Sicheri, F.; Nesvizhskii, A. I.; Aebersold, R.; Raught, B.; Gingras, A. C. A PP2A phosphatase high density interaction network identifies a novel striatin-interacting phosphatase and kinase complex linked to the cerebral cavernous malformation 3 (CCM3) protein. *Mol. Cell. Proteomics* **2009**, *8* (1), 157−71.
(28) Pardo, M.; Lang, B.; Yu, L.; Prosser, H.; Bradley, A.; Babu, M. M.; Choudhary, J. An expanded Oct4 interaction network: implications for stem cell biology, development, and disease. *Cell Stem Cell* **2010**, *6* (4), 382−95.
(29) Mak, A. B.; Ni, Z.; Hewel, J. A.; Chen, G. I.; Zhong, G.; Karamboulas, K.; Blakely, K.; Smiley, S.; Marcon, E.; Roudeva, D.; Li, J.; Olsen, J. B.; Wan, C.; Punna, T.; Isserlin, R.; Chetyrkin, S.; Gingras, A. C.; Emili, A.; Greenblatt, J.; Moffat, J. A lentiviral functional proteomics approach identifies chromatin remodeling complexes important for the induction of pluripotency. *Mol. Cell. Proteomics* **2010**, *9* (5), 811−23.
(30) Cheeseman, I. M.; Niessen, S.; Anderson, S.; Hyndman, F.; Yates, J. R. 3rd; Oegema, K.; Desai, A. A conserved protein network controls assembly of the outer kinetochore and its ability to sustain tension. *Genes Dev.* **2004**, *18* (18), 2255−68.
(31) Rees, J. S.; Lowe, N.; Armean, I. M.; Roote, J.; Johnson, G.; Drummond, E.; Spriggs, H.; Ryder, E.; Russell, S. St; Johnston, D.; Lilley, K. S. In vivo analysis of proteomes and interactomes using Parallel Affinity Capture (iPAC) coupled to mass spectrometry. *Mol. Cell. Proteomics* **2011**, *10* (6), M110 002386.
(32) Gavin, A. C.; Bosche, M.; Krause, R.; Grandi, P.; Marzioch, M.; Bauer, A.; Schultz, J.; Rick, J. M.; Michon, A. M.; Cruciat, C. M.; Remor, M.; Hofert, C.; Schelder, M.; Brajenovic, M.; Ruffner, H.; Merino, A.; Klein, K.; Hudak, M.; Dickson, D.; Rudi, T.; Gnau, V.; Bauch, A.; Bastuck, S.; Huhse, B.; Leutwein, C.; Heurtier, M. A.; Copley, R. R.; Edelmann, A.; Querfurth, E.; Rybin, V.; Drewes, G.; Raida, M.; Bouwmeester, T.; Bork, P.; Seraphin, B.; Kuster, B.; Neubauer, G.; Superti-Furga, G. Functional organization of the yeast proteome by systematic analysis of protein complexes. *Nature* **2002**, *415* (6868), 141−7.
(33) Ho, Y.; Gruhler, A.; Heilbut, A.; Bader, G. D.; Moore, L.; Adams, S. L.; Millar, A.; Taylor, P.; Bennett, K.; Boutilier, K.; Yang, L.; Wolting, C.; Donaldson, I.; Schandorff, S.; Shewnarane, J.; Vo, M.; Taggart, J.; Goudreault, M.; Muskat, B.; Alfarano, C.; Dewar, D.; Lin, Z.; Michalickova, K.; Willems, A. R.; Sassi, H.; Nielsen, P. A.; Rasmussen, K. J.; Andersen, J. R.; Johansen, L. E.; Hansen, L. H.; Jespersen, H.; Podtelejnikov, A.; Nielsen, E.; Crawford, J.; Poulsen, V.; Sorensen, B. D.; Matthiesen, J.; Hendrickson, R. C.; Gleeson, F.; Pawson, T.; Moran, M. F.; Durocher, D.; Mann, M.; Hogue, C. W.; Figeys, D.; Tyers, M. Systematic identification of protein complexes in Saccharomyces cerevisiae by mass spectrometry. *Nature* **2002**, *415* (6868), 180−3.
(34) Graumann, J.; Dunipace, L. A.; Seol, J. H.; McDonald, W. H.; Yates, J. R. 3rd; Wold, B. J.; Deshaies, R. J. Applicability of tandem affinity purification MudPIT to pathway proteomics in yeast. *Mol. Cell. Proteomics* **2004**, *3* (3), 226−37.
