

Supplementary Note

Samples and mass spectrometry analysis

The Universal Proteomics Standard 1 (UPS1), which contains a mixture of 48 human proteins at equimolar concentrations was acquired from Sigma-Aldrich (St. Louis, MO). The MassPREP *E.coli* Digest Standard was from Waters (Milford, MA) and the MS Compatible Human Protein Extract, Digest was from Promega (Madison, WI). UPS1 standards were resuspended in 100 mM Tris pH 8, reduced with 5 mM TCEP for 30 min at 65°C, then alkylated with 50 mM iodoacetamide for 60 min in the dark. UPS1 standards were digested overnight with 1 µg trypsin (Promega, Madison, WI) in 100 mM Tris pH 8 at 37°C. UPS1, *E.coli* and human peptides were acidified with formic acid to 1% final concentration and loaded in various amounts, alone or in combination, onto an in-house made 75 µm x 12 cm analytical column emitter with 3 µm ReproSil-Pur C18-AQ (Dr.Maisch HPLC GmbH, Germany). A NanoLC-Ultra 1D plus (Eksigent, Dublin CA) nano-pump was used to deliver a 90 min gradient from 2% to 35% acetonitrile with 0.1% formic acid, followed by a 30 min wash with 80% acetonitrile and re-equilibration to 2% acetonitrile. Each sample condition was acquired twice on a TripleTOF 5600 mass spectrometer (AB SCIEX, Concord, Ontario, Canada), once using a Data Dependent Acquisition (DDA) method and once using the SWATH Data Independent Acquisition (DIA) method. The DDA run consisted of one 250 ms MS1 TOF survey scan from 400–1250 Da followed by ten 100 ms MS2 candidate ion scans (former precursors were excluded for 15 sec). The DIA run consisted of one 250 ms MS1 TOF survey scan followed by 33 sequential MS2 windows of 25 amu covering a mass range from 400–1250 Da at 100 ms per SWATH scan. The affinity purified (AP) samples and their mass spectrometry acquisition parameters were described previously¹. Briefly, bait proteins EIF4A2, MEPCE and GFP (as negative control) with FLAG tags were affinity purified from cell lysate and digested on beads with trypsin. Each AP experiment was performed in biological triplicates and the resulting samples were analyzed by both DDA and DIA using the parameters described above, with the exception of a 50 ms MS1 scan in the DIA analysis and the use of a Nanoflex cHiPLC system at 200 nL/min (Eksigent ChromXP C18 3 µm × 75 µm × 15 cm column chip) for the chromatography. See **Supplementary Table 3** for the description of the samples.

Peptide identification from DDA

Wiff files were converted to mzXML using the AB SCIEX MS Data Converter with the “ProteinPilot” option. DDA data were searched against their corresponding sequence databases using MSGFDB² with 50 ppm precursor mass tolerance while allowing for N-term acetylation (Nterm+42.012) and methionine oxidation (M+16.02) as variable modifications and carbamidomethylation on cysteine (C+57.021) as a fixed modification. The sequence database for UPS1 proteins was downloaded from the Sigma website (<http://www.sigmaaldrich.com/life-science/proteomics/mass-spectrometry/ups1-and-ups2-proteomic.html>) and UPS1-only samples were searched against this database concatenated with the *E. coli* database for accuracy of

FDR estimation with the Target/Decoy Approach^{3,4}. The *E. coli* protein sequence database was downloaded from NCBI with Taxonomy ID: 511145, ver. 08/25/2009 and the human database was downloaded from NCBI RefSeq, ver. 10/29/2010. All searches results were filtered by enforcing a 1% peptide-level false discovery rate (FDR) using the Target/Decoy approach (TDA)³.

Spectral libraries for MSPLIT-DIA

DDA search results were used to create three sample-specific spectral libraries: 1) an “*E. coli*” spectral library with peptides from UPS1 samples and *E. coli* lysate peptides; 2) a “*human*” spectral library with peptides from UPS1 samples and human lysate peptides; 3) an “*AP-specific*” spectral library with peptides from all AP-MS samples (See **Supplementary Table 3**). In all cases, the Peptide-Spectrum Matches (PSMs) from the corresponding DDA runs were pooled by retaining only the spectrum with the lowest MSGFDB probability for each unique peptide sequence and precursor charge state, and a peptide-level false discovery rate (FDR) of 1% was enforced using TDA³. We note that this is different from simply combining all peptides identified at 1% FDR from each run, which is undesirable as the errors in the combined peptide lists will grow proportionally to the number of runs included in the library. Our approach required us to use a stricter score threshold than that used for each individual run, which therefore resulted in some peptides identified in individual runs not being included in the pooled spectral library (see **Supplementary Figure 3b**). Each library spectrum was processed to reduce noise by removing all peaks that did not match to a theoretical fragment ion from the identified peptide within 0.05 Da mass tolerance; fragment ion types used to compute the theoretical ions were: b, y, b-H₂O, b-NH₃, y-H₂O, y-NH₃. For each ion type, charge states from +1 to the charge state of the precursor were considered. After constructing the target library, a decoy library of equal size was generated and appended to the library. Briefly, for each target spectrum in the library, an artificial decoy spectrum was generated using the shuffle-and-reposition method described previously⁵. Unless otherwise indicated, spectral libraries for MSPLIT-DIA matching were built from the union of all DDA runs of the same samples.

To further assess the applicability of MSPLIT-DIA using community-wide spectral libraries, a comprehensive and generic SWATH-Atlas library constructed by the Aebersold group from multiple human DDA runs on a TripleTOF 5600+ instrument (many of which from fractionated samples) was downloaded from <http://www.swathatlas.org/>⁶.

Identification of multiplexed spectra by MSPLIT-DIA

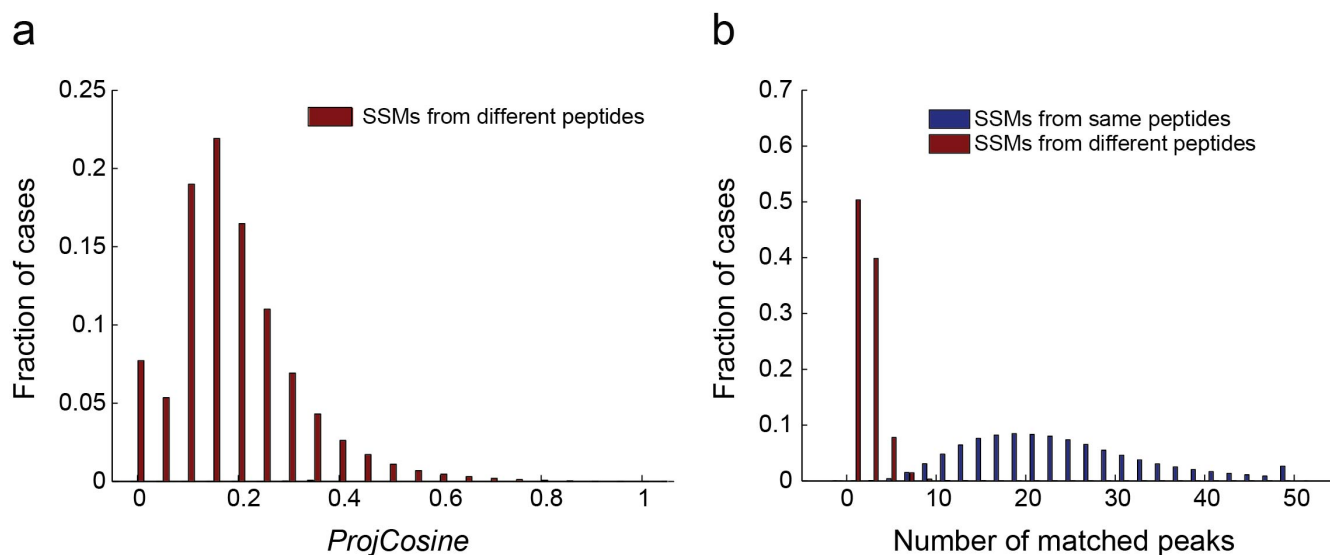
An overview of the proposed spectral library search method to identify multiplexed spectra in DIA data is illustrated in **Fig. 1**; a typical multiplex spectrum after MSPLIT-DIA identification is shown in **Supplementary Figure 1**. We define a multiplexed spectrum as an MS/MS spectrum derived from more than one peptide.

Identification of multiplexed spectra is performed by comparing each multiplexed spectrum to a library of reference single-peptide spectra and returning all spectra in the library that have significant similarity to the multiplexed spectrum. For each DIA spectrum, all library spectra with precursor mass in the DIA isolation window are considered valid candidates (e.g. the width of the isolation window used in this study is 25 Da). Note that, in contrast to targeted approaches for analyzing DIA data, MSPLIT-DIA does not require retention-time information for library peptides because MSPLIT-DIA can scan through the whole retention-time range for possible spectrum-spectrum matches (SSM).

Specifically, we represent an MS/MS spectrum as a set of peaks: $S = \{p_1, p_2, \dots, p_n\}$ where each peak is a mass and intensity value pair. We define a spectral library L as a set of annotated spectra: $\mathcal{L} = \{l_1, l_2, \dots, l_m\}$, where each spectrum l_i is from one unique peptide and precursor charge state. To identify a multiplexed spectrum M , we first reduce noise in the spectrum using a window filtering method⁷: each peak is compared to peaks in a neighborhood of ± 25 Da and kept if it is within the 15 most intense peaks in its neighborhood. Then the similarity between M and each library spectrum l within the precursor mass tolerance of M is computed. However because M is potentially composed of multiple peptides, conventional similarity metrics such as the cosine similarity do not work well due to the presence of many unmatched peaks from other peptides in M . The concept of *projected-spectrum*⁸ was therefore used to address this issue. Given M and a library spectrum l , the projected-spectrum M_l is determined as follows: for each peak p in the library spectrum l , we select peaks in M within a given mass error tolerance δ of p , where δ depends on instrument resolution (δ was set to 50 ppm for the data analyzed here). If multiple peaks in M happen to match to p , only the peak with the highest intensity is retained in M_l . Other methods such as taking the matched peak with the smallest mass error to the library peak or merging all matched peaks by summing up their intensities were also tested but resulted in only minor differences in the performance of peptide identification. Intuitively, spectrum projection extracts a set of matching peaks in M that are most likely to belong to the same peptide as l while ignoring peaks that likely arise from other co-eluting peptides in M . After normalization of the spectra to Euclidian norm one, the similarity of l to the projected spectrum M_l can be evaluated using the conventional cosine similarity function⁹. A variant of the cosine similarity function, *projCosine*, can thus be defined to measure the similarity between a multiplexed spectrum M and a single-peptide spectrum l :

$$\text{projCosine}(M, l) = \text{cosine}(M_l, l)$$

The *projCosine* is computed between M and all library spectra within the selected precursor mass range, and all spectrum-spectrum matches (SSMs) with similarity greater than a significance threshold are considered matches to M . This threshold was determined using a target-decoy approach (see FDR estimation section below)³. A minimum of ten matched peaks between the library spectrum and the multiplexed spectrum was also required. As shown in **Supplementary Note Figure Sn1**, nearly all matching SSMs (spectral matches from the same peptides) have ten or more matching peaks.



