Title

Bait Correlation Improves Interactor Identification by Tandem Mass Tag-Affinity Purification-Mass Spectrometry

Journal

Journal of Proteome Research, 19(4)

ISSN

1535-3893

Authors

Mei, Liangyong
Montoya, Maureen R
Quanrud, Guy M
et al.

Publication Date

2020-04-03

DOI

10.1021/acs.jproteome.9b00825

Bait correlation improves interactor identification by Tandem Mass Tag-Affinity Purification-Mass Spectrometry

Liangyong Mei[‡,†,a] *, Maureen R. Montoya*[‡,a]*, Guy M. Quanrud*[a]*, Minh Tran*[a]*, Athena Villa-Sharma*[a]*, Ming Huang*[b]*, and Joseph C. Genereux*[a,b]*

[a]Department of Chemistry, University of California, Riverside, CA 92521

[b]Environmental Toxicology Graduate Program, University of California, Riverside, CA 92521

*To whom correspondence should be addressed: Joseph C. Genereux, 501 Big Springs Rd, Riverside, CA 92507. josephg@ucr.edu, 951-827-3759.

ABSTRACT

The quantitative multiplexing capacity of isobaric Tandem Mass Tags (TMT) has increased the throughput of affinity purification mass spectrometry (AP-MS) to characterize protein interaction networks of immunoprecipitated bait proteins. However, variable bait levels between replicates can convolute interactor identification. We compared the Student's t-test and Pearson's R correlation as methods to generate t-statistics and assessed the significance of interactors following TMT-AP-MS. Using a simple linear model of protein recovery in immunoprecipitates to simulate reporter ion ratio distributions, we found that correlation-derived t-statistics protect against bait variance while robustly controlling Type I errors (false positives). We experimentally determined the performance of these two approaches for determining t-statistics under two experimental conditions: irreversible prey association to the Hsp40 mutant DNAJB8^{H31Q} followed by stringent washing, and reversible association to 14-3-3ζ with gentle washing. Correlation-derived t-statistics performed at least as well as Student's t-statistics for each sample, and with substantial improvement in performance for experiments with high bait level variance. Deliberately varying bait levels over a large range fails to improve selectivity but does increase robustness between runs. The use of correlation-derived t-statistics should improve identification of interactors using TMT-AP-MS. Data are available via ProteomeXchange with identifier PXD016613.

Key Words: isobaric tags; tandem mass tags; quantitative proteomics; protein-protein interactions; affinity purification – mass spectrometry; DNAJB8; 14-3-3ζ

INTRODUCTION

Protein-protein interactions (PPIs) allow the assembly of protein complexes, mediate signaling pathways, and define the client distribution of enzymes involved in post-translational modifications[1–5]. Identifying and characterizing PPIs provides valuable molecular insights into cell function and physiology. However, the large dynamic range of PPI strength and abundance challenges the discrimination of bona fide from artefactual interactions. This challenge is central to the interpretation of most PPI datasets, including those acquired by the most prevalent techniques: yeast two-hybrid [6–8], sandwich assays such as LUMIER[39,10], microarrays[11], chemical crosslinking[12,13], and Affinity Purification coupled with Mass Spectrometry (AP-MS)[14–19]. In the latter approach, a "bait" protein is isolated, and its co-isolating "prey" interactors are identified by MS. Because many proteins associate with beads independent of bait, this approach is prone to false positives (Type I errors). To decrease false positives, more stringent wash buffers can be used, but this in turn increases false negatives (Type II errors). Alternatively, false positives can be minimized by carefully filtering out potential prey that are known to strongly associate with beads, or by comparison to controls obtained under similar conditions[20,21]. Large datasets comprised of dozens to thousands of AP-MS runs reveal the identities and intensities of non-specific binding proteins associated with an individual AP-MS platform, which can then be compared to each individual AP-MS experiment[21–23]. This was most dramatically demonstrated in the BioPlex, which determined over 70,000 interactions from over 5,000 proteins (and counting) using the Comparative Proteomics Analysis (ComPASS) methodology[24,25]. For smaller datasets, however, while aggressive statistical filters ensure that only high-quality interactors are reported, they can also lead to few identified prey [26,27].

TMT-AP-MS allows direct quantitative comparisons of prey recovery across multiple replicates in a single run, simplifying evaluation of whether potential interactors are preferentially recovered with the bait[22,28]. However, variation in bait levels between conditions, particularly when using transient transfection, leads to variability in recovered prey levels. For experiments comparing interaction networks for a single bait between cellular conditions, this variability has been controlled by normalizing prey levels to bait levels[28,29]; this approach is not available when differentiating real vs. artefactual interactors, as minimal bait levels are present under mock

transfection conditions. Several studies have been successful at identifying protein complex composition by global correlation analysis following native protein chromatography. In this approach, proteins are chromatographically separated under conditions that preserve native complexes, with the expectation that proteins that co-fractionate are likely to interact. A similar approach is to perform a large number of immunoprecipitations, and then to determine which proteins co-IP reproducibly; in this case correlations are made between each prey:prey combination rather than solely between prey and bait. [30–33]. This approach relies upon the tendency of interacting proteins to maintain stoichiometry. However, both of these approaches require large scale experiments and probe global interaction maps, while researchers frequently want to identify interactors of a specific bait protein with the fewest number of replicates. We thus consider whether using a correlation-based analysis can assist in identifying prey following individual TMT-AP-MS runs for a single bait. Herein, we report that evaluating potential interactors on the basis of their Pearson's correlation with bait levels decreases Type II errors without increasing Type I errors.

EXPERIMENTAL

Reagents. Buffer components and other biochemical reagents were all purchased from Fisher, VWR, or Millipore Sigma. Nanopure water and sterilized consumables were used for all biochemical experiments.

Simulations. Calculations were performed in Mathematica (See **Supporting Methods**). TMT runs were simulated by populating reporter ion ratios for three types of protein: bait, non-specific interactors (ni), and specific interactors (prey). Background signals for each protein, bait levels in an individual replicate, the ratio of each prey to the bait, and replicate-level error for all measurements were drawn from truncated normal distributions. For each parameter set, 100 TMT 6-plex samples were populated with 3 channels including "bait" and 3 channels serving as "mock"s, with no bait expressed. Area under the curve was calculated by integration of a 100-point parametrization of the Receiver Operating Characteristic curve, with prey taken as true-positives and non-specific interactors taken as false positives. Mathematica code may be found at https://github.com/josephgenereux/ROC-Simulations-in-Mei-et-al.

Molecular Cloning. DNAJB8 was amplified from pcDNA5/FRT/TO V5 DNAJB8[34] and inserted into the pFLAG.CMV2 vector by PIPE cloning[35]. The H31Q mutation was introduced into DNAJB8 using site-directed mutagenesis. 14-3-3ζ was amplified from cDNA derived from HepG2 (ATCC) using TRIzol (Fisher), and inserted into the pFLAG.CMV2 vector by PIPE cloning. eGFP.pDEST30 has been reported[36]. All constructs were subjected to analytical digest and sequenced (Retrogen) to confirm identity. Primers were purchased from IDT, and all cloning enzymes purchased from New England Biolabs. Primer sequences are available in **Table S1**.

Human Tissue Culture. HEK293T cells (ATCC) were cultured in DMEM (Corning) supplemented with 10% fetal bovine serum (FBS; Seradigm), 2 mM L-Glutamine (Corning), and penicillin (100 IU/mL)-streptomycin (100 μg/mL; Corning), and used within 30 passages. Cells were checked monthly for mycoplasma contamination by PCR assay. Plasmid DNA was introduced into cells by the method of calcium phosphate transfection. Transfection efficiency >80% was confirmed in all experiments by GFP positive control.

Immunoprecipitation. Cells were harvested from confluent 10 cm dishes at 36 to 48 h post-transfection by scraping in TBS buffer with 1 mM EDTA. High stringency lysis was performed in RIPA buffer (150 mM NaCl, 50 mM Tris pH 7.5, 1% Triton X100, 0.5% sodium deoxycholate, 0.1 % SDS). Low stringency lysis and washes was performed with 0.1% Triton X-100 in TBS (10 mM Tris pH 7.5, 150 mM NaCl). For conditions involving dithiobis succinimidyl propionate (DSP) crosslinking, cells were incubated in 1 mM DSP/1%DMSO/PBS for 30 min. with rotating at ambient temperature, and then quenched by addition of Tris pH 8.0 (to 90 mM) and rotating for 15 min. at ambient temperature. Cells were lysed for 30 min on ice in lysis buffer supplemented with fresh protease inhibitors (Roche). Lysate was separated from cell debris centrifugation at 21,100 x g for 15 min at 4°C. Protein was quantified by Bradford assay (Bio-rad). Lysates were pre-cleared with 15 uL sepharose-4B beads (Millipore Sigma) for 30 min at 4°C, then centrifuged at 1,500 × g for 1 min to pellet beads, denatured by boiling for 10 min at 100 °C in 20% SDS, followed by immunoprecipitation (at 0.1% SDS) with 15 μL M2 anti-FLAG Magnetic Beads (Millipore Sigma) and overnight rotation at 4 °C. The denaturation step was excluded for low-stringency immunopurifications. Immunodepletion was validated by immunoblot. Beads were washed the next day four times with lysis buffer, 10 min each wash rotating at ambient temperature. Bound proteins were eluted

from the beads by boiling for 5 min at 100°C in 30 uL of Laemmli concentrate (120 mM Tris pH 6.8, 60% glycerol, 12% SDS, brilliant phenol blue to color). About 17% of eluates were reserved for silver stain analysis, while the remainder was prepared for mass spectrometry.