(35) Gavin, A. C.; Aloy, P.; Grandi, P.; Krause, R.; Boesche, M.; Marzioch, M.; Rau, C.; Jensen, L. J.; Bastuck, S.; Dumpelfeld, B.; Edelmann, A.; Heurtier, M. A.; Hoffman, V.; Hoefert, C.; Klein, K.; Hudak, M.; Michon, A. M.; Schelder, M.; Schirle, M.; Remor, M.; Rudi, T.; Hooper, S.; Bauer, A.; Bouwmeester, T.; Casari, G.; Drewes, G.; Neubauer, G.; Rick, J. M.; Kuster, B.; Bork, P.; Russell, R. B.; Superti-Furga, G. Proteome survey reveals modularity of the yeast cell machinery. *Nature* **2006**, *440* (7084), 631−6.
(36) Krogan, N. J.; Cagney, G.; Yu, H.; Zhong, G.; Guo, X.; Ignatchenko, A.; Li, J.; Pu, S.; Datta, N.; Tikuisis, A. P.; Punna, T.; Peregrin-Alvarez, J. M.; Shales, M.; Zhang, X.; Davey, M.; Robinson, M. D.; Paccanaro, A.; Bray, J. E.; Sheung, A.; Beattie, B.; Richards, D. P.; Canadien, V.; Lalev, A.; Mena, F.; Wong, P.; Starostine, A.; Canete, M. M.; Vlasblom, J.; Wu, S.; Orsi, C.; Collins, S. R.; Chandran, S.; Haw, R.; Rilstone, J. J.; Gandi, K.; Thompson, N. J.; Musso, G. St; Onge, P.; Ghanny, S.; Lam, M. H.; Butland, G.; Altaf-Ul, A. M.; Kanaya, S.; Shilatifard, A.; O'Shea, E.; Weissman, J. S.; Ingles, C. J.; Hughes, T. R.; Parkinson, J.; Gerstein, M.; Wodak, S. J.; Emili, A.; Greenblatt, J. F. Global landscape of protein complexes in the yeast Saccharomyces cerevisiae. *Nature* **2006**, *440* (7084), 637−43.
(37) Kuhner, S.; van Noort, V.; Betts, M. J.; Leo-Macias, A.; Batisse, C.; Rode, M.; Yamada, T.; Maier, T.; Bader, S.; Beltran-Alvarez, P.; Castano-Diez, D.; Chen, W. H.; Devos, D.; Guell, M.; Norambuena, T.; Racke, I.; Rybin, V.; Schmidt, A.; Yus, E.; Aebersold, R.; Herrmann, R.; Bottcher, B.; Frangakis, A. S.; Russell, R. B.; Serrano, L.; Bork, P.; Gavin, A. C. Proteome organization in a genome-reduced bacterium. *Science* **2009**, *326* (5957), 1235−40.
(38) Sowa, M. E.; Bennett, E. J.; Gygi, S. P.; Harper, J. W. Defining the human deubiquitinating enzyme interaction landscape. *Cell* **2009**, *138* (2), 389−403.
(39) Collins, S. R.; Kemmeren, P.; Zhao, X. C.; Greenblatt, J. F.; Spencer, F.; Holstege, F. C.; Weissman, J. S.; Krogan, N. J. Toward a comprehensive atlas of the physical interactome of Saccharomyces cerevisiae. *Mol. Cell. Proteomics* **2007**, *6* (3), 439−50.
(40) Ewing, R. M.; Chu, P.; Elisma, F.; Li, H.; Taylor, P.; Climie, S.; McBroom-Cerajewski, L.; Robinson, M. D.; O'Connor, L.; Li, M.; Taylor, R.; Dharsee, M.; Ho, Y.; Heilbut, A.; Moore, L.; Zhang, S.; Ornatsky, O.; Bukhman, Y. V.; Ethier, M.; Sheng, Y.; Vasilescu, J.; Abu-Farha, M.; Lambert, J. P.; Duewel, H. S.; Stewart, I. I.; Kuehl, B.; Hogue, K.; Colwill, K.; Gladwish, K.; Muskat, B.; Kinach, R.; Adams, S. L.; Moran, M. F.; Morin, G. B.; Topaloglou, T.; Figeys, D. Large-scale mapping of human protein-protein interactions by mass spectrometry. *Mol. Syst. Biol.* **2007**, *3*, 89.