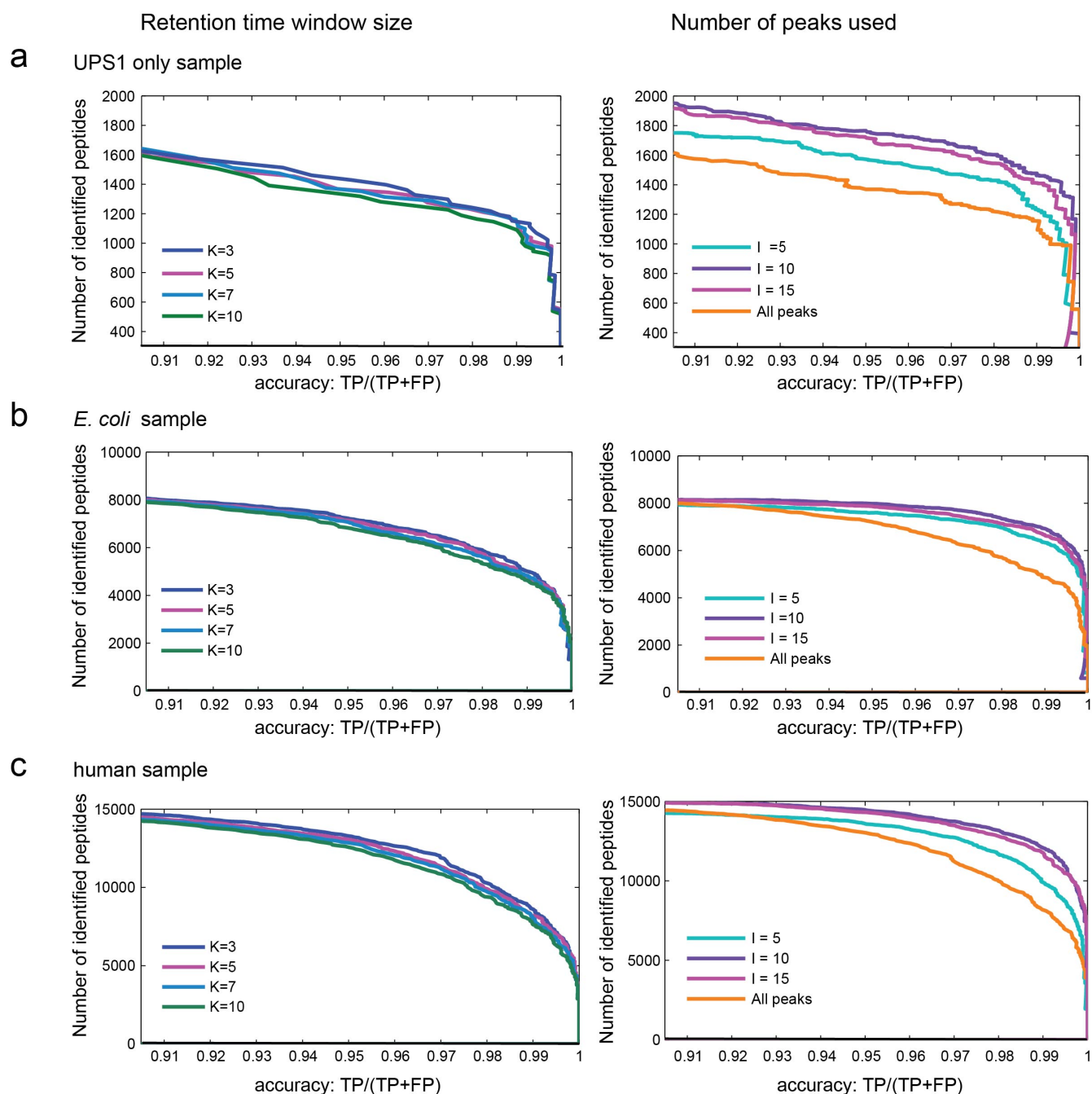
Supplementary Note Figure SN1: Spectrum-Spectrum-Matches (SSM) statistics. Using the set of all Peptide-Spectrum Matches (PSMs) identified at 1% peptide-level FDR in the DDA datasets (see Supplementary Table 3), we randomly selected 10,000 correct and 10,000 incorrect Spectrum-Spectrum Matches (SSMs). A matching SSM is a pair of PSMs identified to the same peptide sequence and charge state; a non-matching SSM is a pair of PSMs identified to different peptide sequences. (a) Projected cosines of the non-matching SSMs. The projected cosines similarity for non-matching SSMs (red) is rarely greater than 0.7, thus this was used as the minimum similarity threshold to detect dependent matches to a multiplexed spectrum. (b) Distribution of the number of matched peaks for matching or non-matching SSMs. The distribution of the number of matching peaks for matching (blue) and non-matching (red) SSMs shows that most matching SSMs have at least 10 matching peaks while non-matching SSMs have less than 10. Therefore, MSPLIT-DIA requires that a library spectrum have at least ten matching peaks to a multiplexed spectrum.

When identifying multiplexed spectra, it is critical to avoid using the same set of peaks in a multiplexed spectrum to support the identification of two (or more) different peptides, as this would lead to artificially high numbers of peptides matched to the same multiplexed spectrum. Examples of such cases include peptides that differ in their sequence only by a reversal of adjacent amino acids such as ID/DI in PEPTIDE and PEPTDIE respectively. In such cases, library spectra from these peptides may match to a multiplexed spectrum with similar good scores even if only one of these peptides is present in the sample. This issue is commonly ignored when analyzing DDA spectra because spectra are assumed to contain only one peptide and thus usually only the best-scoring peptide is considered to be a match. However, when dealing with multiplexed spectra, if there are entries in the spectral library that share many fragment ions, correct matches to these spectra can easily lead to dependent false positive multiplexed matches. In order to limit the dependency between multiple matches to the same multiplexed spectrum, the pairwise similarity between all library spectra that are matched to a multiplexed spectrum is computed. When two library spectra have high similarity (i.e. $projCosine > 0.7$), only the spectrum with highest $projCosine$ to M is returned. The similarity threshold was chosen based on the observation that non-matching SSMs (spectrum matches from different peptides) rarely have a $projCosine$ similarity greater than 0.7 (see **Supplementary Note Figure SN1**). To further assess the performance of this threshold, we determined the fraction of MSPLIT-DIA SSMs at 1% FDR containing dependent peptides (i.e., spectra from peptides differing only by one acetylation or one oxidation) in the same multiplexed spectrum and found that it occurred in less than 0.1% of cases (well below our FDR threshold).

In DIA, each peptide is acquired and analyzed in multiple MS/MS spectra along its elution profile. Thus the ion chromatograms of peaks corresponding to fragment ions from the same peptide precursor correlate in retention time^{10, 11}. To utilize this information, rather than analyzing each DIA spectrum as an independent spectrum, consecutive DIA spectra were analyzed together using a sliding-window approach. For a particular DIA spectrum M_t at retention time t we define its neighbor set $N_{t,K} = \{M_{t-K}, \dots, M_t, \dots, M_{t+K}\}$ as M_t plus K spectra acquired right before M_t and K spectra acquired right after M_t with the same DIA precursor isolation window. Given a library spectrum $l = \{p_1, p_2, \dots, p_n\}$, the projected spectrum for each M_j in the neighbor set can be computed. Therefore each peak p_i in the library spectrum l has a set of up to $2K+1$ matched peaks from the neighbor set. A vector of $2K+1$ elements can then be generated by taking the intensity of matched peaks from each $M_j, t - K \leq i \leq t + K$ (if no matching peak is found, an intensity of zero is used). This vector is then scaled to Euclidean norm one to define the retention time profile $T(p_i, M_t, K)$ for a library peak p_i when matched to M_t . The *cosine* similarity between the time profiles of two peaks p_i and p_j can then be computed to assess whether the pair of peaks co-elute along the retention time window $t-K$ to $t+K$:

$$timeCosine(M_t, p_i, p_j) = cosine(T(p_i, M_t, K), T(p_j, M_t, K))$$

The function *timeCosineSet* can then be computed for a set of peaks by taking the average *timeCosine* for all possible pairs of peaks in the set. Thus *timeCosineSet* can be used to assess whether the matched peaks between a library spectrum and a multiplexed spectrum co-elute and help separate correct matches from false positive matches (by random chance, a set of false matched peaks should not co-elute along chromatographic time). We also found that considering the *timeCosineSet* for only the top I most intense peaks in the library spectrum was more informative than using all peaks. This can be due to the fact that low-intensity peaks are more likely to be noise peaks and/or to have interference in multiplexed DIA spectra. Different values of I and K were tested and we found $I=10$ and $K=5$ to be robust across the samples of varying complexity analyzed here (see **Supplementary Note Figure SN1**). Finally, to combine information from spectral similarity and retention time correlation, we defined the combined score as the product of *projCosine* and *timeCosineSet* and used this score to rank the SSMs resulting from library search.



Supplementary Note Figure SN1. Parameters for *timeCosineSet*.

MSPLIT-DIA evaluates the retention-time correlation of matched peaks between a library spectrum and a DIA spectrum (i.e. *timeCosineSet*) to assess whether these matched peaks co-elute in time and are thus likely to come from the same peptide precursor. Two parameters affecting the calculation of *timeCosineSet* were varied: K – the retention-time window in which the correlation is computed (left-side panels; all peaks were considered) and l – the number of most intense peaks with which the correlation is computed (right-side panels, K was fixed at 5). The number of peptides identified at different FDR was used to assess the effect of different parameters on performance. We found l=10 and K=5 to be the most suitable choices for all samples analyzed here.

Estimation of FDR in MSPLIT-DIA

The threshold for reporting Spectrum-Spectrum-Matches (SSMs) was determined by enforcing False Discovery Rate (FDR) using the Target-Decoy Approach (TDA)³. Since by design each peptide is present in many MS/MS

spectra in DIA data, multiple true SSMs are likely to correspond to the same peptide while false SSMs (i.e., random matches) are more likely to correspond to different peptides. As such, an SSM-level FDR will likely yield a substantially higher FDR for peptide identifications⁴. We thus opted to enforce a more conservative peptide-level FDR by considering only the best-scoring SSM for each unique peptide and precursor charge for FDR enforcement via TDA. Once the score threshold is determined, all SSMs with scores above the threshold are accepted as matches for SSM-level analysis (note that this only adds more SSMs from the same peptide and does not change the peptide-level FDR).

The TDA for spectral library search was originally proposed⁵ for analyzing DDA data where most MS/MS spectra originate from a single peptide. However most spectra in DIA data of moderately complex samples are multiplexed spectra (see **Fig. 1b**). To test whether the TDA approach can be extended to DIA data, we searched the *E. coli* lysate DIA data against a spectral library containing peptides from both *E. coli* and UPS1 proteins. Since UPS1 peptides are not expected to be present in the *E. coli* samples (all UPS1 proteins are from other species), identified peptides mapped to the UPS1 proteins are considered as false matches. In this spectral library, approximately 17.9% of peptides are from UPS1 proteins and thus it is expected that the same fraction of all false positive matches will be mapped to UPS1 proteins by random chance. Therefore the empirical number of false positive library matches in DIA data can be estimated by dividing the number of identified peptides mapped to UPS1 proteins by their relative proportion in the library. At a 1% FDR as enforced by TDA, we identified 12 peptides from UPS1 proteins out of a total of 6274 peptides from the 1 μ g *E. coli* lysate data, which corresponds to an empirical $FDR = \frac{12/0.179}{6274} = 1.07\%$, thus supporting the accuracy of the TDA estimation of FDR.

Using retention-time information for MSPLIT-DIA peptide identification

Even though MSPLIT-DIA is designed to operate without prior knowledge of retention-time (RT) for library peptides, it can utilize RT information to improve sensitivity when it is available in the spectral library (See **Supplementary Figure 9**). MSPLIT-DIA can optionally conduct a two-pass search. The first-pass search is as described in the previous section in which MSPLIT-DIA scans the whole RT range for possible matches. Then, using the list of peptides identified at 1% FDR, MSPLIT-DIA builds a linear regression model between the library RT and the RT at which the peptides are detected in the DIA data using the least-squares approach, thus automatically correcting for possible systematic RT differences. MSPLIT-DIA then uses the identified peptides to automatically learn the distribution of RT differences between library and DIA peptides and thus determines a RT tolerance window that includes at least 95% of peptides identified in the first pass. In the second-pass search, only matches with library and DIA RT differences that fall within this tolerance window are considered. Since MSPLIT-DIA uses RT information when it is available in the library, results obtained with the SWATH-Atlas spectral library used RT information unless otherwise noted.

Protein identification

Even though the main focus of the current work is on peptide identification, we also assessed the impact of the different methods on protein identification using a simple method for protein inference from the identified peptides¹². First only unique peptides (i.e., mapping to only one protein entry) identified at 1% peptide-level FDR were considered for protein identification. Then, similar to computing the peptide-level FDR (where each peptide is assigned a score which is the highest score from all SSMs matching to that peptide), protein scores were assigned to the best score from all unique peptides mapped to each protein. We note that while this is a simple way to assign protein scores, it has actually been recently shown¹³ to outperform popular approaches such as the more complex ProteinProphet calculation. Finally, protein-level FDR was determined using the standard TDA as for peptides and SSMs.

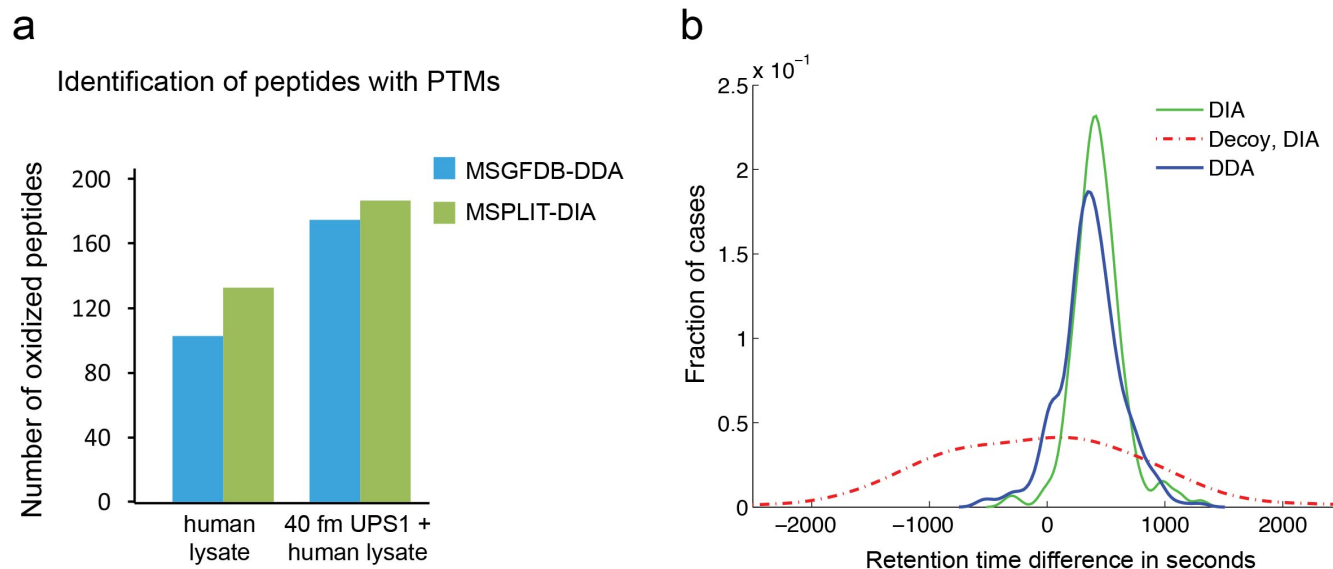
Significance analysis of interaction data

The relative abundance of each protein in each of the triplicate purifications of the EIF4A2 or MEPCE baits in relation to the triplicate purifications of GFP alone (negative control sample) was calculated using spectral counting followed by SAINTexpress¹⁴ analysis to detect proteins that display statistically significant enrichment over control. Only unique peptides (peptides not shared between multiple proteins) were considered, and the abundance of each protein was defined as the sum of spectral counts from all its unique peptides. SAINTexpress analysis was performed with default parameters from www.crapome.org¹⁵, and the results were filtered at 1% Bayesian FDR (**Supplementary Table 1**). The overlap between DDA and DIA data for these high-confidence proteins was analyzed, and the common proteins and proteins specifically detected in the MSPLIT-DIA analysis were further analyzed by DAVID¹⁶ for Gene Ontology enrichment (Molecular Function GO category, with broadest terms filtered out; GO_MF_FAT). The resulting enrichment values are shown in **Supplementary Table 2**. All hits were manually inspected.