Silver Stain. Eluates were boiled for 5 min at 100°C with 0.17 M DTT and separated by SDS-PAGE. Gels were fixed overnight in 30% ethanol/10% acetic acid or for a few hours with 50% methanol/12% acetic acid. Gels were washed in 35% ethanol three times for 20 min each, sensitized for 2 min (0.02% $Na_2S_2O_3$ in H_2O), washed three times for 1 min each in H_2O, and stained for 30 min to overnight in Ag staining solution (0.2% $AgNO_3$, 0.076% formalin). After two 5 min rinses in H_2O, gels were developed with 6% $NaCO_3$/0.05% formalin/0.0004% $Na_2S_2O_3$. Development was stopped with 5% acetic acid. Gels were imaged on a white-light transilluminator (UVP).

TMT-MuDPIT. Immunoprecipitates were prepared for TMT-AP-MS according to standard protocols[28,37]. Labeled and pooled digests were fractionated by SCX in line with a reversed-phase analytical column to enable two-dimensional separation prior to electrospray ionization. Peptides were analyzed using an LTQ Orbitrap Velos Pro in data-dependent mode. The top ten peaks from each full precursor scan were fragmented by HCD to acquire fragmentation spectra. Peptide-spectra matches were evaluated by ProLuCID[38,39] using a Uniprot proteome database supplemented with common contaminants and a full decoy set, and filtered (DTA Select version 2.0.27[40]) to 1% false discovery rate for peptide identifications. TMT reporter ion ratios were quantified in Census[41], and only unique peptides were considered. Full conditions and parameters are provided in **Supporting Experimental Methods**. The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE[42] partner repository with the dataset identifier PXD016613 and 10.6019/PXD016613.

Parallel Reaction Monitoring. Human DNAJB8 peptides appropriate for PRM (not including H31) were chosen using the Picky online interface[43,44] (**Table S2**). Lysates were prepared for PRM according to standard protocols[28,37]. Digests (20 μg) were injected onto a homemade C18 trapping column for desalting and then separated on a reversed-phase analytical column prior to electrospray ionization. Peptides were analyzed according to a scheduled isolation method using an LTQ Orbitrap Velos Pro. Chromatograms were integrated in

Xcalibur. An internal reference peptide[45] (VFFAEDVGSNK) was spiked into each sample and used for normalization. Full conditions and parameters are provided in **Supporting Experimental Methods**.

Statistical Methods. Ordinary t-statistics for each protein were determined from $t = (\bar{x} - \bar{y})\sqrt{\frac{n}{s_x^2 + s_y^2}}$ where $\bar{x}$ and $\bar{y}$ are the sample means for TMT reporter ion intensities, s_x and s_y are the respective sample standard deviations, and n is the number of measurements. p-values were then inferred based on the two-tailed Student's t-distribution. Raw TMT intensities ratios were analyzed with the LIMMA package in R to generate moderated t-statistics and p-values[46,47]. This data package transforms TMT reporter ion intensities to a logarithmic scale due to the assumption that fold changes are constant between conditions[48], and then moderates the standard errors for each protein against a global estimated standard error. This moderated standard error is then used to generate a t-statistic and p-values in the standard manner.

Because the Pearson's coefficient of non-correlated data (the null hypothesis) is normally distributed[49], a simple t-statistic can be directly calculated from: $t = R\sqrt{\frac{n-2}{1-R^2}}$ (for derivation see ref [50]), where $n - 2$ is the degrees of freedom and R is the correlation coefficient, $r = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2}\sqrt{\sum_{i=1}^{n}(y_i - \bar{y})^2}}$. This t-statistic is equivalent to the ratio of the measured slope from the linear fit divided by its standard error. p-values are then inferred by comparing the t-statistic to the two-tailed Student's t-distribution with $n - 2$ degrees of freedom.

q-values (q_{BH}) were determined from p-values using the Benjamini-Hochberg methodology, $q_k = p_k \frac{N}{k}$, where k is the rank for each protein, arranged as increasing p-value, and N is the total number of proteins examined (representing the number of hypotheses)[51]. These q-values are then adjusted to preserve monotonicity by replacing each q_i with $\min\{q_j | j>i\}$, a transformation that has been shown to offer greater power without sacrificing false discovery rate control[52]. q-values were originally developed to allow construction of a set $\{k \mid q_k < Q\}$ such that the False Discovery Rate of {k} is below Q. However, individual q-values still provide a measure of the local false discovery rate even in the absence of a set value of Q [53].

Box and whisker plots are presented with lines marking median values, X marking average values, boxes from the first to third quartiles, whiskers extending to minimum and maximum values (excluding outliers), and outliers defined at points greater than 1.5-fold the interquartile range beyond the first and third quartiles. These outliers are shown explicitly. Coefficients of Variation are calculated as the standard deviation divided by the mean.

RESULTS AND DISCUSSION

We constructed a simple model (See **Supporting Methods**) for the recovery of prey proteins (defined as proteins whose specific recovery is linearly dependent on the bait) and non-interacting proteins (those whose recovery is independent of the bait). Simulations indicate that both Pearson's and Student's t-statistics perform well when bait level variation is low (**Figure 1**). However, as bait level variance between replicates increases, Student's derived t-statistics are less able to distinguish prey from non-specific interactors, as compared to Pearson's correlation-derived t-statistics. This finding was robust against a wide range of parameters sets (**Figure S1**). Hence, we considered that perhaps Pearson-derived t-statistics would enable better discrimination of true and false interactors from TMT-AP-MS data.

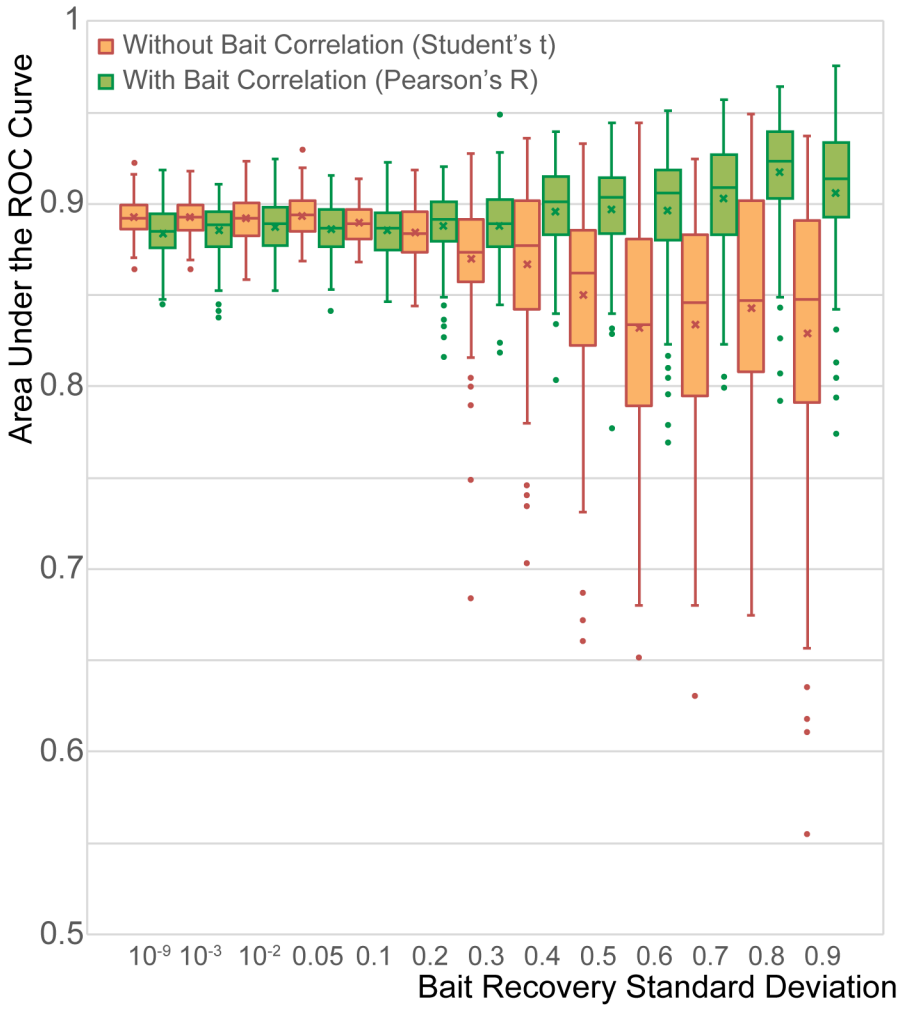


Figure 1. Box and whisker plots of interactor identification accuracy, as determined using t-statistics derived from either Student's t test or Pearson's correlation, as indicated. Areas Under the Curve (AUCs) for Receiver Operating Characteristic (ROC) Curves, which reflect interactor accuracy, were generated from simulations (100 replicates) of 500 non-specific (non-interactor) and 500 specific (prey) proteins recovered from bait immunoprecipitation under conditions of bait level variation. For each simulation, we generate a ROC curve. This curve compares *sensitivity* (true positive rate) to *1 – specificity* (false positive rate), based on assignment of non-interactors and prey before the simulation[54]. The area under this curve represents the accuracy with which non-specific and specific interactors are distinguished. Areas are averaged across the 100 replicate simulations. The abscissa represents the standard deviation of the distribution from which individual bait levels are drawn. Other parameters are provided in **Table S3**. At highly reproducible bait recovery levels (low standard deviation), the

bait correlation approach (green) slightly underperforms a traditional t-statistic (orange), while as the variability of bait recovery increases, t-statistics derived from Pearson's correlations outperform Student's t-statistics.