(41) Hutchins, J. R.; Toyoda, Y.; Hegemann, B.; Poser, I.; Heriche, J. K.; Sykora, M. M.; Augsburg, M.; Hudecz, O.; Buschhorn, B. A.; Bulkescher, J.; Conrad, C.; Comartin, D.; Schleiffer, A.; Sarov, M.; Pozniakovsky, A.; Slabicki, M. M.; Schloissnig, S.; Steinmacher, I.; Leuschner, M.; Ssykor, A.; Lawo, S.; Pelletier, L.; Stark, H.; Nasmyth, K.; Ellenberg, J.; Durbin, R.; Buchholz, F.; Mechtler, K.; Hyman, A. A.; Peters, J. M. Systematic analysis of human protein complexes identifies chromosome segregation proteins. *Science* **2010**, *328* (5978), 593−9.
(42) Lambert, J. P.; Fillingham, J.; Siahbazi, M.; Greenblatt, J.; Baetz, K.; Figeys, D. Defining the budding yeast chromatin-associated interactome. *Mol. Syst. Biol.* **2010**, *6*, 448.
(43) Lambert, J. P.; Mitchell, L.; Rudner, A.; Baetz, K.; Figeys, D. A novel proteomics approach for the discovery of chromatin-associated protein networks. *Mol. Cell. Proteomics* **2009**, *8* (4), 870−82.
(44) Skarra, D. V.; Goudreault, M.; Choi, H.; Mullin, M.; Nesvizhskii, A. I.; Gingras, A. C.; Honkanen, R. E. Label-free quantitative proteomics and SAINT analysis enable interactome mapping for the human Ser/Thr protein phosphatase 5. *Proteomics* **2011**, *11* (8), 1508−16.

(45) van den Berg, D. L.; Snoek, T.; Mullin, N. P.; Yates, A.; Bezstarosti, K.; Demmers, J.; Chambers, I.; Poot, R. A. An Oct4-centered protein interaction network in embryonic stem cells. *Cell Stem Cell* **2010**, *6* (4), 369−81.
(46) Wang, T.; Gu, S.; Ronni, T.; Du, Y. C.; Chen, X. In vivo dual-tagging proteomic approach in studying signaling pathways in immune response. *J. Proteome Res.* **2005**, *4* (3), 941−9.
(47) Wilm, M. Quantitative proteomics in biological research. *Proteomics* **2009**, *9* (20), 4590−605.
(48) Washburn, M. P. Driving biochemical discovery with quantitative proteomics. *Trends Biochem. Sci.* **2011**, *36* (3), 170−7.
(49) Cox, J.; Mann, M. Quantitative, high-resolution proteomics for data-driven systems biology. *Annu. Rev. Biochem.* **2011**, *80*, 273−99.
(50) Ranish, J. A.; Yi, E. C.; Leslie, D. M.; Purvine, S. O.; Goodlett, D. R.; Eng, J.; Aebersold, R. The study of macromolecular complexes by quantitative proteomics. *Nat. Genet.* **2003**, *33* (3), 349−55.
(51) Brand, M.; Ranish, J. A.; Kummer, N. T.; Hamilton, J.; Igarashi, K.; Francastel, C.; Chi, T. H.; Crabtree, G. R.; Aebersold, R.; Groudine, M. Dynamic changes in transcription factor complexes during erythroid differentiation revealed by quantitative proteomics. *Nat. Struct. Mol. Biol.* **2004**, *11* (1), 73−80.
(52) Jasavala, R.; Martinez, H.; Thumar, J.; Andaya, A.; Gingras, A. C.; Eng, J. K.; Aebersold, R.; Han, D. K.; Wright, M. E. Identification of putative androgen receptor interaction protein modules: cytoskeleton and endosomes modulate androgen receptor signaling in prostate cancer cells. *Mol. Cell. Proteomics* **2007**, *6* (2), 252−71.
(53) Pflieger, D.; Junger, M. A.; Muller, M.; Rinner, O.; Lee, H.; Gehrig, P. M.; Gstaiger, M.; Aebersold, R. Quantitative proteomic analysis of protein complexes: concurrent identification of interactors and their state of phosphorylation. *Mol. Cell. Proteomics* **2008**, *7* (2), 326−46.