Identification of peptides with PTMs by MSPLIT-DIA

MSPLIT-DIA can be used to search for peptides with post-translational modifications (PTMs) in DIA data, even if the modified spectrum is not initially present in the library. The PTMs to search for are specified by the user as inputs. Similar to spectral library search methods for DDA data^{17, 18}, MSPLIT-DIA predicts the library spectra for modified peptides based on library spectra of the unmodified peptides. Specifically, for a library spectrum $l = \{p_1, p_2, \dots, p_m\}$ of an unmodified peptide U of length n . Let $U(\Delta, i)$ be a modified version of peptide U with a PTM of mass Δ on the i^{th} amino acid. The library spectrum for $U(\Delta, i)$ is predicted as follows: 1) associate each peak $p \in l$ with a theoretical fragment ion t_j^c where t denotes the type of the fragment ion, c denotes the charge state and j denotes the amino acid position of the fragment ion (e.g., a doubly-charged b_3

ions will be represented as b_3^2). If multiple fragment ions can be matched to a peak, the one with the closest mass is used; 2) a peak is defined as modified if it corresponds to a fragment ion such that: $\{t_j^c | j \geq i\}$ if t is a prefix (n-terminal) ion or $\{t_j^c | j > n - i\}$ if t is a suffix (c-terminal) ion. The library spectrum of the modified peptide is constructed by shifting all the modified peaks by Δ/c along the m/z axis while retaining the peak intensities from the unmodified peptide spectrum. Results for the identification of oxidized peptides are shown in **Supplementary Note Figure SN3**.



Supplementary Note Figure SN3. Identification of peptides with PTMs.

One of the goals of discovery tools in proteomics is to detect peptides with variable modifications. For targeted approaches, these modified peptides need to be included a priori in the spectral library. A key feature of MSPLIT-DIA is that peptide detection can be performed without using any retention time constraints in DIA searches and without requiring normalized retention times in the spectral library (thus making it much easier to build these libraries from public datasets). In particular, this allows MSPLIT-DIA searches for peptides that are not in the spectral library but can be extrapolated from peptides in the library. For example, even though MSPLIT-DIA was not designed for thorough identification of post-translationally modified peptides, we were able to use library spectra of unmodified peptides to predict and identify spectra of the oxidized version of the same peptides. To demonstrate this, we allowed for methionine oxidation as a variable modification while searching the human lysate data against the human lysate library consisting of only unmodified peptides. The library spectra for modified peptides were predicted by shifting corresponding peaks to the appropriate m/z location. (a) Number of human lysate and UPS1 peptides with oxidized methionine detected by MSGFDB-DDA and MSPLIT-DIA; MSPLIT-DIA was able to identify 7-28% more peptides with methionine oxidation than MSGFDB-DDA. (b) Retention time difference between the modified and unmodified species detected. Since it was previously shown that peptides with a particular PTM display a systematic shift in retention time (RT) compared to the corresponding unmodified peptides¹⁹, we plotted the distribution of the RT difference between identified peptides with or without methionine oxidation in DIA and DDA runs and verified that these display very similar distributions (blue and green lines). By contrast, the modified and unmodified peptides pairs from decoy matches (red dashed line) display a much wider distribution than that of pairs from target matches.

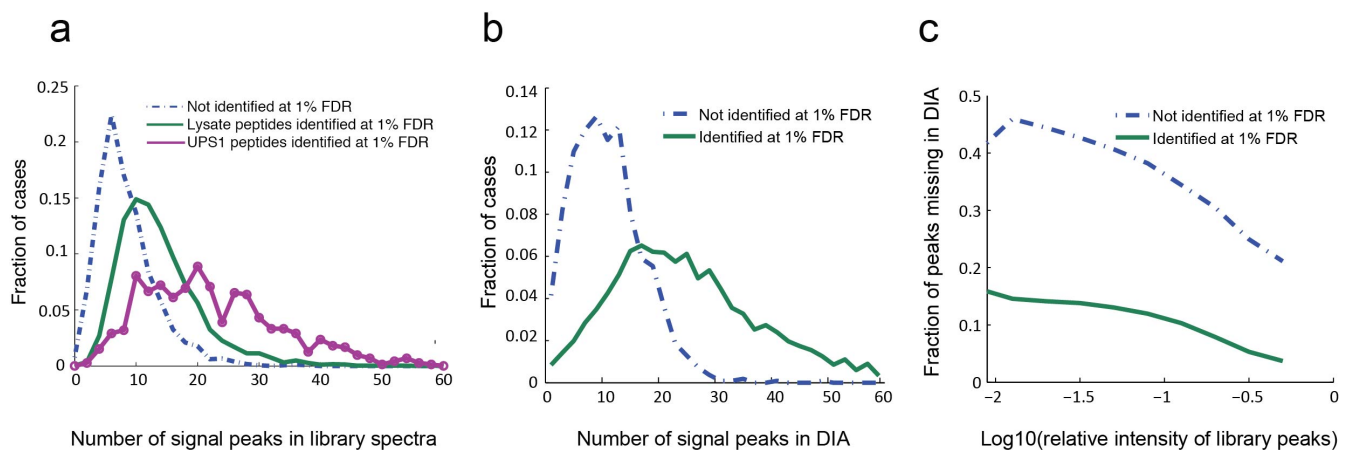
Reproducibility of peptide identifications in DIA and DDA

To compute reproducibility in peptide identification, multiple injections of *E. coli* and human lysate samples were analyzed both by DDA and DIA. We defined a set of identified peptides in the sample as peptides that are identified in at least one DDA run and at least one DIA run at 1% FDR. A peptide is reproducibly identified in K runs if it passes the 1% FDR in all K runs. Reproducibility was computed separately for DDA and DIA runs. To investigate the relationship between peptide abundance and reproducibility, we defined peptide abundance as the average spectral counts of a peptide over the multiple DIA runs under consideration. Peptides were then

binned according to their abundance and the reproducibility of peptide identification across different abundance bins was assessed (see **Supplementary Figure 4**). We also investigated the major factors affecting the reproducibility of peptide identification. A peptide may not be reproducibly identified across multiple runs because: 1) it was not selected for MS/MS analysis (in DDA) or was not detectable above background noise (in the case of DIA) or 2) it was not identified by the computational method at the selected FDR. We investigated possibility #2 by varying the FDR threshold and monitoring reproducibility, which indicated that possibility #2 was the source of most of the apparent non-reproducibility in DIA but not DDA data (**Supplementary Figure 4**).

Effect of spectral quality in peptide identification

In order to investigate the factors that affect peptide identification in DIA data, the list of identified peptides from the parallel DDA run was used as a reference to assess their detectability in each SWATH run. The reference peptide list was separated into two groups: those that were identified by MSPLIT-DIA at 1% FDR in DIA and those that were not. For each peptide that was not identified by MSPLIT-DIA at 1% FDR, retention time information from the DDA run was used to help locate the peptides in DIA by only considering DIA spectra within 5 min of the recorded DDA retention time for the corresponding peptide. The best-scoring DIA spectrum was then considered an unidentified match for that peptide. Two statistics were calculated for the set of identified and unidentified peptides in DIA: i) the number of signal peaks in the library spectrum and ii) the number of matched peaks between a library spectrum and a DIA spectrum that are also signal peaks in the DIA spectrum. A signal peak is defined as a peak with signal-to-noise ratio of at least two. Intensity level for noise peaks was estimated as the median intensity of the 50% lowest-intensity peaks in a spectrum. Signal-to-noise ratio was then computed by dividing the intensity of each peak by the intensity level for noise peaks (see **Supplementary Note Figure SN4**).



Supplementary Note Figure SN4. Spectrum quality in library and DIA spectra.

The 40 fm UPS1 + *E. coli* (see **Supplementary Table 3**) was used as a representative run to illustrate the investigation of peptide detectability in DIA data (analysis from other runs shows similar trends). Peptides identified in the parallel DDA run were used as a reference set and divided into two groups: those that were identified by MSPLIT-DIA at 1% FDR and those that were not. Each peptide was investigated with respect to its best-matching DIA spectrum in the data. Peptides not identified by MSPLIT-DIA were assigned to their DIA spectra with retention-time information from the parallel DDA run to

reduce the chance of false matches. Factors that distinguish the two groups of peptides were investigated as they represent properties that affect peptide identification in DIA data. Panel (a) shows the distribution of the number of signal peaks in library spectra, where signal peaks are MS/MS peaks with signal-to-noise ratio greater than two. We note that the number of signal peaks serves as a measure of the quality of library spectra. Library spectra often contain more fragment ion peaks than the number of signal peaks but these peaks are present at low intensity (i.e., often close to or at the level of noise). In general, peptides identified by MSPLIT-DIA (green, magenta) have better library spectra than those that were not identified by MSPLIT-DIA (blue). Furthermore, identified peptides were further separated into lysate peptides (green) and UPS1 peptides (magenta): UPS1 peptides tend to have better library spectra than lysate peptides identified by MSPLIT-DIA, which could explain why MSPLIT-DIA had better sensitivity at identifying UPS1 peptides than lysate peptides. (b) Number of signal peaks in DIA spectra for identified (green) and unidentified (blue) peptides. (c) Peaks in library spectra but missing in the corresponding DIA spectra were binned by their intensity relative to the most-intense peak in the library spectrum. As expected, unidentified DIA spectra had a substantially higher fraction of missing peaks (irrespective of relative peak intensity) than observed in identified DIA spectra.

Peptide identification with DIA-Umpire

Pseudo-MS/MS spectra generated by DIA-Umpire²⁰ were searched by X!Tandem²¹ and MSGF+²², allowing tryptic peptides with up to one missed cleavage, oxidation of methionine and N-terminal acetylation as variable modifications, and cysteine alkylation as a fixed modification. The precursor-ion mass tolerance was set to 50 ppm and the fragment-ion mass tolerance to 40 ppm. Search results were converted into pepXML format and analyzed with PeptideProphet²³ via the TPP²⁴ for each of the three quality categories of MS/MS spectra. iProphet²⁵ was used to merge search results from X! Tandem and MSGF+. Unique peptide identifications for each DIA run were filtered at 1% FDR, estimated via the target-decoy approach. The number of proteins with at least one unique peptide was determined directly from the PeptideProphet results.

Peptide extraction with PeakView, OpenSWATH and Skyline

Targeted peptide extraction was performed using PeakView^{1, 11}, OpenSWATH²⁶ or Skyline²⁷. To enable comparable analysis of peptide detection in DIA data, all assay libraries were derived from the same peptide spectral libraries that were used for MSPLIT-DIA searches (see above). All unmodified peptides were considered valid candidates for extraction. The five most intense peaks annotated as b or y-ions for each peptide were selected to generate the assay libraries. All assay libraries were generated in a tab-delimited text format appropriate for each tool. We note that for PeakView, the full spectral libraries were imported because PeakView automatically selects the top five fragment ions for extraction.