To evaluate this hypothesis, we initially designed a TMT-AP-MS experiment that minimizes non-specific interactors while providing a large number of true interactors. Hsp40 co-chaperones are responsible for recruiting about a third of the proteome to the Hsp70 chaperoning pathway[55]. These protein clients are then handed off from Hsp40 to Hsp70 to promote folding or degradation. Mutation of the Hsp70-binding motif of Hsp40, however, inhibits client release[56]. We constructed ^{Flag}DNAJB8^{H31Q}, where the H31Q mutation blocks association with Hsp70[57]. We further employed the cell permeable crosslinker DSP to immortalize interactions between prey and ^{Flag}DNAJB8^{H31Q} prior to lysis and immunopurification[58]. Stringent washing with the high-detergent buffer RIPA was employed to minimize non-specific interactions with the beads. To evaluate the fidelity of our bait recovery, elution, digestion, and labeling protocol, we transfected three concentrations of ^{Flag}DNAJB8^{H31Q} DNA into HEK293T cells, and compared ^{Flag}DNAJB8^{H31Q} TMT bait reporter ion ratios against the amount of ^{Flag}DNAJB8^{H31Q} in the lysates as measured by Parallel Reaction Monitoring (PRM). Both measures provide similar ratios. The two lowest levels of ^{Flag}DNAJB8^{H31Q} indicate similar recovery, while recovery decreases for the highest level of transfected ^{Flag}DNAJB8^{H31Q} (**Figure S2a-c**).

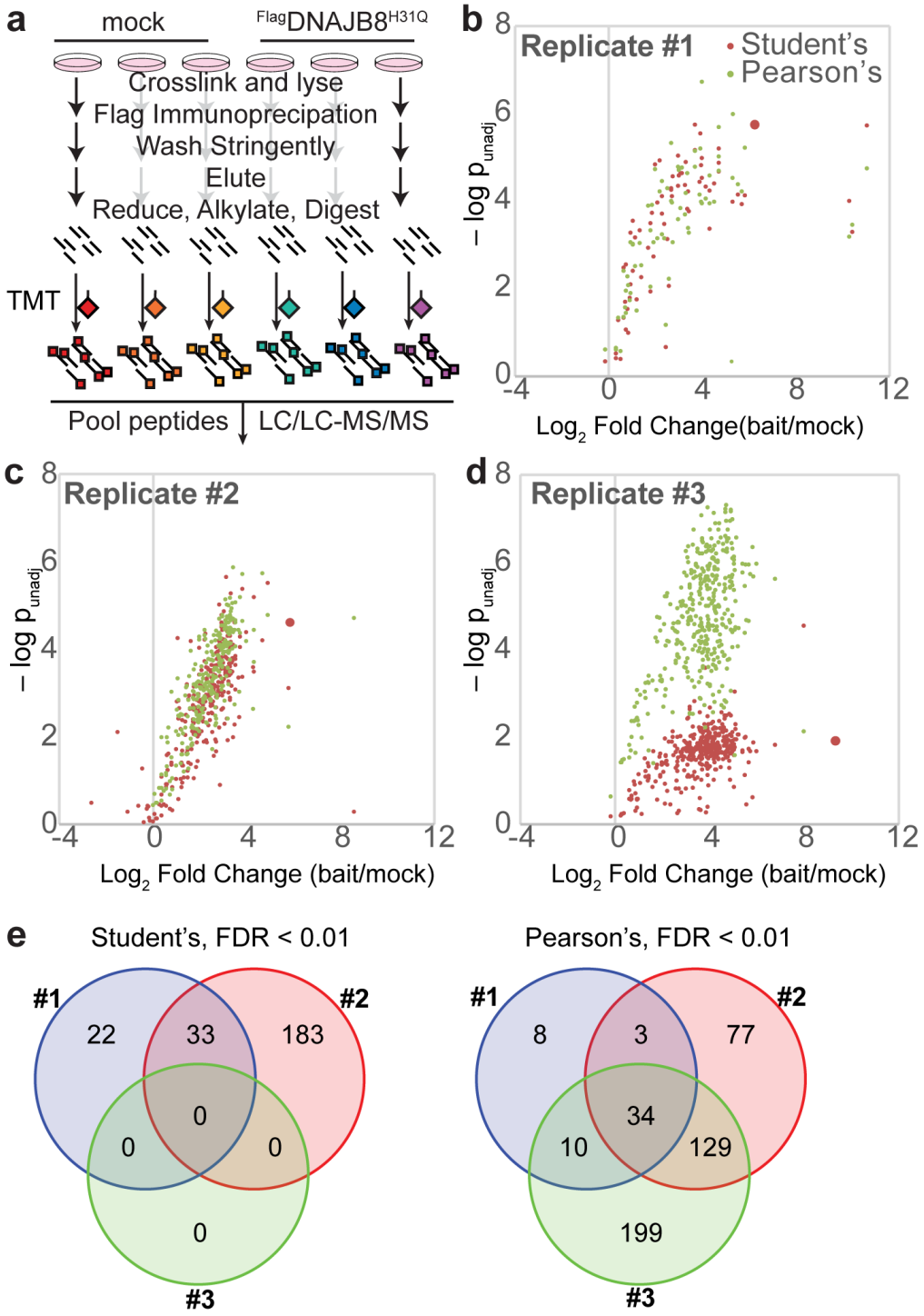


Figure 2. Comparison of p-values obtained through Student's t vs. Pearson's correlation analysis. (**a**) Schematic of TMT-AP-MS to characterize DNAJB8^{H31Q} interactors. Cells were crosslinked in 1 mM DSP for 30 min. and quenched with Tris buffer prior to lysis. (**b-d**) Volcano plots for DNAJB8^{H31Q} interactome replicates collected as described in panel **a**, with unadjusted p-values determined from unpaired, two-way moderated Student's t (top)

and Pearson's correlation (bottom). The large point indicates the bait, using the p-value determined from Student's t analysis. (**e**) Comparison of identified prey (FDR < 0.01) between replicates for the two analysis methods, with FDR determined from the method of Benjamini and Hochberg[51].

We generated three replicate six-plex sets as follows. Nine plates each of HEK293T cells were transfected with either bait (^{Flag}DNAJB8^{H31Q}) or mock (GFP) (**Figure 2a**). After DSP crosslinking, lysing and immunopurification, the eluates were reduced, alkylated, digested, and TMT-labeled. TMT-labeled peptides were pooled in six-plex to yield three replicates that were then analyzed by shotgun proteomics (**Figure 2b-d**, **Figure S2d**, and **Table S4**). Note that here, "replicate" refers to a single run that includes six independent "samples". Student's t-test, followed by Benjamini-Hochberg analysis[51] to determine q-values (q_{BH}) and false discovery rates (FDR), was applied to determine likely DNAJB8^{H31Q} interactors. When the number of samples is low, ordinary t-tests suffer from poor estimation of variance. This estimate is improved by moderating the variance of each individual protein's integrated TMT ion intensity with the global variance[46,59]. Consistent with prior reports[59], we find that moderation slightly decreases the p-value for most proteins, while sharply increasing the p-value for proteins featuring anomalously low variance (**Figure S3a)**. The first two replicates yield several dozen significant interactors of DNAJB8^{H31Q} (using a threshold of FDR < 1%), and the second replicate captures 60% (33 out of 55) of the interactors identified in the first replicate (**Figure 2e**). In the third replicate, however, no significant interactors are identified. This is due to higher bait level variance; the coefficient of variation of the bait in this replicate is 60% as opposed to <15% for the first two replicates. Even the DNAJB8^{H31Q} *bait* is not significantly different between the mock and bait transfection conditions (**Figure 2d**). Rather, in each replicate the p-value for the bait itself (large points in **Figure 2b-d**) defines the approximate lower limit for what p-values are calculated for the various prey. The cause of this variance could be due to any number of factors: variance in efficiencies of transfection, lysis, immunoprecipitation, elution, and digestion. When we use Pearson's correlation to derive t-statistics, we find little change in p-values for Replicates #1 and #2, where bait variability is low (**Figure 2b,c and Figure S3b**). ^{Flag}DNAJB8^{H31Q} interactors in Replicate #3, however, have far lower p-values when derived from Pearson's correlation that from a Student's t-test (**Figure 2d and Figure S3b**). The q_{BH} values

for the high variability sample are similarly decreased when t-statistics are generated from using Pearson's correlation (**Figure S3c**). Now, 92% (34 out of 37) of prey shared between Replicates #1 and #2 appear in the set of Replicate #3 prey with FDR < 1% (**Figure 2e**). There is little change in the overlap between Replicates #1 and #2. Of the 33 proteins that fall below the 1% threshold for both Replicates #1 and #2 using the Student's t-statistic, 32 fall below the threshold for both replicates when using the Pearson's correlation derived t-statistic. This demonstrates that using Pearson's correlation to determine t-statistics from TMT-AP-MS data can account for bait level variation, as predicted by our simulations (**Figure 1**).