(54) Karp, N. A.; Huber, W.; Sadowski, P. G.; Charles, P. D.; Hester, S. V.; Lilley, K. S. Addressing accuracy and precision issues in iTRAQ quantitation. *Mol. Cell. Proteomics* **2010**, *9* (9), 1885−97.
(55) Boersema, P. J.; Raijmakers, R.; Lemeer, S.; Mohammed, S.; Heck, A. J. Multiplex peptide stable isotope dimethyl labeling for quantitative proteomics. *Nat. Protoc.* **2009**, *4* (4), 484−94.
(56) Dubois, F.; Vandermoere, F.; Gernez, A.; Murphy, J.; Toth, R.; Chen, S.; Geraghty, K. M.; Morrice, N. A.; MacKintosh, C. Differential 14-3-3 affinity capture reveals new downstream targets of phosphatidylinositol 3-kinase signaling. *Mol. Cell. Proteomics* **2009**, *8* (11), 2487−99.
(57) Blagoev, B.; Kratchmarova, I.; Ong, S. E.; Nielsen, M.; Foster, L. J.; Mann, M. A proteomics strategy to elucidate functional protein-protein interactions applied to EGF signaling. *Nat. Biotechnol.* **2003**, *21* (3), 315−8.
(58) Schulze, W. X.; Deng, L.; Mann, M. Phosphotyrosine interactome of the ErbB-receptor kinase family. *Mol. Syst. Biol.* **2005**, *1* (2005), 0008.
(59) He, Y. F.; Bao, H. M.; Xiao, X. F.; Zuo, S.; Du, R. Y.; Tang, S. W.; Yang, P. Y.; Chen, X. Biotin tagging coupled with amino acid-coded mass tagging for efficient and precise screening of interaction proteome in mammalian cells. *Proteomics* **2009**, *9* (24), 5414−24.
(60) de Boer, E.; Rodriguez, P.; Bonte, E.; Krijgsveld, J.; Katsantoni, E.; Heck, A.; Grosveld, F.; Strouboulis, J. Efficient biotinylation and single-step purification of tagged transcription factors in mammalian cells and transgenic mice. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100* (13), 7480−5.
(61) Rodriguez, P.; Braun, H.; Kolodziej, K. E.; de Boer, E.; Campbell, J.; Bonte, E.; Grosveld, F.; Philipsen, S.; Strouboulis, J. Isolation of transcription factor complexes by in vivo biotinylation tagging and direct binding to streptavidin beads. *Methods Mol. Biol.* **2006**, *338*, 305−23.
(62) Hubner, N. C.; Bird, A. W.; Cox, J.; Splettstoesser, B.; Bandilla, P.; Poser, I.; Hyman, A.; Mann, M. Quantitative proteomics combined with BAC TransgeneOmics reveals in vivo protein interactions. *J. Cell Biol.* **2010**, *189* (4), 739−54.
(63) Cox, J.; Mann, M. MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat. Biotechnol.* **2008**, *26* (12), 1367−72.
(64) Vermeulen, M.; Eberl, H. C.; Matarese, F.; Marks, H.; Denissov, S.; Butter, F.; Lee, K. K.; Olsen, J. V.; Hyman, A. A.; Stunnenberg, H. G.; Mann, M. Quantitative interaction proteomics and genome-wide profiling of epigenetic histone marks and their readers. *Cell* **2010**, *142* (6), 967−80.
(65) Tomecki, R.; Kristiansen, M. S.; Lykke-Andersen, S.; Chlebowski, A.; Larsen, K. M.; Szczesny, R. J.; Drazkowska, K.; Pastula, A.; Andersen, J. S.; Stepien, P. P.; Dziembowski, A.; Jensen, T. H. The human core exosome interacts with differentially localized processive RNases: hDIS3 and hDIS3L. *Embo J.* **2010**, *29* (14), 2342−57.
(66) Lubas, M.; Christensen, M. S.; Kristiansen, M. S.; Domanski, M.; Falkenby, L. G.; Lykke-Andersen, S.; Andersen, J. S.; Dziembowski, A.; Jensen, T. H. Interaction profiling identifies the human nuclear exosome targeting complex. *Mol. Cell* **2011**, *43* (4), 624−37.
(67) Pagliuca, F. W.; Collins, M. O.; Lichawska, A.; Zegerman, P.; Choudhary, J. S.; Pines, J. Quantitative proteomics reveals the basis for the biochemical specificity of the cell-cycle machinery. *Mol. Cell* **2011**, *43* (3), 406−17.