PeakView extraction was done as described before¹ (using PeakView v.2.1.0.1104 with SWATH microapp version 2.0.0.2003) with 50 ppm mass tolerance within ± 5 min of the assigned retention time (RT). PeakView results were filtered at 1% FDR¹¹. For OpenSWATH, the daily build version of 10/8/2014 was used. Assay libraries were first converted to TraML format using the utility program ConvertTSVToTraML. Decoy assay libraries were generated using the OpenSwathDecoyGenerator program with default parameters. Targeted extraction was done using the OpenSwathWorkflow program with a RT window of ± 5 min and 0.05 Da fragment mass tolerance. Final scores for the extracted peak groups were calculated using pyprophet (v0.11.1) and results were filtered using a 1% FDR. For the UPS and lysate samples analysis with sample-specific

libraries, OpenSWATH was not able to finish processing the data so we were not able to include its results for these samples. The performance of OpenSWATH was benchmarked against other tools using data from the AP-MS samples (see **Supplementary Figure 9**). For Skyline analysis (64bit, v2.6.0.6709), decoy assay libraries were generated using the "Add Decoy Peptides" function and targeted extraction was performed using 0.05 Da fragment mass tolerance and ± 5 min retention time window. The "Train Model" function was used to re-train a new scoring model for each sample and all peak groups were re-scored and re-integrated using the new scoring model. Results were filtered and exported at 1% FDR.

To ensure the same FDR estimation was used for all methods, we determined peptide-level FDR using the Target/Decoy approach (TDA). For each unique peptide sequence, we chose the highest scoring match or peak group. At a particular score threshold T, FDR was computed as the number of decoy library matches divided by the number of target library matches above the score threshold T.

When the assay libraries or spectral libraries are not built from the parallel DDA runs of the same samples as the DIA runs, there is the possibility of retention-time drift between the libraries and the DIA data (e.g., SWATH-Atlas library vs AP-MS samples). In these cases, we performed RT-calibration before doing targeted extraction. However, in order to perform RT-calibration, one usually needs to know *a priori* a set of common peptides that are present in the samples (e.g. iRT peptides²⁸). Since in our case no iRT peptides were added into the sample, we manually selected a set of peptides (twelve for lysate samples and ten for the AP-MS samples) identified by MSPLIT-DIA and used them as calibration peptides. We selected the peptides such that 1) they were identified by MSPLIT-DIA at 1% FDR across all samples of the same type (i.e. on all human samples or on all *E. coli* samples; we generated a set of peptides for each type of samples); 2) they span the entire chromatographic range; 3) they roughly spread out evenly across the chromatographic range and have high *proj-cosine* similarity (i.e. > 0.85). For PeakView, the "Calculate RT Fit" function was used to calculate RT-calibration. For OpenSWATH the selected peptides were input as "iRT" peptides to the using the "-tr_irt options". For Skyline, the selected peptides were input as "Standard peptides" in the iRT calculator and "Retention time predictor" functionality. As shown in **Supplementary Figure 9** by the results labeled with "(MSPLIT-DIA-assisted RT alignment)", retention time calibration using peptides identified by MSPLIT-DIA significantly improved targeted peptide extraction for all tools without the need to spike in additional calibration peptides.

Importing MSPLIT-DIA results for targeted quantification

Since MSPLIT-DIA is designed as a discovery tool to perform peptide identification in DIA data, it is robust to retention time changes between the spectral library and DIA data. As such, peptides identified by MSPLIT-DIA can be used to generate sample-specific assay libraries to perform targeted quantification using tools like PeakView, Skyline or OpenSWATH. Sample-specific libraries reduce the size of the assay libraries input into targeted quantification tools and also provide the actual retention time where the peptides were detected in the

DIA data (thus addressing any potential issues regarding retention-time drift between the peptide in the library and DIA data). MSPLIT-DIA currently exports its results to PeakView, Skyline and OpenSWATH in their text-based format (*i.e.*, ‘.txt’ and ‘.csv’, respectively). For each peptide that was identified at 1% peptide-level FDR, MSPLIT-DIA exports: 1) the retention time at which the best scoring peptide is identified in the DIA run and 2) the fragment mass and intensity information from that peptide from the corresponding DDA library spectrum (default) or from the DIA spectrum where the peptide was identified.

Data and software availability

Mass spectrometry data used were deposited in the MassIVE repository at <http://massive.ucsd.edu/ProteoSAFe/status.jsp?task=a4b32b9ba82e4885a8956b97ca71a1f8>, which can be accessed via FTP with username: MSV000078545 and password: 123456. The MSPLIT-DIA software is available for both download and online access at <http://proteomics.ucsd.edu/software-tools/msplit-dia/>.

References:

1. Lambert, J.P. et al. Mapping differential interactomes by affinity purification coupled with data-independent mass spectrometry acquisition. *Nat Methods* **10**, 1239-1245 (2013).
2. Kim, S. et al. The generating function of CID, ETD, and CID/ETD pairs of tandem mass spectra: applications to database search. *Mol Cell Proteomics* **9**, 2840-2852 (2010).
3. Elias, J.E. & Gygi, S.P. Target-decoy search strategy for increased confidence in large-scale protein identifications by mass spectrometry. *Nat Methods* **4**, 207-214 (2007).
4. Jeong, K., Kim, S. & Bandeira, N. False discovery rates in spectral identification. *BMC Bioinformatics* **13 Suppl 16**, S2 (2012).
5. Lam, H., Deutsch, E.W. & Aebersold, R. Artificial decoy spectral libraries for false discovery rate estimation in spectral library searching in proteomics. *J Proteome Res* **9**, 605-610 (2010).
6. Rosenberger, G. et al. A repository of assays to quantify 10,000 human proteins by SWATH-MS. *Sci Data* **1**, 140031 (2014).
7. Tanner, S. et al. InsPecT: identification of posttranslationally modified peptides from tandem mass spectra. *Anal Chem* **77**, 4626-4639 (2005).
8. Wang, J., Perez-Santiago, J., Katz, J.E., Mallick, P. & Bandeira, N. Peptide identification from mixture tandem mass spectra. *Mol Cell Proteomics* **9**, 1476-1485 (2010).
9. Lam, H. et al. Development and validation of a spectral library searching method for peptide identification from MS/MS. *Proteomics* **7**, 655-667 (2007).
10. Geromanos, S.J. et al. The detection, correlation, and comparison of peptide precursor and product ions from data independent LC-MS with data dependant LC-MS/MS. *Proteomics* **9**, 1683-1695 (2009).
11. Gillet, L.C. et al. Targeted data extraction of the MS/MS spectra generated by data-independent acquisition: a new concept for consistent and accurate proteome analysis. *Mol Cell Proteomics* **11**, O111016717 (2012).
12. Gupta, N. & Pevzner, P.A. False discovery rates of protein identifications: a strike against the two-peptide rule. *J Proteome Res* **8**, 4173-4181 (2009).
13. Shanmugam, A.K., Yocum, A.K. & Nesvizhskii, A.I. Utility of RNA-seq and GPMDB protein observation frequency for improving the sensitivity of protein identification by tandem MS. *J Proteome Res* **13**, 4113-4119 (2014).
14. Teo, G. et al. SAINTexpress: improvements and additional features in Significance Analysis of INTeractome software. *J Proteomics* **100**, 37-43 (2014).
15. Mellacheruvu, D. et al. The CRAPome: a contaminant repository for affinity purification-mass

- spectrometry data. *Nat Methods* **10**, 730-736 (2013).
16. Huang, D.W. et al. DAVID Bioinformatics Resources: expanded annotation database and novel algorithms to better extract biology from large gene lists. *Nucleic Acids Res* **35**, W169-175 (2007).
 17. Ahrne, E., Nikitin, F., Lisacek, F. & Muller, M. QuickMod: A tool for open modification spectrum library searches. *J Proteome Res* **10**, 2913-2921 (2011).
 18. Ye, D. et al. Open MS/MS spectral library search to identify unanticipated post-translational modifications and increase spectral identification rate. *Bioinformatics* **26**, i399-406 (2010).
 19. Fu, Y. et al. DeltAMT: a statistical algorithm for fast detection of protein modifications from LC-MS/MS data. *Mol Cell Proteomics* **10**, M110 000455 (2011).
 20. Tsou, C.C. et al. DIA-Umpire: comprehensive computational framework for data-independent acquisition proteomics. *Nat Methods* **12**, 258-264, 257 p following 264 (2015).
 21. Craig, R. & Beavis, R.C. TANDEM: matching proteins with tandem mass spectra. *Bioinformatics* **20**, 1466-1467 (2004).
 22. Kim, S. & Pevzner, P.A. MS-GF+ makes progress towards a universal database search tool for proteomics. *Nat Commun* **5**, 5277 (2014).
 23. Keller, A., Nesvizhskii, A.I., Kolker, E. & Aebersold, R. Empirical statistical model to estimate the accuracy of peptide identifications made by MS/MS and database search. *Anal Chem* **74**, 5383-5392 (2002).
 24. Deutsch, E.W. et al. A guided tour of the Trans-Proteomic Pipeline. *Proteomics* **10**, 1150-1159 (2010).
 25. Shteynberg, D. et al. iProphet: multi-level integrative analysis of shotgun proteomic data improves peptide and protein identification rates and error estimates. *Mol Cell Proteomics* **10**, M111 007690 (2011).
 26. Rost, H.L. et al. OpenSWATH enables automated, targeted analysis of data-independent acquisition MS data. *Nat Biotechnol* **32**, 219-223 (2014).
 27. MacLean, B. et al. Skyline: an open source document editor for creating and analyzing targeted proteomics experiments. *Bioinformatics* **26**, 966-968 (2010).
 28. Escher, C. et al. Using iRT, a normalized retention time for more targeted measurement of peptides. *Proteomics* **12**, 1111-1121 (2012).

Supplementary Table 1. High confidence protein identifications. Affinity purification coupled to mass spectrometry was performed on biological triplicates of the FLAG-EIF4A2 bait, the FLAG-MEPCE bait and FLAG-GFP as a negative control. Mass spectrometry was acquired using a standard Data Dependent Acquisition (DDA) protocol followed by MSGFDB analysis (blue shading) or by Data Independent Acquisition (SWATH) using MSPLIT-DIA (green shading). SAINTexpress (Teo et al., *J Proteomics*, 2014) was used to score the data based on the spectral counts distribution in the purifications of the bait compared to the negative controls. The averaged SAINT score (AvgP) across all replicates was used for Bayesian FDR (BFDR) modelling; 1% FDR was set as a reporting cutoff for high confidence hits. For each bait-prey interaction, once an interaction passes the confidence threshold in at least one of the acquisition methods, all data pertaining to this bait-prey interaction across DDA and DIA are returned for comparative purposes. The Prey Protein accessions are from Uniprot; the Prey Gene is from HUGO (at NCBI). The number of spectra matched to the prey in each of the biological replicates is indicated in the "Spectra" column ("|" is a delimiter) and the averaged spectral count is shown in the AvgSpec column. The spectra across the 3 GFP negative controls are shown (Control Spectra) alongside the Averaged SAINT score (AvgP), the maximal SAINT score across the 3 replicates (maxP) and the BFDR. The acquisition type (DDA or DIA) is listed and the interactions reaching the confidence threshold are labeled (pink) in the "high confidence" column.