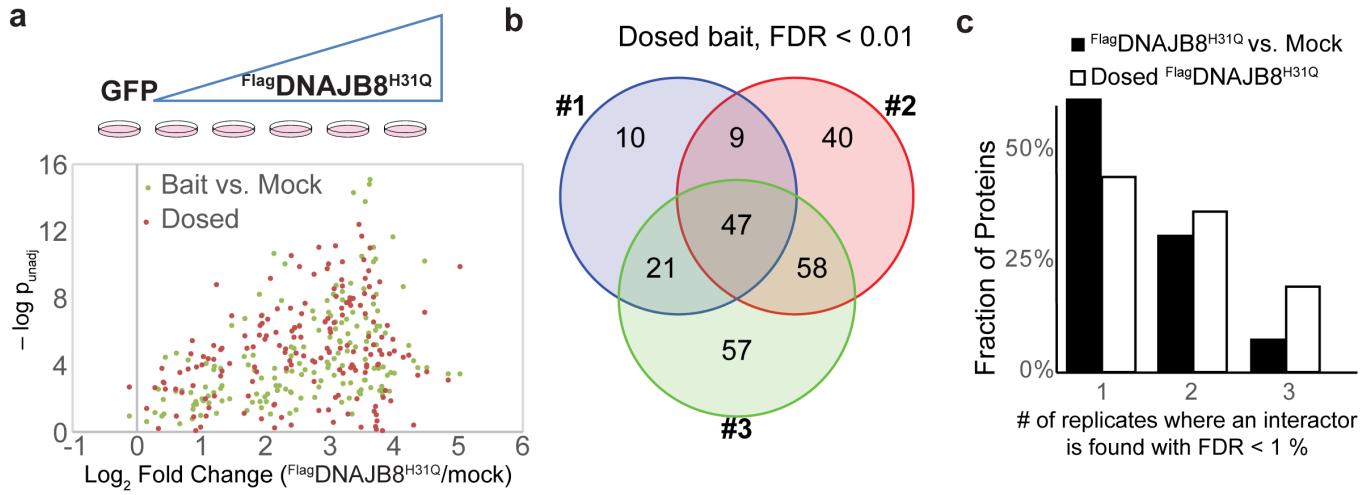


Figure 3. Comparison of p-values obtained through Pearson's correlation analysis from experiments analyzing three bait vs. three mock conditions, or intentional bait dosing for DNAJB8^{H31Q} AP-MS. (**a**) Transfection schematic for the dosed bait approach (upper), and comparison of unadjusted p-values obtained from either the combined bait vs. mock experiments (n = 3) and the dosed bait approach (n = 3). (**b**) Comparison of identified prey (FDR < 0.01) between replicates. (**c**) Histogram of the fraction of identified prey that was found in the given number of replicates.

We hypothesized that deliberately varying the bait level and then determining significance by employing Pearson-derived t-statistics might further improve interactor identification. Although this approach loses the ability to determine "fold change" between mock and bait expression conditions, this fold change reflects non-

specific interactions of a potential prey with beads as much as it reflects specific interactions with bait, and hence is not inherently useful for determining meaningful interactions. To test our hypothesis, we transfected DNA encoding ^{Flag}DNAJB8^{H31Q} over a range of concentrations (0 µg, 2 µg, 4 µg, 6 µg, 8 µg, 10 µg DNA per 10 cm dish) into HEK293T cells, and quantified interactors by TMT-AP-MS for three replicates (**Figure 3a and Figure S4**). For each replicate and at the highest DNA concentration, the measured amount of ^{Flag}DNAJB8^{H31Q} by TMT reporter ion ratios decreases as compared to the amount of ^{Flag}DNAJB8^{H31Q} protein at lower DNA concentrations, suggesting that the cell does not well handle this high amount of ^{Flag}DNAJB8^{H31Q}. Surprisingly, this dosing strategy yields similar but not superior results to the more traditional bait vs. mock approach (**Figure 3a,b**). Nevertheless, a larger fraction of prey was identified as significant in each run for the dosing approach as opposed to the bait vs. mock approach (**Figure 3c**), indicating that dosing might be valuable when robustness across data sets is key.

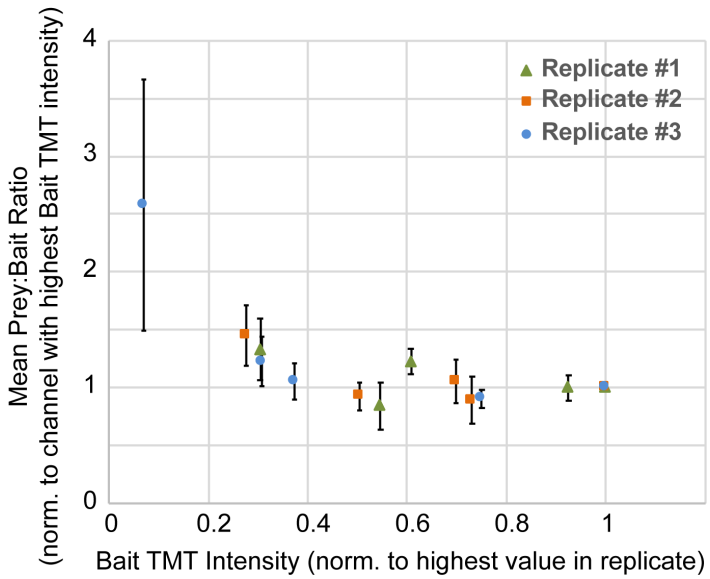


Figure 4. Average (mean) interactor-to-bait TMT reporter ion intensity ratios for all identified interactors (FDR < 0.01) from bait dosing experiments (as in **Figure 3a**), normalized to the channel with the highest levels of DNAJB8^{H31Q} as determined from TMT intensities. The abscissa provides the Bait TMT Intensity for each channel, normalized within replicates to the TMT channel with the highest DNAJB8^{H31Q} intensity. Averages are performed across all interactors, and then across the three replicates. Error bars represent standard deviation across all prey (#1: n = 87; #2: n = 154; #3: n = 184).

In our simulated data (**Figure 1**), accuracy of interactor assignment using Pearson-derived t-statistics increased with higher bait variation, but we did not observe any such improvement in practice, perhaps due to a loss of the assumed linearity between recovered prey and bait levels. Within each replicate, we normalized the interactor:bait ratio for each recovered interactor to the ratio with the highest levels of DNAJB8^{H31Q} by TMT intensity. If an interactor varies linearly with bait levels over this concentration range, then we would expect this ratio to remain constant. Instead, we find that the mean interactor:bait ratio decreases with increasing bait levels, at least for the lower levels of bait (**Figure 4**). This non-linearity is severe enough to suppress bait-interactor correlation, and explains why deliberately dosing bait over a wide range of levels does not improve interactor t-statistics. In molecular terms, it is likely that at higher levels of bait expression, the bait begins to saturate endogenous interactors, thus decreasing the interactor:bait ratio.