(68) Gunaratne, J.; Goh, M. X.; Swa, H. L.; Lee, F. Y.; Sanford, E.; Wong, L. M.; Hogue, K. A.; Blackstock, W. P.; Okumura, K. Protein interactions of phosphatase and tensin homologue (PTEN) and its cancer-associated G20E mutant compared by using stable isotope labeling by amino acids in cell culture-based parallel affinity purification. *J. Biol. Chem.* **2011**, *286* (20), 18093−103.
(69) Breitkreutz, A.; Choi, H.; Sharom, J. R.; Boucher, L.; Neduva, V.; Larsen, B.; Lin, Z. Y.; Breitkreutz, B. J.; Stark, C.; Liu, G.; Ahn, J.; Dewar-Darch, D.; Reguly, T.; Tang, X.; Almeida, R.; Qin, Z. S.; Pawson, T.; Gingras, A. C.; Nesvizhskii, A. I.; Tyers, M. A global protein kinase and phosphatase interaction network in yeast. *Science* **2010**, *328* (5981), 1043−6.
(70) Ohta, S.; Bukowski-Wills, J. C.; Sanchez-Pulido, L.; Alves Fde, L.; Wood, L.; Chen, Z. A.; Platani, M.; Fischer, L.; Hudson, D. F.; Ponting, C. P.; Fukagawa, T.; Earnshaw, W. C.; Rappsilber, J. The protein composition of mitotic chromosomes determined using multiclassifier combinatorial proteomics. *Cell* **2010**, *142* (5), 810−21.
(71) Tackett, A. J.; DeGrasse, J. A.; Sekedat, M. D.; Oeffinger, M.; Rout, M. P.; Chait, B. T. I-DIRT, a general method for distinguishing between specific and nonspecific protein interactions. *J. Proteome Res.* **2005**, *4* (5), 1752−6.
(72) Fang, L.; Wang, X.; Yamoah, K.; Chen, P. L.; Pan, Z. Q.; Huang, L. Characterization of the human COP9 signalosome complex using affinity purification and mass spectrometry. *J. Proteome Res.* **2008**, *7* (11), 4914−25.
(73) Mousson, F.; Kolkman, A.; Pijnappel, W. W.; Timmers, H. T.; Heck, A. J. Quantitative proteomics reveals regulation of dynamic components within TATA-binding protein (TBP) transcription complexes. *Mol. Cell. Proteomics* **2008**, *7* (5), 845−52.
(74) Wang, X.; Huang, L. Identifying dynamic interactors of protein complexes by quantitative mass spectrometry. *Mol. Cell. Proteomics* **2008**, *7* (1), 46−57.
(75) Liu, H.; Sadygov, R. G.; Yates, J. R. 3rd A model for random sampling and estimation of relative protein abundance in shotgun proteomics. *Anal. Chem.* **2004**, *76* (14), 4193−201.
(76) Zybailov, B.; Coleman, M. K.; Florens, L.; Washburn, M. P. Correlation of relative abundance ratios derived from peptide ion chromatograms and spectrum counting for quantitative proteomic analysis using stable isotope labeling. *Anal. Chem.* **2005**, *77* (19), 6218−24.
(77) Mosley, A. L.; Florens, L.; Wen, Z.; Washburn, M. P. A label free quantitative proteomic analysis of the Saccharomyces cerevisiae nucleus. *J. Proteomics* **2009**, *72* (1), 110−20.
(78) Florens, L.; Carozza, M. J.; Swanson, S. K.; Fournier, M.; Coleman, M. K.; Workman, J. L.; Washburn, M. P. Analyzing

chromatin remodeling complexes using shotgun proteomics and normalized spectral abundance factors. *Methods* **2006**, *40* (4), 303−11.
(79) Rinner, O.; Mueller, L. N.; Hubalek, M.; Muller, M.; Gstaiger, M.; Aebersold, R. An integrated mass spectrometric and computational framework for the analysis of protein interaction networks. *Nat. Biotechnol.* **2007**, *25* (3), 345−52.
(80) Kuhner, S.; Gavin, A. C. Towards quantitative analysis of proteome dynamics. *Nat. Biotechnol.* **2007**, *25* (3), 298−300.