Bait	Prey Protein	Prey Gene	Spectra	AvgSpec	Control Spectra	AvgP	MaxP	BFDR	Acquisition	High Confidence
EIF4A2	ABCE1_HUMAN	ABCE1	17 15 16	16	1 4 4	1	1	0	DDA	1
EIF4A2	ABCE1_HUMAN	ABCE1	61 55 81	65.67	1 2 5	1	1	0	DIA	1
EIF4A2	ATX2_HUMAN	ATXN2	9 8 5	7.33	0 0 0	1	1	0	DDA	1
EIF4A2	ATX2_HUMAN	ATXN2	17 19 24	20	0 0 0	1	1	0	DIA	1
EIF4A2	ATX2L_HUMAN	ATXN2L	32 37 36	35	0 0 1	1	1	0	DDA	1
EIF4A2	ATX2L_HUMAN	ATXN2L	88 92 114	98	1 0 0	1	1	0	DIA	1
EIF4A2	BA2L2_HUMAN	PRRC2C	38 31 43	37.33	1 0 1	1	1	0	DDA	1
EIF4A2	BA2L2_HUMAN	PRRC2C	83 89 94	88.67	0 0 0	1	1	0	DIA	1
EIF4A2	BAT2_HUMAN	PRRC2A	52 48 47	49	0 0 0	1	1	0	DDA	1
EIF4A2	BAT2_HUMAN	PRRC2A	165 154 150	156.33	0 0 0	1	1	0	DIA	1
EIF4A2	EDRF1_HUMAN	EDRF1	6 10 5	7	0 0 0	1	1	0	DDA	1
EIF4A2	EDRF1_HUMAN	EDRF1	23 21 19	21	0 0 0	1	1	0	DIA	1
EIF4A2	CN166_HUMAN	C14orf166	11 15 14	13.33	6 5 5	0.32	0.63	0.68	DDA	
EIF4A2	CN166_HUMAN	C14orf166	60 56 66	60.67	12 11 12	1	1	0	DIA	1
EIF4A2	CSK21_HUMAN	CSNK2A1	14 17 15	15.33	1 1 1	1	1	0	DDA	1
EIF4A2	CSK21_HUMAN	CSNK2A1	33 35 34	34	0 0 0	1	1	0	DIA	1
EIF4A2	CSK22_HUMAN	CSNK2A2	15 16 13	14.67	1 0 1	1	1	0	DDA	1
EIF4A2	CSK22_HUMAN	CSNK2A2	26 28 24	26	2 3 2	1	1	0	DIA	1
EIF4A2	CSK2B_HUMAN	CSNK2B	4 2 4	3.33	0 0 0	0.98	1	0.02	DDA	
EIF4A2	CSK2B_HUMAN	CSNK2B	5 7 5	5.67	0 0 0	1	1	0	DIA	1
EIF4A2	CUL1_HUMAN	CUL1	6 5 4	5	0 0 0	1	1	0	DIA	1
EIF4A2	DAZP1_HUMAN	DAZAP1	3 3 4	3.33	0 0 0	0.99	1	0.01	DDA	1
EIF4A2	DAZP1_HUMAN	DAZAP1	6 13 11	10	0 0 0	1	1	0	DIA	1
EIF4A2	DDX1_HUMAN	DDX1	21 25 27	24.33	8 8 10	0.66	0.99	0.34	DDA	
EIF4A2	DDX1_HUMAN	DDX1	95 87 94	92	11 25 20	1	1	0	DIA	1
EIF4A2	DDX17_HUMAN	DDX17	12 11 17	13.33	3 4 4	0.86	1	0.14	DDA	
EIF4A2	DDX17_HUMAN	DDX17	32 47 42	40.33	12 12 7	1	1	0	DIA	1
EIF4A2	DDX21_HUMAN	DDX21	14 12 13	13	2 3 3	1	1	0	DDA	1
EIF4A2	DDX21_HUMAN	DDX21	31 26 49	35.33	2 0 2	1	1	0	DIA	1
EIF4A2	DDX3X_HUMAN	DDX3X	35 39 40	38	6 9 8	1	1	0	DDA	1
EIF4A2	DDX3X_HUMAN	DDX3X	138 119 145	134	24 27 24	1	1	0	DIA	1
EIF4A2	DDX3Y_HUMAN	DDX3Y	3 4 3	3.33	0 0 0	0.99	1	0.01	DDA	1
EIF4A2	DDX5_HUMAN	DDX5	19 18 20	19	7 6 7	0.74	0.92	0.26	DDA	
EIF4A2	DDX5_HUMAN	DDX5	59 50 74	61	11 11 7	1	1	0	DIA	1
EIF4A2	DDX50_HUMAN	DDX50	7 6 8	7	0 0 0	1	1	0	DDA	1
EIF4A2	DDX50_HUMAN	DDX50	4 5 10	6.33	5 0 0	0.51	0.84	0.49	DIA	
EIF4A2	DHX29_HUMAN	DHX29	4 8 9	7	0 0 0	1	1	0	DDA	1
EIF4A2	DHX29_HUMAN	DHX29	7 4 20	10.33	0 0 0	1	1	0	DIA	1
EIF4A2	EIF3A_HUMAN	EIF3A	232 237 233	234	15 12 12	1	1	0	DDA	1
EIF4A2	EIF3A_HUMAN	EIF3A	970 960 928	952.67	43 35 37	1	1	0	DIA	1
EIF4A2	EIF3B_HUMAN	EIF3B	103 126 110	113	6 4 3	1	1	0	DDA	1
EIF4A2	EIF3B_HUMAN	EIF3B	490 459 460	469.67	15 12 16	1	1	0	DIA	1
EIF4A2	EIF3C_HUMAN	EIF3C	104 105 116	108.33	10 9 7	1	1	0	DDA	1
EIF4A2	EIF3C_HUMAN	EIF3C	463 429 403	431.67	26 26 22	1	1	0	DIA	1
EIF4A2	EIF3D_HUMAN	EIF3D	57 70 50	59	0 1 0	1	1	0	DDA	1
EIF4A2	EIF3D_HUMAN	EIF3D	228 189 194	203.67	4 3 2	1	1	0	DIA	1
EIF4A2	EIF3E_HUMAN	EIF3E	69 59 67	65	5 3 3	1	1	0	DDA	1
EIF4A2	EIF3E_HUMAN	EIF3E	349 373 396	372.67	16 20 14	1	1	0	DIA	1
EIF4A2	EIF3F_HUMAN	EIF3F	86 83 77	82	6 6 11	1	1	0	DDA	1
EIF4A2	EIF3F_HUMAN	EIF3F	364 366 349	359.67	27 25 16	1	1	0	DIA	1
EIF4A2	EIF3G_HUMAN	EIF3G	125 124 123	124	25 28 28	1	1	0	DDA	1
EIF4A2	EIF3G_HUMAN	EIF3G	493 474 496	487.67	85 82 102	1	1	0	DIA	1
EIF4A2	EIF3H_HUMAN	EIF3H	85 90 75	83.33	6 5 8	1	1	0	DDA	1
EIF4A2	EIF3H_HUMAN	EIF3H	342 328 309	326.33	11 10 9	1	1	0	DIA	1
EIF4A2	EIF3I_HUMAN	EIF3I	70 69 70	69.67	13 15 15	1	1	0	DDA	1
EIF4A2	EIF3I_HUMAN	EIF3I	423 399 406	409.33	72 74 80	1	1	0	DIA	1
EIF4A2	EIF3J_HUMAN	EIF3J	30 29 26	28.33	2 2 1	1	1	0	DDA	1
EIF4A2	EIF3J_HUMAN	EIF3J	138 130 126	131.33	8 7 5	1	1	0	DIA	1
EIF4A2	EIF3K_HUMAN	EIF3K	22 28 27	25.67	4 2 3	1	1	0	DDA	1
EIF4A2	EIF3K_HUMAN	EIF3K	108 98 93	99.67	6 5 6	1	1	0	DIA	1
EIF4A2	EIF3L_HUMAN	EIF3L	104 104 95	101	7 9 3	1	1	0	DDA	1
EIF4A2	EIF3L_HUMAN	EIF3L	449 459 443	450.33	19 18 18	1	1	0	DIA	1
EIF4A2	EIF3M_HUMAN	EIF3M	57 58 53	56	2 3 2	1	1	0	DDA	1
EIF4A2	EIF3M_HUMAN	EIF3M	302 267 282	283.67	5 4 2	1	1	0	DIA	1
EIF4A2	IF4A1_HUMAN	EIF4A1	32 41 42	38.33	7 5 6	1	1	0	DDA	1

EIF4A2	IF4A1_HUMAN	EIF4A1	252 219 268	246.33	22 19 24	1	1	0	DIA	1
EIF4A2	IF4A3_HUMAN	EIF4A3	4 4 7	5	1 1 1	0.75	0.98	0.25	DDA	
EIF4A2	IF4A3_HUMAN	EIF4A3	8 7 9	8	0 0 0	1	1	0	DIA	1
EIF4A2	IF4E_HUMAN	EIF4E	10 10 18	12.67	0 0 0	1	1	0	DDA	1
EIF4A2	IF4E_HUMAN	EIF4E	41 51 45	45.67	0 0 0	1	1	0	DIA	1
EIF4A2	IF4E3_HUMAN	EIF4E3	1 1 0	0.67	0 0 0	0	0	1	DDA	
EIF4A2	IF4E3_HUMAN	EIF4E3	6 6 7	6.33	0 0 0	1	1	0	DIA	1
EIF4A2	IF4G1_HUMAN	EIF4G1	290 301 289	293.33	8 10 9	1	1	0	DDA	1
EIF4A2	IF4G1_HUMAN	EIF4G1	1419 1348 1341	1369.33	30 30 20	1	1	0	DIA	1
EIF4A2	IF4G2_HUMAN	EIF4G2	131 143 143	139	0 0 0	1	1	0	DDA	1
EIF4A2	IF4G2_HUMAN	EIF4G2	620 604 599	607.67	0 0 0	1	1	0	DIA	1
EIF4A2	IF4G3_HUMAN	EIF4G3	128 124 128	126.67	1 1 1	1	1	0	DDA	1
EIF4A2	IF4G3_HUMAN	EIF4G3	534 497 527	519.33	5 2 5	1	1	0	DIA	1
EIF4A2	ELAV1_HUMAN	ELAVL1	11 8 10	9.67	0 0 0	1	1	0	DDA	1
EIF4A2	ELAV1_HUMAN	ELAVL1	38 43 48	43	0 0 0	1	1	0	DIA	1
EIF4A2	FA98A_HUMAN	FAM98A	6 6 7	6.33	0 0 0	1	1	0	DDA	1
EIF4A2	FA98A_HUMAN	FAM98A	29 19 34	27.33	0 1 2	1	1	0	DIA	1
EIF4A2	FMR1_HUMAN	FMR1	7 8 11	8.67	0 0 0	1	1	0	DDA	1
EIF4A2	FMR1_HUMAN	FMR1	28 19 32	26.33	0 0 0	1	1	0	DIA	1
EIF4A2	FXR1_HUMAN	FXR1	5 8 7	6.67	0 1 1	0.98	1	0.02	DDA	
EIF4A2	FXR1_HUMAN	FXR1	15 10 21	15.33	2 0 0	1	1	0	DIA	1
EIF4A2	FXR2_HUMAN	FXR2	6 11 10	9	0 0 0	1	1	0	DDA	1
EIF4A2	FXR2_HUMAN	FXR2	26 12 24	20.67	1 0 5	0.96	1	0.04	DIA	
EIF4A2	G3BP2_HUMAN	G3BP2	7 6 11	8	0 0 0	1	1	0	DDA	1
EIF4A2	G3BP2_HUMAN	G3BP2	27 29 36	30.67	0 0 0	1	1	0	DIA	1
EIF4A2	GBLP_HUMAN	GNB2L1	18 16 20	18	4 6 8	0.76	0.99	0.24	DDA	
EIF4A2	GBLP_HUMAN	GNB2L1	62 62 77	67	15 18 15	1	1	0	DIA	1
EIF4A2	HELC1_HUMAN	HELC1_HUMAN	10 10 13	11	0 0 0	1	1	0	DDA	1
EIF4A2	HELC1_HUMAN	HELC1_HUMAN	25 22 15	20.67	0 2 0	1	1	0	DIA	1
EIF4A2	ROA2_HUMAN	HNRNPA2B1	32 31 40	34.33	8 6 8	1	1	0	DDA	1
EIF4A2	ROA2_HUMAN	HNRNPA2B1	150 143 162	151.67	46 38 38	1	1	0	DIA	1
EIF4A2	ROA3_HUMAN	HNRNPA3	8 5 5	6	1 1 0	0.97	1	0.03	DDA	
EIF4A2	ROA3_HUMAN	HNRNPA3	34 32 38	34.67	7 1 3	1	1	0	DIA	1
EIF4A2	ROAA_HUMAN	HNRNPAB	8 5 8	7	2 2 2	0.67	0.9	0.33	DDA	
EIF4A2	ROAA_HUMAN	HNRNPAB	37 29 42	36	4 7 4	1	1	0	DIA	1
EIF4A2	HNRDL_HUMAN	HNRNPDL	3 4 3	3.33	0 0 0	0.99	1	0.01	DDA	1
EIF4A2	HNRDL_HUMAN	HNRNPDL	4 10 15	9.67	4 1 3	0.65	1	0.35	DIA	
EIF4A2	HNRPL_HUMAN	HNRNPL	13 16 15	14.67	2 1 1	1	1	0	DDA	1
EIF4A2	HNRPL_HUMAN	HNRNPL	62 47 59	56	0 0 0	1	1	0	DIA	1
EIF4A2	HNRPM_HUMAN	HNRNPM	39 23 42	34.67	5 5 3	1	1	0	DDA	1
EIF4A2	HNRPM_HUMAN	HNRNPM	81 122 158	120.33	4 12 8	1	1	0	DIA	1
EIF4A2	HNRPR_HUMAN	HNRNPR	12 13 15	13.33	2 3 2	1	1	0	DDA	1
EIF4A2	HNRPR_HUMAN	HNRNPR	37 23 27	29	1 6 3	1	1	0	DIA	1
EIF4A2	IF2B1_HUMAN	IGF2BP1	23 24 26	24.33	0 0 1	1	1	0	DDA	1
EIF4A2	IF2B1_HUMAN	IGF2BP1	86 86 95	89	1 0 0	1	1	0	DIA	1
EIF4A2	IF2B2_HUMAN	IGF2BP2	13 14 16	14.33	0 0 0	1	1	0	DDA	1
EIF4A2	IF2B2_HUMAN	IGF2BP2	53 41 48	47.33	0 0 0	1	1	0	DIA	1
EIF4A2	IF2B3_HUMAN	IGF2BP3	9 9 11	9.67	0 0 0	1	1	0	DDA	1
EIF4A2	IF2B3_HUMAN	IGF2BP3	35 24 34	31	0 0 0	1	1	0	DIA	1
EIF4A2	ILF2_HUMAN	ILF2	13 14 15	14	1 1 1	1	1	0	DDA	1
EIF4A2	ILF2_HUMAN	ILF2	20 14 17	17	0 0 0	1	1	0	DIA	1
EIF4A2	ILF3_HUMAN	ILF3	13 16 17	15.33	0 2 1	1	1	0	DDA	1
EIF4A2	ILF3_HUMAN	ILF3	49 43 52	48	5 6 3	1	1	0	DIA	1
EIF4A2	LARP1_HUMAN	LARP1	12 14 14	13.33	0 0 0	1	1	0	DDA	1
EIF4A2	LARP1_HUMAN	LARP1	35 33 34	34	0 3 0	1	1	0	DIA	1
EIF4A2	LARP4_HUMAN	LARP4	3 4 3	3.33	0 0 0	0.99	1	0.01	DDA	1
EIF4A2	LARP4_HUMAN	LARP4	0 1 0	0.33	0 0 0	0	0	1	DIA	
EIF4A2	LAR4B_HUMAN	LARP4B	2 4 5	3.67	0 0 0	0.98	1	0.02	DDA	
EIF4A2	LAR4B_HUMAN	LARP4B	20 8 24	17.33	0 0 0	1	1	0	DIA	1
EIF4A2	LIGA_HUMAN	EIF2D	3 5 5	4.33	0 0 0	1	1	0	DDA	1
EIF4A2	LIGA_HUMAN	EIF2D	6 3 17	8.67	0 2 0	0.83	1	0.17	DIA	
EIF4A2	LRC47_HUMAN	LRRC47	2 3 2	2.33	2 2 3	0	0.01	1	DDA	
EIF4A2	LRC47_HUMAN	LRRC47	5 4 4	4.33	0 0 0	1	1	0	DIA	1
EIF4A2	LSM12_HUMAN	LSM12	10 14 10	11.33	0 0 0	1	1	0	DDA	1
EIF4A2	LSM12_HUMAN	LSM12	28 25 24	25.67	0 0 0	1	1	0	DIA	1
EIF4A2	MATR3_HUMAN	MATR3	20 23 27	23.33	3 1 1	1	1	0	DDA	1
EIF4A2	MATR3_HUMAN	MATR3	64 67 91	74	0 1 0	1	1	0	DIA	1
EIF4A2	NONO_HUMAN	NONO	42 34 64	46.67	8 7 6	1	1	0	DDA	1
EIF4A2	NONO_HUMAN	NONO	188 238 350	258.67	55 49 44	1	1	0	DIA	1
EIF4A2	NUFP2_HUMAN	NUFIP2	15 17 19	17	2 1 0	1	1	0	DDA	1
EIF4A2	NUFP2_HUMAN	NUFIP2	30 22 37	29.67	4 3 0	1	1	0	DIA	1
EIF4A2	OBSL1_HUMAN	OBSL1	2 2 1	1.67	0 0 0	0.85	0.94	0.15	DDA	
EIF4A2	OBSL1_HUMAN	OBSL1	6 6 4	5.33	0 0 0	1	1	0	DIA	1
EIF4A2	PABP1_HUMAN	PABPC1	12 13 13	12.67	0 0 1	1	1	0	DDA	1
EIF4A2	PABP1_HUMAN	PABPC1	63 57 58	59.33	0 0 0	1	1	0	DIA	1
EIF4A2	PABP4_HUMAN	PABPC4	32 38 42	37.33	0 0 0	1	1	0	DDA	1
EIF4A2	PABP4_HUMAN	PABPC4	113 103 116	110.67	0 0 0	1	1	0	DIA	1
EIF4A2	PDCD4_HUMAN	PDCD4	202 174 216	197.33	0 0 0	1	1	0	DDA	1