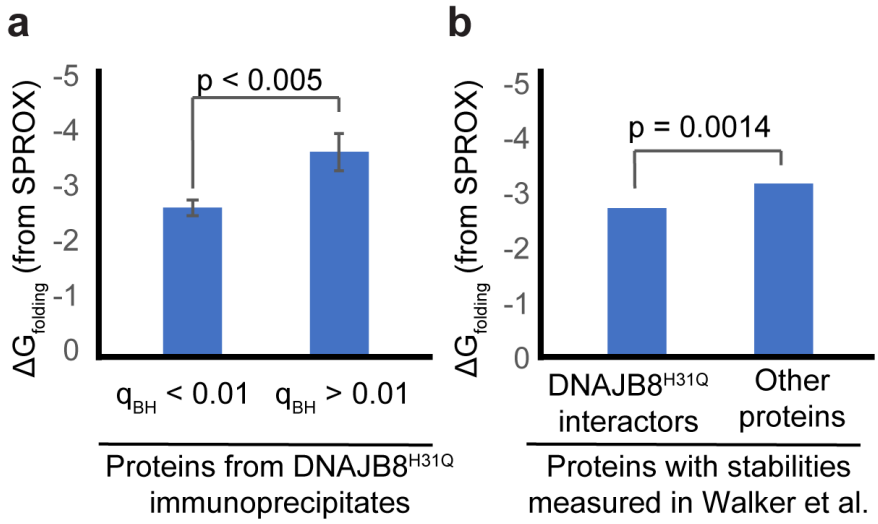


Figure 5. SPROX-derived $\Delta G_{folding}$ for the DNAJB8^{H31Q}-interacting proteome. (**a**) Mean $\Delta G_{folding}$ for likely DNAJB8^{H31Q} -interacting prey compared to the rest of the identified proteome in immunoprecipitates. Likely prey are those with a $q_{BH} < 0.01$. **b** Mean $\Delta G_{folding}$ for likely ($q_{BH} > 0.01$) DNAJB8^{H31Q} -interacting prey as compared to all other proteins with $\Delta G_{folding}$ identified by SPROX. Data were assessed by two-tailed Student's t-test, and error bars represent standard deviations.

A lack of a validated DNAJB8 interactor data set makes it challenging to judge the accuracy of our discovered interactors. However, as an Hsp40 chaperone, DNAJB8 should associate with a relatively destabilized proteome.

To evaluate that possibility, we compared our results to a recently published data set from Walker et al. that directly measured proteome-wide $\Delta G_{folding}$ by Stability of Proteins by Rate of Oxidation (SPROX)[60]. In SPROX, H_2O_2 is added to cell lysate, and the relative oxidation of methionines to the sulfoxide is determined by quantitative shotgun proteomics[61]. Methionine oxidation of a peptide reveals the extent to which that peptide is solvent exposed. If the dependence of methionine oxidation on chaotrope concentration follows a 2-state transition, then $\Delta G_{folding}$ can be inferred. We divided the identified proteins from DNAJB8^{H31Q} immunoprecipitates into likely interactors ($q_{BH} < 0.01$; 476 proteins) and less reliable interactors ($q_{BH} > 0.01$; 159 proteins) on the basis of their Benjamini-Hochberg q-values. Out of these, 163 and 46 proteins respectively had stabilities reported in Walker et al. The likely DNAJB8^{H31Q} interactors are significantly destabilized compared to the less reliable DNAJB8 ^{H31Q} interactors (**Figure 5a**). The converse holds as well; the $\Delta G_{folding}$ for peptides found in our DNAJB8 interactors is significantly less negative than for the peptides not found in our interactors (**Figure 5b**). Hence, the interactors that we observe co-immunoprecipitating with DNAJB8^{H31Q} represent a more destabilized proteome, consistent with DNAJB8^{H31Q}'s functional role as a chaperone for misfolded protein.

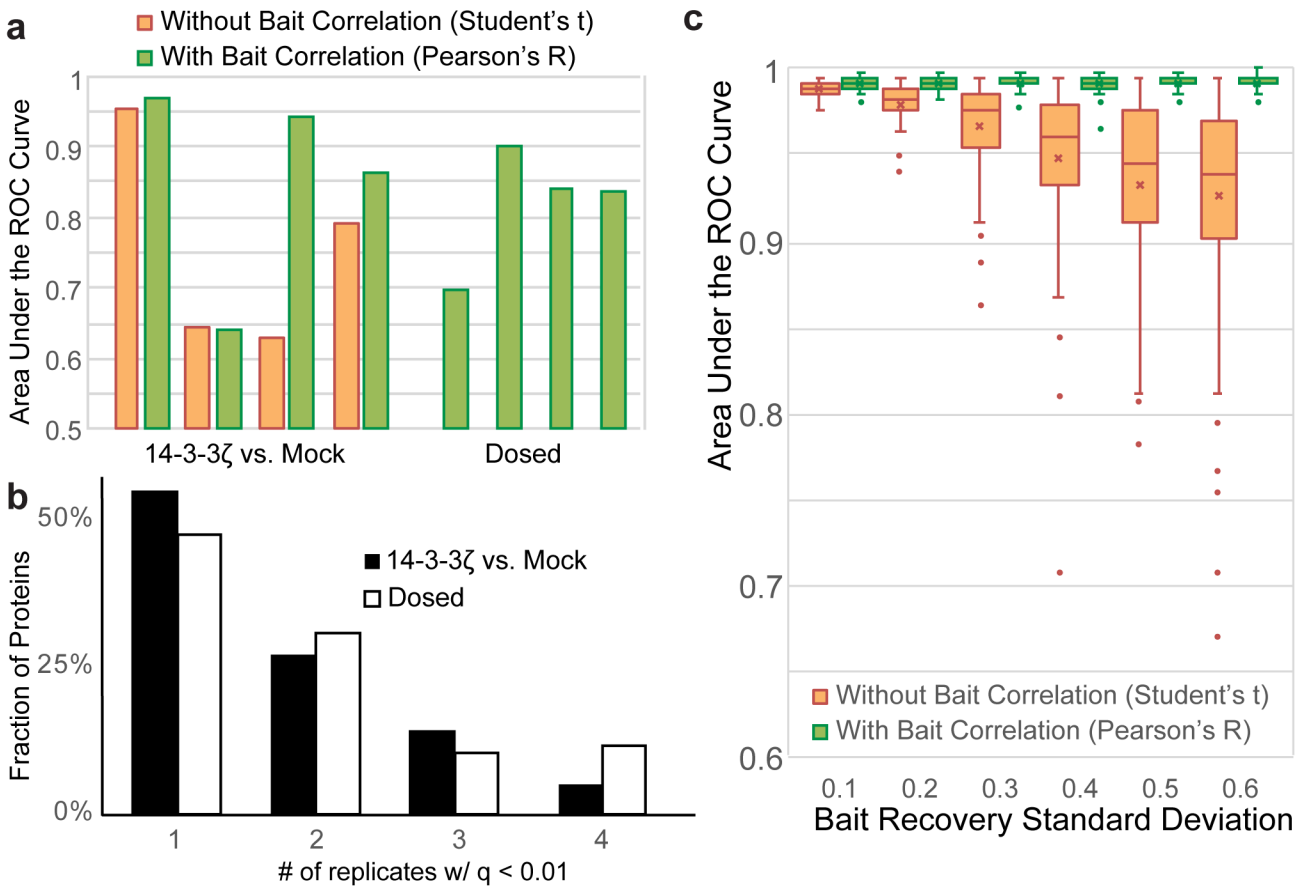


Figure 6. Comparison of Student's and Pearson's-derived t-statistics for differentiating true and false positive interactors of 14-3-3ζ from TMT-AP-MS. **a** Areas under the Curve (AUCs) for Receiver Operating Characteristic Curves for 14-3-3ζ interactors identified by Student's (orange) and Pearson's-derived (green) t-statistics. Four replicates are show for bait vs. mock experiments (as in **Figure 2a**) and bait dosing experiments (as **in Figure 3a**). Higher AUCs indicate higher accuracy at distinguishing true (from BioGrid[62]) and false (from CRAPome[21]) interactors. **b** Histogram of the fraction of identified prey that was found in the given number of replicates. **c** Box and whisker plots of interactor identification accuracy, as determined using t-statistics derived from either Student's t test or Pearson's correlation, as indicated. Areas Under the Curve (AUCs) for Receiver Operating Characteristic Curves, which reflect interactor accuracy, were generated from simulations (100 replicates) of 500 non-specific (non-interactor) and 500 specific (prey) proteins recovered from bait immunoprecipitation under conditions of bait level variation. The abscissa represents the standard deviation of the distribution from which

individual bait levels are drawn. Parameters are drawn from the new data sets generated in this work, as described in **Supporting Methods** and **Table S6**.