(81) Hubner, N. C.; Mann, M. Extracting gene function from protein-protein interactions using Quantitative BAC InteraCtomics (QUBIC). *Methods* **2011**, *53* (4), 453−9.
(82) Paoletti, A. C.; Parmely, T. J.; Tomomori-Sato, C.; Sato, S.; Zhu, D.; Conaway, R. C.; Conaway, J. W.; Florens, L.; Washburn, M. P. Quantitative proteomic analysis of distinct mammalian Mediator complexes using normalized spectral abundance factors. *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103* (50), 18928−33.
(83) Zybailov, B.; Mosley, A. L.; Sardiu, M. E.; Coleman, M. K.; Florens, L.; Washburn, M. P. Statistical analysis of membrane proteome expression changes in Saccharomyces cerevisiae. *J. Proteome Res.* **2006**, *5* (9), 2339−47.
(84) Prochasson, P.; Florens, L.; Swanson, S. K.; Washburn, M. P.; Workman, J. L. The HIR corepressor complex binds to nucleosomes generating a distinct protein/DNA complex resistant to remodeling by SWI/SNF. *Genes Dev.* **2005**, *19* (21), 2534−9.
(85) Schneider, J.; Wood, A.; Lee, J. S.; Schuster, R.; Dueker, J.; Maguire, C.; Swanson, S. K.; Florens, L.; Washburn, M. P.; Shilatifard, A. Molecular regulation of histone H3 trimethylation by COMPASS and the regulation of gene expression. *Mol. Cell* **2005**, *19* (6), 849−56.
(86) Sardiu, M. E.; Cai, Y.; Jin, J.; Swanson, S. K.; Conaway, R. C.; Conaway, J. W.; Florens, L.; Washburn, M. P. Probabilistic assembly of human protein interaction networks from label-free quantitative proteomics. *Proc. Natl. Acad. Sci. U.S.A.* **2008**, *105* (5), 1454−9.
(87) Mosley, A. L.; Sardiu, M. E.; Pattenden, S. G.; Workman, J. L.; Florens, L.; Washburn, M. P. Highly reproducible label free quantitative proteomic analysis of RNA polymerase complexes. *Mol. Cell. Proteomics* **2011**, *10* (2), M110 000687.
(88) Guruharsha, K. G.; Rual, J. F.; Zhai, B.; Mintseris, J.; Vaidya, P.; Vaidya, N.; Beekman, C.; Wong, C.; Rhee, D. Y.; Cenaj, O.; McKillip, E.; Shah, S.; Stapleton, M.; Wan, K. H.; Yu, C.; Parsa, B.; Carlson, J. W.; Chen, X.; Kapadia, B.; Vijayraghavan, K.; Gygi, S. P.; Celniker, S. E.; Obar, R. A.; Artavanis-Tsakonas, S. A Protein Complex Network of Drosophila melanogaster. *Cell* **2011**, *147* (3), 690−703.
(89) Behrends, C.; Sowa, M. E.; Gygi, S. P.; Harper, J. W. Network organization of the human autophagy system. *Nature* **2010**, *466* (7302), 68−76.
(90) Malovannaya, A.; Li, Y.; Bulynko, Y.; Jung, S. Y.; Wang, Y.; Lanz, R. B.; O'Malley, B. W.; Qin, J. Streamlined analysis schema for high-throughput identification of endogenous protein complexes. *Proc. Natl. Acad. Sci. U.S.A.* **2010**, *107* (6), 2431−6.
(91) Malovannaya, A.; Lanz, R. B.; Jung, S. Y.; Bulynko, Y.; Le, N. T.; Chan, D. W.; Ding, C.; Shi, Y.; Yucer, N.; Krenciute, G.; Kim, B. J.; Li, C.; Chen, R.; Li, W.; Wang, Y.; O'Malley, B. W.; Qin, J. Analysis of the human endogenous coregulator complexome. *Cell* **2011**, *145* (5), 787−99.
(92) Griffin, N. M.; Yu, J.; Long, F.; Oh, P.; Shore, S.; Li, Y.; Koziol, J. A.; Schnitzer, J. E. Label-free, normalized quantification of complex mass spectrometry data for proteomic analysis. *Nat. Biotechnol.* **2010**, *28* (1), 83−9.
(93) Paul, F. E.; Hosp, F.; Selbach, M. Analyzing protein-protein interactions by quantitative mass spectrometry. *Methods* **2011**, *54* (4), 387−95.