EIF4A2	PDCD4_HUMAN	PDCD4	844 1053 1061	986	0 3 1	1	1	0	DIA	1
EIF4A2	PSPC1_HUMAN	PSPC1	6 3 8	5.67	1 2 4	0.34	0.78	0.66	DDA	1
EIF4A2	PSPC1_HUMAN	PSPC1	5 17 18	13.33	1 0 0	1	1	0	DIA	1
EIF4A2	PTBP1_HUMAN	PTBP1	19 25 24	22.67	1 2 4	1	1	0	DDA	1
EIF4A2	PTBP1_HUMAN	PTBP1	112 105 108	108.33	3 0 2	1	1	0	DIA	1
EIF4A2	PTBP2_HUMAN	PTBP2	2 3 2	2.33	0 0 0	0.95	0.99	0.05	DDA	1
EIF4A2	PTBP2_HUMAN	PTBP2	8 8 8	8	0 0 0	1	1	0	DIA	1
EIF4A2	PURA_HUMAN	PURA	1 1 1	1	0 0 0	0	0	1	DDA	1
EIF4A2	PURA_HUMAN	PURA	9 10 12	10.33	0 0 0	1	1	0	DIA	1
EIF4A2	RL15_HUMAN	RPL15	1 1 1	1	0 0 0	0	0	1	DDA	1
EIF4A2	RL15_HUMAN	RPL15	6 7 6	6.33	0 0 0	1	1	0	DIA	1
EIF4A2	RL18_HUMAN	RPL18	2 2 4	2.67	1 0 0	0.76	0.99	0.24	DDA	1
EIF4A2	RL18_HUMAN	RPL18	4 3 4	3.67	0 0 0	1	1	0	DIA	1
EIF4A2	RL27_HUMAN	RPL27	2 1 2	1.67	1 1 1	0.08	0.12	0.92	DDA	1
EIF4A2	RL27_HUMAN	RPL27	5 6 4	5	0 0 0	1	1	0	DIA	1
EIF4A2	RL4_HUMAN	RPL4	2 2 5	3	1 0 1	0.51	0.95	0.49	DDA	1
EIF4A2	RL4_HUMAN	RPL4	10 10 12	10.67	0 1 0	1	1	0	DIA	1
EIF4A2	RLA1_HUMAN	RPLP1	1 1 2	1.33	0 1 0	0.36	0.64	0.64	DDA	1
EIF4A2	RLA1_HUMAN	RPLP1	7 5 7	6.33	0 0 0	1	1	0	DIA	1
EIF4A2	RS13_HUMAN	RPS13	14 13 13	13.33	7 5 6	0.03	0.05	0.97	DDA	1
EIF4A2	RS13_HUMAN	RPS13	65 61 67	64.33	19 21 26	1	1	0	DIA	1
EIF4A2	RS23_HUMAN	RPS23	8 9 8	8.33	3 2 3	0.64	0.81	0.36	DDA	1
EIF4A2	RS23_HUMAN	RPS23	40 44 41	41.67	8 10 15	1	1	0	DIA	1
EIF4A2	RS7_HUMAN	RPS7	18 20 19	19	10 13 9	0	0	1	DDA	1
EIF4A2	RS7_HUMAN	RPS7	133 117 133	127.67	44 41 41	1	1	0	DIA	1
EIF4A2	RS9_HUMAN	RPS9	15 9 12	12	5 6 7	0.06	0.17	0.94	DDA	1
EIF4A2	RS9_HUMAN	RPS9	36 41 40	39	12 13 12	1	1	0	DIA	1
EIF4A2	SAFB1_HUMAN	SAFB	9 10 11	10	0 0 0	1	1	0	DIA	1
EIF4A2	SFPQ_HUMAN	SFPQ	38 38 59	45	11 10 13	1	1	0	DDA	1
EIF4A2	SFPQ_HUMAN	SFPQ	150 187 268	201.67	39 23 26	1	1	0	DIA	1
EIF4A2	SRSF1_HUMAN	SRSF1	11 8 12	10.33	2 3 1	0.97	1	0.03	DDA	1
EIF4A2	SRSF1_HUMAN	SRSF1	16 23 33	24	5 4 2	1	1	0	DIA	1
EIF4A2	SRSF7_HUMAN	SRSF7	3 5 4	4	0 0 0	1	1	0	DDA	1
EIF4A2	SRSF7_HUMAN	SRSF7	10 25 17	17.33	0 0 0	1	1	0	DIA	1
EIF4A2	HNRPO_HUMAN	SYNCRIP	26 30 29	28.33	10 14 13	0.02	0.05	0.98	DDA	1
EIF4A2	HNRPO_HUMAN	SYNCRIP	75 70 77	74	11 11 16	1	1	0	DIA	1
EIF4A2	TADBP_HUMAN	TARDBP	3 5 4	4	1 3 2	0.09	0.2	0.91	DDA	1
EIF4A2	TADBP_HUMAN	TARDBP	7 8 20	11.67	1 2 0	0.99	1	0.01	DIA	1
EIF4A2	UAP56_HUMAN	DDX39B	0 2 0	0.67	0 0 0	0.31	0.94	0.69	DDA	1
EIF4A2	UAP56_HUMAN	DDX39B	9 10 11	10	0 0 0	1	1	0	DIA	1
EIF4A2	RENT1_HUMAN	UPF1	5 4 7	5.33	0 0 0	1	1	0	DDA	1
EIF4A2	RENT1_HUMAN	UPF1	4 5 6	5	0 0 0	1	1	0	DIA	1
EIF4A2	UBP10_HUMAN	USP10	7 6 9	7.33	0 0 0	1	1	0	DDA	1
EIF4A2	UBP10_HUMAN	USP10	20 20 36	25.33	1 0 0	1	1	0	DIA	1
MEPCE	AFF4_HUMAN	AFF4	8 7 9	8	0 0 0	1	1	0	DDA	1
MEPCE	AFF4_HUMAN	AFF4	13 8 18	13	0 0 0	1	1	0	DIA	1
MEPCE	C1QBP_HUMAN	C1QBP	371 337 335	347.67	39 56 46	1	1	0	DDA	1
MEPCE	C1QBP_HUMAN	C1QBP	2351 1809 2333	2164.33	183 177 210	1	1	0	DIA	1
MEPCE	CCNT1_HUMAN	CCNT1	22 20 20	20.67	0 0 0	1	1	0	DDA	1
MEPCE	CCNT1_HUMAN	CCNT1	84 79 82	81.67	0 1 0	1	1	0	DIA	1
MEPCE	CCNT2_HUMAN	CCNT2	9 9 6	8	0 0 0	1	1	0	DDA	1
MEPCE	CCNT2_HUMAN	CCNT2	21 17 28	22	0 0 0	1	1	0	DIA	1
MEPCE	CDK9_HUMAN	CDK9	29 28 27	28	0 0 0	1	1	0	DDA	1
MEPCE	CDK9_HUMAN	CDK9	106 113 115	111.33	4 1 2	1	1	0	DIA	1
MEPCE	CHCH2_HUMAN	CHCH2	3 4 4	3.67	0 0 0	1	1	0	DDA	1
MEPCE	CHCH2_HUMAN	CHCH2	12 23 26	20.33	0 0 0	1	1	0	DIA	1
MEPCE	NEUA_HUMAN	CMAS	19 22 26	22.33	0 1 1	1	1	0	DDA	1
MEPCE	NEUA_HUMAN	CMAS	68 60 84	70.67	0 0 0	1	1	0	DIA	1
MEPCE	CU057_HUMAN	YBEY	1 1 2	1.33	0 0 0	0.75	0.94	0.25	DDA	1
MEPCE	CU057_HUMAN	YBEY	19 16 19	18	0 0 0	1	1	0	DIA	1
MEPCE	DDX21_HUMAN	DDX21	20 22 22	21.33	2 3 3	1	1	0	DDA	1
MEPCE	DDX21_HUMAN	DDX21	82 63 76	73.67	2 0 2	1	1	0	DIA	1
MEPCE	DDX23_HUMAN	DDX23	16 20 22	19.33	3 1 3	1	1	0	DDA	1
MEPCE	DDX23_HUMAN	DDX23	48 46 50	48	17 12 9	1	1	0	DIA	1
MEPCE	DDX5_HUMAN	DDX5	13 16 13	14	7 6 7	0.03	0.08	0.97	DDA	1
MEPCE	DDX5_HUMAN	DDX5	30 31 37	32.67	11 11 7	1	1	0	DIA	1
MEPCE	DDX56_HUMAN	DDX56	2 1 3	2	0 0 0	0.86	0.99	0.14	DDA	1
MEPCE	DDX56_HUMAN	DDX56	4 5 4	4.33	1 0 0	1	1	0	DIA	1
MEPCE	DHX9_HUMAN	DHX9	41 44 48	44.33	12 10 13	1	1	0	DDA	1
MEPCE	DHX9_HUMAN	DHX9	151 148 167	155.33	45 36 38	1	1	0	DIA	1
MEPCE	EDC4_HUMAN	EDC4	3 12 14	9.67	1 0 0	0.98	1	0.02	DDA	1
MEPCE	EDC4_HUMAN	EDC4	4 24 41	23	0 0 0	1	1	0	DIA	1
MEPCE	HME1_HUMAN	EN1	4 4 4	4	0 0 0	1	1	0	DIA	1
MEPCE	NUCG_HUMAN	ENDOG	33 28 32	31	0 0 0	1	1	0	DDA	1
MEPCE	NUCG_HUMAN	ENDOG	115 132 141	129.33	0 0 0	1	1	0	DIA	1
MEPCE	BGAL_HUMAN	GLB1	3 3 3	3	0 0 0	0.99	0.99	0.01	DDA	1
MEPCE	BGAL_HUMAN	GLB1	2 5 7	4.67	0 0 0	0.98	1	0.02	DIA	1
MEPCE	HEX1_HUMAN	HEXIM1	12 11 11	11.33	0 0 0	1	1	0	DDA	1