We considered whether less stringent AP-MS conditions, wherein a substantial number of identified proteins will be non-specific interactors, would affect the performance of bait-prey correlation in identifying specific prey. Towards this end, we chose a bait protein that reversibly binds its clients, 14-3-3ζ[63]. To increase the number of potential Type I errors, we did not crosslink and only washed with a gentle, low-detergent wash buffer. Twelve plates each of HEK293T cells were transfected with either ^{Flag}14-3-3ζ or mock (GFP) and immunoprecipitated with anti-Flag beads. Eluates were reduced, alkylated, digested, and peptides TMT-labeled. These peptides were pooled to generate four replicates that each contained 3 samples with bait and 3 mock, and each sample was analyzed by MuDPIT LC-MS/MS. Fewer interactors in general are observed in each individual replicate, as compared to the ^{Flag}DNAJB8^{H31Q} pulldowns (**Figure S5a-e** and **Table S5**). There is no global change in interactor p-values between the Pearson and Student approaches, nor is there a change in the number of proteins passing the FDR < 0.01 threshold (**Figure S6a**). Unlike DNAJB8, there are many 14-3-3ζ interaction sets in the literature[64,65]. We generated a true positive interactor list from BioGrid, requiring that proteins be recovered with 14-3-3ζ in at least three different AP-MS studies or at least two AP-IB studies (58 total proteins, or which 19 were found in at least two of our runs)[22,62]. To generate the true negative interactor list, we considered proteins that were reported as significant Dynabead contaminants from total human cell lysate in the Contaminant Repository from Affinity Proteomics (CRAPome)[21], using a filter of appearing in at least 8/24 reported control runs and removing known 14-3-3ζ interactors (983 total proteins, of which 55 are observed in at least 2 of our runs). Most proteins without filtering appear in three or fewer runs, with a steep drop as the threshold is increased. We chose 8/24 as the threshold based on the inflection point in the cumulative distribution function relating proteins observed vs. number of CRAPome runs. Using these sets, we generated Receiver Operator Characteristic Curves for each replicate AP-MS experiment, using either Student's t-statistics, or incorporating bait correlation by using Pearson's-derived t-statistics (**Figure S6b**). While bait correlation had no effect on the Area Under the

Curve for two replicates, it substantially improved differentiation between false and true positives for the other two replicates (**Figure 6a**). Not surprisingly, given our findings with DNAJB8^{H31Q} (**Figure 3**), dosing in variable levels of bait did not improve prey identification as opposed to the bait vs. mock experimental setup. We further considered combining all 14-3-3ζ replicates to identify consistent high-quality interactors (Bonferonni-adjusted p-value < 0.001; 34 proteins) and unlikely interactors (unadjusted p-value > 0.7; 42 proteins). Here, we are using Bonferonni adjustment because it is a highly conservative metric. With this set of interactors/non-interactors, Receiver Operating Curves show larger areas under the curve when evaluated using Pearson's correlation-derived t-statistics as opposed to Student's t-statistics (**Figure S6c,d**). Finally, as with DNAJB8^{H31Q}, the dosing approach yields more reproducible high-confidence interactors between replicates than the more traditional bait vs. mock approach (**Figure 6b**).

The large number of non-interactors in the 14-3-3ζ TMT-AP-MS experiments allow us to estimate the mean and standard deviation of non-interactor TMT reporter ion ratios. Nonspecific levels of bait TMT reporter ion ratios in non-transfected mock samples, and both levels and distributions of prey in mock samples without bait transfection, were similarly determined from the DNAJB8^{H31Q} TMT-AP-MS experimental data (see **Supporting Methods** and **Table S4**). We used this experimental data to test the normality assumption underlying the data in Figure 1, finding that non-interactor mean intensities and prey-bait ratios are both close to a truncated normal distribution (**Figures S7a,b**). Interestingly the mean intensities of DNAJB8^{H31Q} interactors in the bait-free conditions deviate sharply below what would be expected from a truncated normal distribution (**Figures S7c**). High levels of protein in one TMT channel can allow quantification of that protein in other channels, even if the levels in those channels would not normally be adequate for data dependent isolation or quantification[66]. In this case, high levels of interactors in the bait pull-downs could be allowing quantification of negligible levels in the untransfected samples. Using these measured parameters, we revisited the simulations of **Figure 1** to determine whether under realistic parameters we still observe improvement in prey identification by using bait correlation. The use of Pearson's correlation-derived t-statistics continues to outperform Student's t-statistics over a wide range of bait level variances (**Figure 6c**).

CONCLUSIONS

To some extent, all proteins that share a compartment interact inside the crowded environment of the cell. After membrane disruption, even proteins from separate compartments that would never encounter each other within the cell can display high affinities in the lysate. Distinguishing meaningful interactions continues to be challenging. The highest-confidence sets include stringent controls, multimodal characterization, and independent validation. Practically, however, an individual researcher seeking high-value interactor targets for a single bait needs methodology that is simple and reliable. The development of the CRAPome suite enabled reasonable confidences to be inferred for individual label-free experiments by spectral counting and comparison to a mock experiment database, but this approach does not mitigate the low sensitivity of spectral counting. Isobaric quantification provides an accessible approach to quantitatively compare several replicates in a single run. We have demonstrated that variation in transfected bait levels poses a challenge to reliable interactor identification during TMT-AP-MS. Incorporating bait correlation by deriving t-statistics from Pearson's correlation improves the sensitivity and specificity of prey identification. This approach allows small-scale TMT-AP-MS experiments to be "rescued" under conditions of bait variation.

Supporting Information is available free of charge on the ACS Publications website at http://pubs.acs.org:

- Detailed experimental methods, code for simulations, **Tables S1-3**, and **S6** and **Figures S1-7** (pdf)
- **Tables S4** and **S5**, TMT reporter ion intensity ratios for immunoprecipitations (xlsx) are included in the supporting information as xlsx documents.

AUTHOR INFORMATION

*To whom correspondence should be addressed: Joseph C. Genereux, 501 Big Springs Rd, Riverside, CA 92507. josephg@ucr.edu, 951-827-3759.

†Present Address: Department of Chemistry, University of Arizona, Tucson, AZ

‡ These authors contributed equally.

ACKNOWLEDGEMENTS

pcDNA5/FRT/TO V5 DNAJB8 (Addgene # 19531) was generously shared by H. Kampinga. eGFP.pDEST30 plasmid was generously shared by R.L. Wiseman. This work was supported by a Society of Analytical Chemists of Pittsburgh Starter Grant (J.C.G.) and the University of California, Riverside.

REFERENCES

(1) de Lichtenberg, U.; Jensen, L. J.; Brunak, S.; Bork, P. Dynamic Complex Formation during the Yeast Cell Cycle. *Science* **2005**, *307* (5710), 724–727.

(2) Fields, S.; Song, O. K. A Novel Genetic System to Detect Protein Protein Interactions. *Nature* **1989**, *340* (6230), 245–246.

(3) Phizicky, E. M.; Fields, S. Protein-Protein Interactions - Methods for Detection and Analysis. *Microbiol. Rev.* **1995**, *59* (1), 94–123.

(4) Scott, J. D.; Pawson, T. Cell Signaling in Space and Time: Where Proteins Come Together and When They're Apart. *Science* **2009**, *326* (5957), 1220–1224.

(5) von Mering, C.; Krause, R.; Snel, B.; Cornell, M.; Oliver, S. G.; Fields, S.; Bork, P. Comparative Assessment of Large-Scale Data Sets of Protein-Protein Interactions. *Nature* **2002**, *417* (6887), 399–403.

(6) Chien, C. T.; Bartel, P. L.; Sternglanz, R.; Fields, S. The 2-Hybrid System - a Method to Identify and Clone Genes for Proteins That Interact with a Protein of Interest. *Proc. Natl. Acad. Sci. U. S. A.* **1991**, *88* (21), 9578–9582.

(7) Fields, S.; Sternglanz, R. The 2-Hybrid System - an Assay for Protein-Protein Interactions. *Trends Genet.* **1994**, *10* (8), 286–292.

(8) Parrish, J. R.; Gulyas, K. D.; Finley, R. L. Yeast Two-Hybrid Contributions to Interactome Mapping. *Curr. Opin. Biotechnol.* **2006**, *17* (4), 387–393.

(9) Barrios-Rodiles, M.; Brown, K. R.; Ozdamar, B.; Bose, R.; Liu, Z.; Donovan, R. S.; Shinjo, F.; Liu, Y. M.; Dembowy, J.; Taylor, I. W.; et al. High-Throughput Mapping of a Dynamic Signaling Network in Mammalian Cells. *Science* **2005**, *307* (5715), 1621–1625.

(10) Braun, P.; Tasan, M.; Dreze, M.; Barrios-Rodiles, M.; Lemmens, I.; Yu, H. Y.; Sahalie, J. M.; Murray, R. R.; Roncari, L.; de Smet, A. S.; et al. An Experimentally Derived Confidence Score for Binary Protein-Protein Interactions. *Nat. Methods* **2009**, *6* (1), 91–98.

(11) Tang, Y.; Qiu, J.; Machner, M.; LaBaer, J. Discovering Protein-Protein Interactions Using Nucleic Acid Programmable Protein Arrays. *Curr. Protoc. Cell Biol.* **2017**, *74*, 15.21.1-15.21.14. https://doi.org/10.1002/cpcb.14.

(12) Sinz, A. Chemical Cross-Linking and Mass Spectrometry to Map Three-Dimensional Protein Structures and Protein–Protein Interactions. *Mass Spectrom. Rev.* **2006**, *25* (4), 663–682. https://doi.org/10.1002/mas.20082.

(13) Iacobucci, C.; Piotrowski, C.; Aebersold, R.; Amaral, B. C.; Andrews, P.; Bernfur, K.; Borchers, C.; Brodie, N. I.; Bruce, J. E.; Cao, Y.; et al. First Community-Wide, Comparative Cross-Linking Mass Spectrometry Study. *Anal. Chem.* **2019**, *91* (11), 6953–6961. https://doi.org/10.1021/acs.analchem.9b00658.