MEPCE	HEX1_HUMAN	HEXIM1	47 47 50	48	0 1 0	1	1	0	DIA	1
MEPCE	HMX3_HUMAN	HMX3	2 2 1	1.67	0 0 0	0.85	0.94	0.15	DDA	
MEPCE	HMX3_HUMAN	HMX3	9 9 6	8	0 0 0	1	1	0	DIA	1
MEPCE	ROA2_HUMAN	HNRNPA2B1	24 24 29	25.67	8 6 8	1	1	0	DDA	1
MEPCE	ROA2_HUMAN	HNRNPA2B1	127 110 128	121.67	46 38 38	1	1	0	DIA	1
MEPCE	ROAA_HUMAN	HNRNPAB	3 4 6	4.33	2 2 2	0.18	0.47	0.82	DDA	
MEPCE	ROAA_HUMAN	HNRNPAB	18 23 40	27	4 7 4	1	1	0	DIA	1
MEPCE	HNRPM_HUMAN	HNRNPM	28 28 31	29	5 5 3	1	1	0	DDA	1
MEPCE	HNRPM_HUMAN	HNRNPM	86 49 70	68.33	4 12 8	1	1	0	DIA	1
MEPCE	HNRPR_HUMAN	HNRNPR	24 24 23	23.67	2 3 2	1	1	0	DDA	1
MEPCE	HNRPR_HUMAN	HNRNPR	46 46 56	49.33	1 6 3	1	1	0	DIA	1
MEPCE	HNRL1_HUMAN	HNRNPUL1	9 14 13	12	1 0 1	1	1	0	DDA	1
MEPCE	HNRL1_HUMAN	HNRNPUL1	34 26 34	31.33	2 5 5	1	1	0	DIA	1
MEPCE	HP1B3_HUMAN	HP1BP3	2 3 3	2.67	0 0 1	0.84	0.94	0.16	DDA	
MEPCE	HP1B3_HUMAN	HP1BP3	6 5 8	6.33	0 0 0	1	1	0	DIA	1
MEPCE	ILF2_HUMAN	ILF2	8 8 8	8	1 1 1	1	1	0	DDA	1
MEPCE	ILF2_HUMAN	ILF2	21 14 26	20.33	0 0 0	1	1	0	DIA	1
MEPCE	ILF3_HUMAN	ILF3	8 13 13	11.33	0 2 1	1	1	0	DDA	1
MEPCE	ILF3_HUMAN	ILF3	28 27 34	29.67	5 6 3	1	1	0	DIA	1
MEPCE	IMA2_HUMAN	KPNA2	56 47 49	50.67	5 6 7	1	1	0	DDA	1
MEPCE	IMA2_HUMAN	KPNA2	183 194 194	190.33	14 22 18	1	1	0	DIA	1
MEPCE	IMA4_HUMAN	KPNA3	8 4 5	5.67	0 0 1	1	1	0	DDA	1
MEPCE	IMA4_HUMAN	KPNA3	27 14 34	25	0 0 2	1	1	0	DIA	1
MEPCE	IMA3_HUMAN	KPNA4	5 8 4	5.67	0 0 1	1	1	0	DDA	1
MEPCE	IMA3_HUMAN	KPNA4	26 25 35	28.67	0 0 0	1	1	0	DIA	1
MEPCE	LARP7_HUMAN	LARP7	121 114 122	119	1 1 1	1	1	0	DDA	1
MEPCE	LARP7_HUMAN	LARP7	566 582 619	589	1 1 0	1	1	0	DIA	1
MEPCE	LBR_HUMAN	LBR	4 7 5	5.33	0 0 0	1	1	0	DDA	1
MEPCE	LBR_HUMAN	LBR	11 4 10	8.33	5 2 1	0.59	0.92	0.41	DIA	
MEPCE	LSM2_HUMAN	LSM2	5 6 6	5.67	0 1 1	0.97	0.99	0.03	DDA	
MEPCE	LSM2_HUMAN	LSM2	41 35 34	36.67	3 3 4	1	1	0	DIA	1
MEPCE	LSM5_HUMAN	LSM5	0 1 1	0.67	0 0 0	0	0	1	DDA	
MEPCE	LSM5_HUMAN	LSM5	6 5 7	6	0 0 0	1	1	0	DIA	1
MEPCE	LSM6_HUMAN	LSM6	5 4 7	5.33	0 0 0	1	1	0	DDA	1
MEPCE	LSM6_HUMAN	LSM6	8 10 14	10.67	0 0 0	1	1	0	DIA	1
MEPCE	MATR3_HUMAN	MATR3	5 9 4	6	3 1 1	0.53	0.99	0.47	DDA	
MEPCE	MATR3_HUMAN	MATR3	4 7 9	6.67	0 1 0	1	1	0	DIA	1
MEPCE	MET10_HUMAN	METTL10	35 39 38	37.33	0 0 0	1	1	0	DDA	1
MEPCE	MET10_HUMAN	METTL10	78 80 105	87.67	0 0 0	1	1	0	DIA	1
MEPCE	MSH2_HUMAN	MSH2	2 2 2	2	2 1 1	0.04	0.04	0.96	DDA	
MEPCE	MSH2_HUMAN	MSH2	7 3 6	5.33	0 0 0	1	1	0	DIA	1
MEPCE	NAA38_HUMAN	NAA38	17 11 14	14	2 3 2	1	1	0	DDA	1
MEPCE	NAA38_HUMAN	NAA38	52 48 59	53	4 5 3	1	1	0	DIA	1
MEPCE	NUCL_HUMAN	NCL	56 45 47	49.33	16 16 14	1	1	0	DDA	1
MEPCE	NUCL_HUMAN	NCL	137 147 165	149.67	43 45 45	1	1	0	DIA	1
MEPCE	NH2L1_HUMAN	NHP2L1	2 3 4	3	0 0 0	0.98	1	0.02	DDA	
MEPCE	NH2L1_HUMAN	NHP2L1	18 11 26	18.33	0 0 0	1	1	0	DIA	1
MEPCE	NPM_HUMAN	NPM1	16 17 20	17.67	5 4 4	1	1	0	DDA	1
MEPCE	NPM_HUMAN	NPM1	81 75 83	79.67	18 11 11	1	1	0	DIA	1
MEPCE	PARP1_HUMAN	PARP1	24 22 15	20.33	6 10 8	0.55	0.96	0.45	DDA	
MEPCE	PARP1_HUMAN	PARP1	41 53 52	48.67	16 8 10	1	1	0	DIA	1
MEPCE	PDIP2_HUMAN	POLDIP2	7 5 8	6.67	1 0 0	1	1	0	DDA	1
MEPCE	PDIP2_HUMAN	POLDIP2	13 8 16	12.33	7 3 0	0.58	0.89	0.42	DIA	
MEPCE	PECI_HUMAN	ECI2	3 4 7	4.67	0 0 1	0.98	1	0.02	DDA	
MEPCE	PECI_HUMAN	ECI2	4 4 8	5.33	0 0 0	1	1	0	DIA	1
MEPCE	PPIH_HUMAN	PPIH	14 12 20	15.33	0 0 0	1	1	0	DDA	1
MEPCE	PPIH_HUMAN	PPIH	66 62 72	66.67	0 0 0	1	1	0	DIA	1
MEPCE	PRPF3_HUMAN	PRPF3	39 38 48	41.67	1 1 0	1	1	0	DDA	1
MEPCE	PRPF3_HUMAN	PRPF3	160 148 173	160.33	1 2 0	1	1	0	DIA	1
MEPCE	PRP31_HUMAN	PRPF31	59 49 54	54	15 13 14	1	1	0	DDA	1
MEPCE	PRP31_HUMAN	PRPF31	204 187 210	200.33	41 46 43	1	1	0	DIA	1
MEPCE	PRP4_HUMAN	PRPF4	45 40 40	41.67	0 1 0	1	1	0	DDA	1
MEPCE	PRP4_HUMAN	PRPF4	144 154 177	158.33	1 1 0	1	1	0	DIA	1
MEPCE	PRP6_HUMAN	PRPF6	23 28 27	26	1 5 2	1	1	0	DDA	1
MEPCE	PRP6_HUMAN	PRPF6	84 74 81	79.67	16 11 9	1	1	0	DIA	1
MEPCE	RL12_HUMAN	RPL12	6 6 5	5.67	1 1 1	0.91	0.94	0.09	DDA	
MEPCE	RL12_HUMAN	RPL12	28 32 28	29.33	4 4 6	1	1	0	DIA	1
MEPCE	RL13_HUMAN	RPL13	13 13 16	14	1 2 2	1	1	0	DDA	1
MEPCE	RL13_HUMAN	RPL13	58 55 59	57.33	1 7 2	1	1	0	DIA	1
MEPCE	RL14_HUMAN	RPL14	4 4 5	4.33	0 0 1	1	1	0	DDA	1
MEPCE	RL14_HUMAN	RPL14	30 31 28	29.67	1 0 2	1	1	0	DIA	1
MEPCE	RL15_HUMAN	RPL15	4 4 5	4.33	0 0 0	1	1	0	DDA	1
MEPCE	RL15_HUMAN	RPL15	21 14 22	19	0 0 0	1	1	0	DIA	1
MEPCE	RL18_HUMAN	RPL18	7 7 8	7.33	1 0 0	1	1	0	DDA	1
MEPCE	RL18_HUMAN	RPL18	17 21 18	18.67	0 0 0	1	1	0	DIA	1
MEPCE	RL18A_HUMAN	RPL18A	4 3 4	3.67	0 0 0	1	1	0	DDA	1
MEPCE	RL21_HUMAN	RPL21	4 2 2	2.67	0 0 0	0.96	1	0.04	DDA	
MEPCE	RL21_HUMAN	RPL21	13 11 12	12	0 0 0	1	1	0	DIA	1