(14) Kratchmarova, I.; Blagoev, B.; Haack-Sorensen, M.; Kassem, M.; Mann, M. Mechanism of Divergent Growth Factor Effects in Mesenchymal Stem Cell Differentiation. *Science* **2005**, *308* (5727), 1472–1477.

(15) Gingras, A. C.; Gstaiger, M.; Raught, B.; Aebersold, R. Analysis of Protein Complexes Using Mass Spectrometry. *Nat. Rev. Mol. Cell Biol.* **2007**, *8* (8), 645–654.

(16) Gavin, A. C.; Maeda, K.; Kuhner, S. Recent Advances in Charting Protein-Protein Interaction: Mass Spectrometry-Based Approaches. *Curr. Opin. Biotechnol.* **2011**, *22* (1), 42–49.

(17) Pankow, S.; Bamberger, C.; Calzolari, D.; Martinez-Bartolome, S.; Lavallee-Adam, M.; Balch, W. E.; Yates, J. R. Delta F508 CFTR Interactome Remodelling Promotes Rescue of Cystic Fibrosis. *Nature* **2015**, *528* (7583), 510-+.

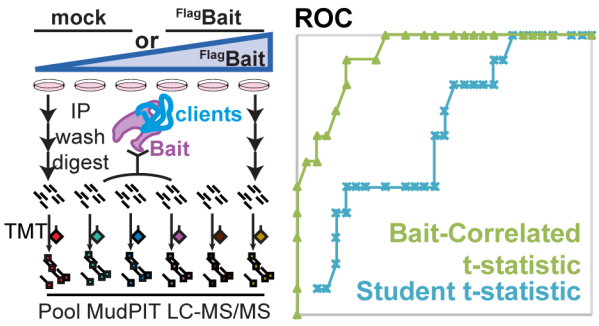
(18) Tarassov, K.; Messier, V.; Landry, C. R.; Radinovic, S.; Molina, M. M. S.; Shames, I.; Malitskaya, Y.; Vogel, J.; Bussey, H.; Michnick, S. W. An in Vivo Map of the Yeast Protein Interactome. *Science* **2008**, *320* (5882), 1465–1470.

(19) Budayeva, H. G.; Cristea, I. M. A Mass Spectrometry View of Stable and Transient Protein Interactions. *Adv. Exp. Med. Biol.* **2014**, *806*, 263–282. https://doi.org/10.1007/978-3-319-06068-2_11.
(20) Lavallée-Adam, M.; Cloutier, P.; Coulombe, B.; Blanchette, M. Modeling Contaminants in AP-MS/MS Experiments. *J. Proteome Res.* **2011**, *10* (2), 886–895. https://doi.org/10.1021/pr100795z.
(21) Mellacheruvu, D.; Wright, Z.; Couzens, A. L.; Lambert, J. P.; St-Denis, N. A.; Li, T.; Miteva, Y. V.; Hauri, S.; Sardiu, M. E.; Low, T. Y.; et al. The CRAPome: A Contaminant Repository for Affinity Purification-Mass Spectrometry Data. *Nat. Methods* **2013**, *10* (8), 730-+.
(22) Huttlin, E. L.; Ting, L.; Bruckner, R. J.; Gebreab, F.; Gygi, M. P.; Szpyt, J.; Tam, S.; Zarraga, G.; Colby, G.; Baltier, K.; et al. The BioPlex Network: A Systematic Exploration of the Human Interactome. *Cell* **2015**, *162* (2), 425–440.
(23) Choi, H.; Larsen, B.; Lin, Z. Y.; Breitkreutz, A.; Mellacheruvu, D.; Fermin, D.; Qin, Z. S.; Tyers, M.; Gingras, A. C.; Nesvizhskii, A. I. SAINT: Probabilistic Scoring of Affinity Purification-Mass Spectrometry Data. *Nat Methods* **2011**, *8* (1), 70–73.
(24) 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. https://doi.org/10.1016/j.cell.2009.04.042.
(25) Huttlin, E. L.; Bruckner, R. J.; Navarrete-Perea, J.; Cannon, J. R.; Baltier, K.; Gebreab, F.; Gygi, M. P.; Thornock, A.; Zarraga, G.; Tam, S.; et al. *Dual Proteome-Scale Networks Reveal Cell-Specific Remodeling of the Human Interactome*; preprint; Systems Biology, 2020. https://doi.org/10.1101/2020.01.19.905109.
(26) Langley, S. R.; Mayr, M. Comparative Analysis of Statistical Methods Used for Detecting Differential Expression in Label-Free Mass Spectrometry Proteomics. *J. Proteomics* **2015**, *129*, 83–92.
(27) Nesvizhskii, A. I. Computational and Informatics Strategies for Identification of Specific Protein Interaction Partners in Affinity Purification Mass Spectrometry Experiments. *Proteomics* **2012**, *12* (10), 1639–1655.
(28) Plate, L.; Rius, B.; Nguyen, B.; Genereux, J. C.; Kelly, J. W.; Wiseman, R. L. Quantitative Interactome Proteomics Reveals a Molecular Basis for ATF6-Dependent Regulation of a Destabilized Amyloidogenic Protein. *Cell Chem. Biol.* **2019**, *26* (7), 913-925.e4. https://doi.org/10.1016/j.chembiol.2019.04.001.
(29) Papachristou, E. K.; Kishore, K.; Holding, A. N.; Harvey, K.; Roumeliotis, T. I.; Chilamakuri, C. S. R.; Omarjee, S.; Chia, K. M.; Swarbrick, A.; Lim, E.; et al. A Quantitative Mass Spectrometry-Based Approach to Monitor the Dynamics of Endogenous Chromatin-Associated Protein Complexes. *Nat. Commun.* **2018**, *9*.
(30) Keilhauer, E. C.; Hein, M. Y.; Mann, M. Accurate Protein Complex Retrieval by Affinity Enrichment Mass Spectrometry (AE-MS) Rather than Affinity Purification Mass Spectrometry (AP-MS). *Mol. Cell. Proteomics* **2015**, *14* (1), 120–135.
(31) Hosp, F.; Scheltema, R. A.; Eberl, H. C.; Kulak, N. A.; Keilhauer, E. C.; Mayr, K.; Mann, M. A Double-Barrel Liquid Chromatography-Tandem Mass Spectrometry (LC-MS/MS) System to Quantify 96 Interactomes per Day. *Mol. Cell. Proteomics* **2015**, *14* (7), 2030–2041.
(32) Youn, J.-Y.; Dunham, W. H.; Hong, S. J.; Knight, J. D. R.; Bashkurov, M.; Chen, G. I.; Bagci, H.; Rathod, B.; MacLeod, G.; Eng, S. W. M.; et al. High-Density Proximity Mapping Reveals the Subcellular Organization of MRNA-Associated Granules and Bodies. *Mol. Cell* **2018**, *69* (3), 517-532.e11. https://doi.org/10.1016/j.molcel.2017.12.020.
(33) Joshi, P.; Greco, T. M.; Guise, A. J.; Luo, Y.; Yu, F.; Nesvizhskii, A. I.; Cristea, I. M. The Functional Interactome Landscape of the Human Histone Deacetylase Family. *Mol. Syst. Biol.* **2013**, *9*, 672. https://doi.org/10.1038/msb.2013.26.
(34) Hageman, J.; Kampinga, H. H. Computational Analysis of the Human HSPH/HSPA/DNAJ Family and Cloning of a Human HSPH/HSPA/DNAJ Expression Library. *Cell Stress Chaperones* **2009**, *14* (1), 1–21. https://doi.org/10.1007/s12192-008-0060-2.