MEPCE	RL27_HUMAN	RPL27	8 7 9	8	1 1 1	0.99	1	0.01	DDA	1
MEPCE	RL27_HUMAN	RPL27	22 20 22	21.33	0 0 0	1	1	0	DIA	1
MEPCE	RL28_HUMAN	RPL28	6 5 6	5.67	1 0 0	1	1	0	DDA	1
MEPCE	RL28_HUMAN	RPL28	14 12 12	12.67	0 0 0	1	1	0	DIA	1
MEPCE	RL29_HUMAN	RPL29	4 4 2	3.33	0 0 0	0.98	1	0.02	DDA	
MEPCE	RL29_HUMAN	RPL29	18 19 19	18.67	5 6 5	0.99	1	0.01	DIA	1
MEPCE	RL3_HUMAN	RPL3	12 17 18	15.67	0 0 0	1	1	0	DDA	1
MEPCE	RL3_HUMAN	RPL3	52 52 63	55.67	0 0 0	1	1	0	DIA	1
MEPCE	RL30_HUMAN	RPL30	2 4 2	2.67	0 0 0	0.96	1	0.04	DDA	
MEPCE	RL30_HUMAN	RPL30	14 8 10	10.67	0 0 0	1	1	0	DIA	1
MEPCE	RL34_HUMAN	RPL34	5 3 5	4.33	0 0 0	1	1	0	DDA	1
MEPCE	RL34_HUMAN	RPL34	11 8 11	10	0 0 0	1	1	0	DIA	1
MEPCE	RL36_HUMAN	RPL36	3 1 3	2.33	0 0 0	0.88	0.99	0.12	DDA	
MEPCE	RL36_HUMAN	RPL36	14 11 13	12.67	0 0 0	1	1	0	DIA	1
MEPCE	RL37_HUMAN	RPL37	1 1 1	1	0 0 1	0	0	1	DDA	
MEPCE	RL37_HUMAN	RPL37	28 24 22	24.67	0 0 0	1	1	0	DIA	1
MEPCE	RL37A_HUMAN	RPL37A	3 3 2	2.67	0 0 0	0.97	0.99	0.03	DDA	
MEPCE	RL37A_HUMAN	RPL37A	9 7 6	7.33	0 0 0	1	1	0	DIA	1
MEPCE	RL4_HUMAN	RPL4	20 19 18	19	1 0 1	1	1	0	DDA	1
MEPCE	RL4_HUMAN	RPL4	107 111 116	111.33	0 1 0	1	1	0	DIA	1
MEPCE	RL5_HUMAN	RPL5	11 12 8	10.33	2 0 2	1	1	0	DDA	1
MEPCE	RL5_HUMAN	RPL5	26 25 28	26.33	3 3 0	1	1	0	DIA	1
MEPCE	RL6_HUMAN	RPL6	13 14 11	12.67	0 0 1	1	1	0	DDA	1
MEPCE	RL6_HUMAN	RPL6	47 41 49	45.67	1 0 1	1	1	0	DIA	1
MEPCE	RL7_HUMAN	RPL7	10 12 12	11.33	0 0 0	1	1	0	DDA	1
MEPCE	RL7_HUMAN	RPL7	51 36 47	44.67	0 0 0	1	1	0	DIA	1
MEPCE	RL7A_HUMAN	RPL7A	12 10 15	12.33	0 0 0	1	1	0	DDA	1
MEPCE	RL7A_HUMAN	RPL7A	43 39 47	43	2 3 2	1	1	0	DIA	1
MEPCE	RL8_HUMAN	RPL8	5 5 5	5	0 0 0	1	1	0	DDA	1
MEPCE	RL8_HUMAN	RPL8	27 25 22	24.67	0 0 0	1	1	0	DIA	1
MEPCE	RL9_HUMAN	RPL9	12 11 12	11.67	2 1 2	1	1	0	DDA	1
MEPCE	RL9_HUMAN	RPL9	35 22 40	32.33	2 3 1	1	1	0	DIA	1
MEPCE	RLA0_HUMAN	RPLP0	7 12 11	10	0 0 1	1	1	0	DDA	1
MEPCE	RLA0_HUMAN	RPLP0	22 24 38	28	0 0 0	1	1	0	DIA	1
MEPCE	RLA1_HUMAN	RPLP1	6 6 5	5.67	0 1 0	1	1	0	DDA	1
MEPCE	RLA1_HUMAN	RPLP1	35 39 44	39.33	0 0 0	1	1	0	DIA	1
MEPCE	RLA2_HUMAN	RPLP2	11 16 15	14	4 3 3	0.95	1	0.05	DDA	
MEPCE	RLA2_HUMAN	RPLP2	93 104 113	103.33	6 8 11	1	1	0	DIA	1
MEPCE	RRP1B_HUMAN	RRP1B	3 2 3	2.67	0 0 0	0.97	0.99	0.03	DDA	
MEPCE	RRP1B_HUMAN	RRP1B	12 8 11	10.33	0 0 0	1	1	0	DIA	1
MEPCE	RL1D1_HUMAN	RSL1D1	3 3 3	3	0 0 0	0.99	0.99	0.01	DDA	1
MEPCE	RL1D1_HUMAN	RSL1D1	5 11 9	8.33	1 8 6	0.01	0.02	0.99	DIA	
MEPCE	SNUT1_HUMAN	SART1	21 27 22	23.33	2 1 3	1	1	0	DDA	1
MEPCE	SNUT1_HUMAN	SART1	81 63 95	79.67	6 3 11	1	1	0	DIA	1
MEPCE	SART3_HUMAN	SART3	137 136 132	135	0 0 0	1	1	0	DDA	1
MEPCE	SART3_HUMAN	SART3	600 611 686	632.33	1 3 1	1	1	0	DIA	1
MEPCE	SET_HUMAN	SET	13 18 21	17.33	4 2 4	0.99	1	0.01	DDA	1
MEPCE	SET_HUMAN	SET	66 70 73	69.67	17 12 17	1	1	0	DIA	1
MEPCE	US20_HUMAN	SNRNP200	30 30 41	33.67	15 16 17	0.1	0.29	0.9	DDA	
MEPCE	US20_HUMAN	SNRNP200	75 67 105	82.33	29 23 22	0.99	1	0.01	DIA	1
MEPCE	SNR40_HUMAN	SNRNP40	1 2 3	2	0 0 0	0.86	0.99	0.14	DDA	
MEPCE	SNR40_HUMAN	SNRNP40	11 6 8	8.33	0 1 0	1	1	0	DIA	1
MEPCE	SRSF7_HUMAN	SRSF7	1 2 3	2	0 0 0	0.86	0.99	0.14	DDA	
MEPCE	SRSF7_HUMAN	SRSF7	5 3 6	4.67	0 0 0	1	1	0	DIA	1
MEPCE	LA_HUMAN	SSB	66 70 66	67.33	11 10 10	1	1	0	DDA	1
MEPCE	LA_HUMAN	SSB	348 356 388	364	42 39 36	1	1	0	DIA	1
MEPCE	SSBP_HUMAN	SSBP1	4 5 6	5	0 0 3	0.67	0.78	0.33	DDA	
MEPCE	SSBP_HUMAN	SSBP1	7 8 16	10.33	0 0 0	1	1	0	DIA	1
MEPCE	HNRPO_HUMAN	SYNCRIP	39 35 33	35.67	10 14 13	0.88	1	0.12	DDA	
MEPCE	HNRPO_HUMAN	SYNCRIP	92 104 118	104.67	11 11 16	1	1	0	DIA	1
MEPCE	TMCO6_HUMAN	TMCO6	3 3 6	4	0 0 0	0.99	1	0.01	DDA	1
MEPCE	TMCO6_HUMAN	TMCO6	13 11 27	17	0 0 0	1	1	0	DIA	1
MEPCE	TUT1_HUMAN	TUT1	0 1 2	1	0 0 0	0.53	0.94	0.47	DDA	
MEPCE	TUT1_HUMAN	TUT1	5 3 6	4.67	0 0 0	1	1	0	DIA	1
MEPCE	SNUT2_HUMAN	USP39	5 11 7	7.67	1 0 0	1	1	0	DDA	1
MEPCE	SNUT2_HUMAN	USP39	19 27 34	26.67	0 0 0	1	1	0	DIA	1
MEPCE	ZNS11_HUMAN	ZNF511	8 8 9	8.33	0 0 0	1	1	0	DDA	1
MEPCE	ZNS11_HUMAN	ZNF511	24 21 36	27	0 0 0	1	1	0	DIA	1

Supplementary Table 2. Gene Ontology enrichment analysis of the high confident interaction partners for MEPCE and EIF4A2 identified in common by MSGFDB-DDA and MSPLIT-DIA (orange shading) or only by MSPLIT-DIA (green shading). GO analysis was performed at DAVID (<http://david.abcc.ncifcrf.gov/>) using GO Molecular Function FAT. Only categories with Pvalue < 0.001 are listed. Note the similarity in categories detected for preys identified by both approaches and by the MSPLIT-DIA approach only, attesting to the recovery of biologically relevant partners by MSPLIT-DIA.

GO term	Count	%	Pvalue	List total	Pop Hits	Pop Total	Fold enrich	Bonferroni	Benjamini	FDR
MEPCE common										
GO:0003723~RNA binding	35	62.5	4.97E-31	52	718	12983	12.17069316	5.31E-29	5.31E-29	5.58E-28
GO:0003735~structural constituent of ribosome	17	30.4	1.42E-18	52	168	12983	25.26453755	1.52E-16	7.58E-17	1.59E-15
GO:0005198~structural molecule activity	17	30.4	1.26E-09	52	634	12983	6.694703955	1.35E-07	4.49E-08	1.42E-06
GO:0017069~snRNA binding	4	7.14	6.72E-06	52	10	12983	99.86923077	7.19E-04	1.80E-04	0.00755487
GO:0031202~RNA splicing factor activity, transesterification mechanism	4	7.14	1.09E-04	52	24	12983	41.61217949	0.01160111	0.00233107	0.12250982
GO:0019843~rRNA binding	4	7.14	1.94E-04	52	29	12983	34.43766578	0.02056318	0.00345693	0.21803499
MEPCE DIA										
GO:0003723~RNA binding	15	57.7	1.14E-12	24	718	12983	11.30135794	9.77E-11	9.77E-11	1.23E-09
GO:0003735~structural constituent of ribosome	8	30.8	1.10E-08	24	168	12983	25.75992063	9.49E-07	4.75E-07	1.19E-05
GO:0005198~structural molecule activity	9	34.6	7.90E-06	24	634	12983	7.679219243	6.79E-04	2.26E-04	0.00851335
EIF4A2 common										
GO:0003743~translation initiation factor activity	19	35.2	8.07E-31	50	61	12983	80.87770492	5.89E-29	5.89E-29	8.42E-28
GO:0003723~RNA binding	32	59.3	2.47E-27	50	718	12983	11.57259053	1.80E-25	9.00E-26	2.57E-24
GO:0008135~translation factor activity, nucleic acid binding	19	35.2	1.16E-26	50	98	12983	50.3422449	8.45E-25	2.82E-25	1.21E-23
GO:0000166~nucleotide binding	28	51.9	2.20E-09	50	2245	12983	3.238521158	1.60E-07	4.01E-08	2.29E-06
GO:0003729~mRNA binding	8	14.8	4.59E-09	50	66	12983	31.47393939	3.35E-07	6.71E-08	4.80E-06
GO:0070035~purine NTP-dependent helicase activity	6	11.1	3.24E-05	50	98	12983	15.89755102	0.00236339	3.94E-04	0.03382735
GO:0008026~ATP-dependent helicase activity	6	11.1	3.24E-05	50	98	12983	15.89755102	0.00236339	3.94E-04	0.03382735
GO:0048027~mRNA 5'-UTR binding	3	5.56	8.33E-05	50	4	12983	194.745	0.00606448	8.69E-04	0.08693943
GO:0045182~translation regulator activity	4	7.41	1.72E-04	50	29	12983	35.81517241	0.01249978	0.00157108	0.17969336
GO:0004386~helicase activity	6	11.1	1.77E-04	50	140	12983	11.12828571	0.0128217	0.00143282	0.18434688
GO:0000339~RNA cap binding	3	5.56	3.85E-04	50	8	12983	97.3725	0.02772683	0.0028079	0.40124589
EIF4A2 DIA										
GO:0003723~RNA binding	22	73.3	8.26E-21	30	718	12983	13.26025998	5.29E-19	5.29E-19	8.39E-18
GO:0003735~structural constituent of ribosome	9	30	2.27E-09	30	168	12983	23.18392857	1.45E-07	7.25E-08	2.30E-06
GO:0003724~RNA helicase activity	4	13.3	3.52E-05	30	29	12983	59.69195402	0.0022511	7.51E-04	0.03576661
GO:0005198~structural molecule activity	9	30	5.33E-05	30	634	12983	6.143375394	0.00340312	8.52E-04	0.0540968
GO:0003729~mRNA binding	4	13.3	4.17E-04	30	66	12983	26.22828283	0.02635007	0.00532645	0.42297984

Supplementary Table 3. This table describes the datasets used in the analysis. Raw mass spectrometry data were deposited in the MassIVE repository at: <http://massive.ucsd.edu/ProteoSAFe/status.jsp?task=a4b32b9ba82e4885a8956b97ca71a1f8>

List of datasets	index	Raw file name
UPS only samples	1	18506_REP3_40fmol_UPS2_IDA_1
	2	18507_REP3_40fmol_UPS2_SWATH_1
	3	18508_REP3_400fmol_UPS2_IDA_1
	4	18509_REP3_400fmol_UPS2_SWATH_1
	5	18510_REP3_40fmol_UPS1_IDA_1
	6	18511_REP3_40fmol_UPS1_SWATH_1
	7	18512_REP3_400fmol_UPS1_IDA_1
	8	18513_REP3_400fmol_UPS1_SWATH_1
UPS spiked in with <i>E. coli</i> lysate	9	18485_REP3_1ug_Ecoli_IDA_NewStock2_2
	10	18486_REP3_1ug_Ecoli_SWATH_NewStock2_2
	11	18487_REP3_40fmol_UPS1_1ug_Ecoli_IDA_NewStock2_1
	12	18488_REP3_40fmol_UPS1_1ug_Ecoli_SWATH_NewStock2_1
	13	18489_REP3_400fmol_UPS1_1ug_Ecoli_IDA_NewStock2_1
	14	18490_REP3_400fmol_UPS1_1ug_Ecoli_SWATH_NewStock2_1
	15	18491_REP3_40fmol_UPS2_1ug_Ecoli_IDA_NewStock2_1
	16	18492_REP3_40fmol_UPS2_1ug_Ecoli_SWATH_NewStock2_1
	17	18493_REP3_400fmol_UPS2_1ug_Ecoli_IDA_NewStock2_1
	18	18494_REP3_400fmol_UPS2_1ug_Ecoli_SWATH_NewStock2_1
UPS spiked in with human lysate	19	18467_REP3_500ng_HumanLysate_IDA_1
	20	18468_REP3_500ng_HumanLysate_SWATH_1
	21	18469_REP3_40fmol_UPS1_500ng_HumanLysate_IDA_1
	22	18470_REP3_40fmol_UPS1_500ng_HumanLysate_SWATH_1
	23	18473_REP3_40fmol_UPS2_500ng_HumanLysate_IDA_1
	24	18474_REP3_40fmol_UPS2_500ng_HumanLysate_SWATH_1
	25	18475_REP3_400fmol_UPS2_500ng_HumanLysate_IDA_1
	26	18476_REP3_400fmol_UPS2_500ng_HumanLysate_SWATH_1
AP-MS data	27	IDA_EIF4aJune7-Biorep1
	28	IDA_EIF4aJune7-Biorep2
	29	IDA_EIF4aJune7-Biorep3
	30	IDA_GFPJune7-Biorep1
	31	IDA_GFPJune7-Biorep2
	32	IDA_GFPJune7-Biorep3
	33	IDA_MEPCEJune7-Biorep1
	34	IDA_MEPCEJune7-Biorep2
	35	IDA_MEPCEJune7-Biorep3
	36	Swath_EIF4aJune7-Biorep2
	37	Swath_EIF4aJune7-Biorep3
	38	Swath_EIFraJune7-Biorep1
	39	Swath_GFPJune7-Biorep1
	40	Swath_GFPJune7-Biorep2
	41	Swath_GFPJune7-Biorep3
	42	Swath_MEPCEJune7-Biorep1
	43	Swath_MEPCEJune7-Biorep2b
	44	Swath_MEPCEJune7-Biorep3