(35) Klock, H. E.; Lesley, S. A. The Polymerase Incomplete Primer Extension (PIPE) Method Applied to High-Throughput Cloning and Site-Directed Mutagenesis. *Methods Mol Biol* **2009**, *498*, 91–103.
(36) Shoulders, M. D.; Ryno, L. M.; Genereux, J. C.; Moresco, J. J.; Tu, P. G.; Wu, C.; Yates, J. R.; Su, A. I.; Kelly, J. W.; Wiseman, R. L. Stress-Independent Activation of XBP1s and/or ATF6 Reveals Three Functionally Diverse ER Proteostasis Environments. *Cell Rep.* **2013**, *3* (4), 1279–1292. https://doi.org/10.1016/j.celrep.2013.03.024.
(37) Washburn, M. P.; Wolters, D.; Yates, J. R. Large-Scale Analysis of the Yeast Proteome by Multidimensional Protein Identification Technology. *Nat. Biotechnol.* **2001**, *19* (3), 242–247. https://doi.org/10.1038/85686.
(38) Eng, J. K.; McCormack, A. L.; Yates, J. R. An Approach to Correlate Tandem Mass Spectral Data of Peptides with Amino Acid Sequences in a Protein Database. *J. Am. Soc. Mass Spectrom.* **1994**, *5* (11), 976–989.
(39) Xu, T.; Park, S. K.; Venable, J. D.; Wohlschlegel, J. A.; Diedrich, J. K.; Cociorva, D.; Lu, B.; Liao, L.; Hewel, J.; Han, X.; et al. ProLuCID: An Improved SEQUEST-like Algorithm with Enhanced Sensitivity and Specificity. *J. Proteomics* **2015**, *129*, 16–24.
(40) Tabb, D. L.; McDonald, W. H.; Yates, J. R. DTASelect and Contrast: Tools for Assembling and Comparing Protein Identifications from Shotgun Proteomics. *J Proteome Res* **2002**, *1* (1), 21–26.
(41) Park, S. K.; Venable, J. D.; Xu, T.; Yates, J. R. A Quantitative Analysis Software Tool for Mass Spectrometry-Based Proteomics. *Nat Methods* **2008**, *5* (4), 319–322.
(42) Perez-Riverol, Y.; Csordas, A.; Bai, J.; Bernal-Llinares, M.; Hewapathirana, S.; Kundu, D. J.; Inuganti, A.; Griss, J.; Mayer, G.; Eisenacher, M.; et al. The PRIDE Database and Related Tools and Resources in 2019: Improving Support for Quantification Data. *Nucleic Acids Res.* **2019**, *47* (D1), D442–D450. https://doi.org/10.1093/nar/gky1106.
(43) Zauber, H.; Kirchner, M.; Selbach, M. Picky: A Simple Online PRM and SRM Method Designer for Targeted Proteomics. *Nat. Methods* **2018**, *15* (3), 156–157. https://doi.org/10.1038/nmeth.4607.
(44) Peterson, A. C.; Russell, J. D.; Bailey, D. J.; Westphall, M. S.; Coon, J. J. Parallel Reaction Monitoring for High Resolution and High Mass Accuracy Quantitative, Targeted Proteomics. *Mol. Cell. Proteomics* **2012**, *11* (11), 1475–1488. https://doi.org/10.1074/mcp.O112.020131.
(45) Sherrod, S. D.; Myers, M. V.; Li, M.; Myers, J. S.; Carpenter, K. L.; MacLean, B.; MacCoss, M. J.; Liebler, D. C.; Ham, A.-J. L. Label-Free Quantitation of Protein Modifications by Pseudo Selected Reaction Monitoring with Internal Reference Peptides. *J. Proteome Res.* **2012**, *11* (6), 3467–3479.
(46) Smyth, G. K. Linear Models and Empirical Bayes Methods for Assessing Differential Expression in Microarray Experiments. *Stat. Appl. Genet. Mol. Biol.* **2004**, *3* (1), 1–25. https://doi.org/10.2202/1544-6115.1027.
(47) Krammer, F.; Pica, N.; Hai, R.; Margine, I.; Palese, P. Chimeric Hemagglutinin Influenza Virus Vaccine Constructs Elicit Broadly Protective Stalk-Specific Antibodies. *J. Virol.* **2013**, *87* (12), 6542–6550.
(48) Oberg, A. L.; Mahoney, D. W.; Eckel-Passow, J. E.; Malone, C. J.; Wolfinger, R. D.; Hill, E. G.; Cooper, L. T.; Onuma, O. K.; Spiro, C.; Therneau, T. M.; et al. Statistical Analysis of Relative Labeled Mass Spectrometry Data from Complex Samples Using ANOVA. *J. Proteome Res.* **2008**, *7* (1), 225–233. https://doi.org/10.1021/pr700734f.
(49) Fisher, R. A. *Statistical Methods for Research Workers*, 14th ed., revised and enlarged.; Oliver and Boyd: Edinburgh, 1970.
(50) Hogg, R. V.; Tanis, E. A. *Probability and Statistical Inference*, 2nd ed.; Macmillan ; Collier Macmillan: New York : London, 1983.
(51) Benjamini, Y.; Hochberg, Y. Controlling the False Discovery Rate - a Practical and Powerful Approach to Multiple Testing. *J. R. Stat. Soc. Ser. B-Stat. Methodol.* **1995**, *57* (1), 289–300.

(52) Yekutieli, D.; Benjamini, Y. Resampling-Based False Discovery Rate Controlling Multiple Test Procedures for Correlated Test Statistics. *J. Stat. Plan. Inference* **1999**, *82* (1–2), 171–196. https://doi.org/10.1016/S0378-3758(99)00041-5.
(53) Storey, J. D.; Tibshirani, R. Statistical Significance for Genomewide Studies. *Proc. Natl. Acad. Sci. U. S. A.* **2003**, *100* (16), 9440–9445.
(54) Swets, J. Measuring the Accuracy of Diagnostic Systems. *Science* **1988**, *240* (4857), 1285–1293. https://doi.org/10.1126/science.3287615.
(55) Kampinga, H. H.; Craig, E. A. The HSP70 Chaperone Machinery: J Proteins as Drivers of Functional Specificity (Vol 11, Pg 579, 2010). *Nat. Rev. Mol. Cell Biol.* **2010**, *11* (10).
(56) Jin, Y.; Awad, W.; Petrova, K.; Hendershot, L. M. Regulated Release of ERdj3 from Unfolded Proteins by BiP. *EMBO J.* **2008**, *27* (21), 2873–2882. https://doi.org/10.1038/emboj.2008.207.
(57) Hageman, J.; Rujano, M. A.; van Waarde, M. A. W. H.; Kakkar, V.; Dirks, R. P.; Govorukhina, N.; Oosterveld-Hut, H. M. J.; Lubsen, N. H.; Kampinga, H. H. A DNAJB Chaperone Subfamily with HDAC-Dependent Activities Suppresses Toxic Protein Aggregation. *Mol. Cell* **2010**, *37* (3), 355–369. https://doi.org/10.1016/j.molcel.2010.01.001.
(58) Liu, H.; Zhang, H.; Niedzwiedzki, D. M.; Prado, M.; He, G.; Gross, M. L.; Blankenship, R. E. Phycobilisomes Supply Excitations to Both Photosystems in a Megacomplex in Cyanobacteria. *Science* **2013**, *342* (6162), 1104–1107. https://doi.org/10.1126/science.1242321.
(59) Kammers, K.; Cole, R. N.; Tiengwe, C.; Ruczinski, I. Detecting Significant Changes in Protein Abundance. *EuPA Open Proteomics* **2015**, *7*, 11–19. https://doi.org/10.1016/j.euprot.2015.02.002.
(60) Walker, E. J.; Bettinger, J. Q.; Welle, K. A.; Hryhorenko, J. R.; Ghaemmaghami, S. Global Analysis of Methionine Oxidation Provides a Census of Folding Stabilities for the Human Proteome. *Proc. Natl. Acad. Sci.* **2019**, *116* (13), 6081–6090. https://doi.org/10.1073/pnas.1819851116.
(61) West, G. M.; Tucker, C. L.; Xu, T.; Park, S. K.; Han, X.; Yates, J. R.; Fitzgerald, M. C. Quantitative Proteomics Approach for Identifying Protein-Drug Interactions in Complex Mixtures Using Protein Stability Measurements. *Proc. Natl. Acad. Sci.* **2010**, *107* (20), 9078–9082. https://doi.org/10.1073/pnas.1000148107.
(62) Chatr-aryamontri, A.; Oughtred, R.; Boucher, L.; Rust, J.; Chang, C.; Kolas, N. K.; O'Donnell, L.; Oster, S.; Theesfeld, C.; Sellam, A.; et al. The BioGRID Interaction Database: 2017 Update. *Nucleic Acids Res.* **2017**, *45* (D1), D369–D379.
(63) Lim, G. E.; Johnson, J. D. 14-3-3ζ: A Numbers Game in Adipocyte Function? *Adipocyte* **2016**, *5* (2), 232–237. https://doi.org/10.1080/21623945.2015.1120913.
(64) Johnson, C.; Tinti, M.; Wood, N. T.; Campbell, D. G.; Toth, R.; Dubois, F.; Geraghty, K. M.; Wong, B. H. C.; Brown, L. J.; Tyler, J.; et al. Visualization and Biochemical Analyses of the Emerging Mammalian 14-3-3-Phosphoproteome. *Mol. Cell. Proteomics* **2011**, *10* (10), M110.005751. https://doi.org/10.1074/mcp.M110.005751.
(65) Mugabo, Y.; Sadeghi, M.; Fang, N. N.; Mayor, T.; Lim, G. E. Elucidation of the 14-3-3ζ Interactome Reveals Critical Roles of RNA-Splicing Factors during Adipogenesis. *J. Biol. Chem.* **2018**, *293* (18), 6736–6750. https://doi.org/10.1074/jbc.M117.816272.
(66) Budnik, B.; Levy, E.; Harmange, G.; Slavov, N. SCoPE-MS: Mass Spectrometry of Single Mammalian Cells Quantifies Proteome Heterogeneity during Cell Differentiation. *Genome Biol.* **2018**, *19* (1), 161. https://doi.org/10.1186/s13059-018-1547-5.

TOC FIGURE


mock
FlagBait
or
FlagBait
IP
wash
digest
clients
Bait
TMT
Pool MudPIT LC-MS/MS
ROC
Bait-Correlated
t-statistic
Student t-statistic