analytical chemistry

pubs.acs.org/ac Article

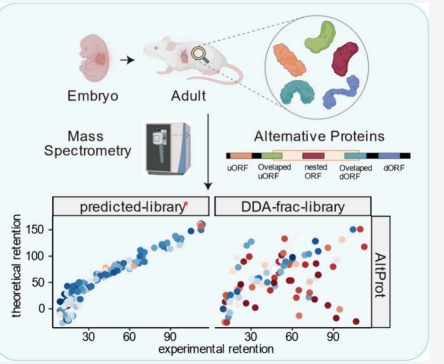
Enhanced Discovery of Alternative Proteins (AltProts) in Mouse Cardiac Development Using Data-Independent Acquisition (DIA) Proteomics

Yuanliang Zhang, Ying Yang, Kecheng Li, Lei Chen, Yang Yang, Chenxi Yang, Zhi Xie, Hongwei Wang, and Qian Zhao*

Cite This: *Anal. Chem.* 2025, 97, 1517−1527 Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: Alternative proteins (AltProts) are a class of proteins encoded by DNA sequences previously classified as noncoding. Despite their historically being overlooked, recent studies have highlighted their widespread presence and distinctive biological roles. So far, direct detection of AltProt has been relying on data-dependent acquisition (DDA) mass spectrometry (MS). However, data-independent acquisition (DIA) MS, a method that is rapidly gaining popularity for the analysis of canonical proteins, has seen limited application in AltProt research, largely due to the complexities involved in constructing DIA libraries. In this study, we present a novel DIA workflow that leverages a fragmentation spectra predictor for the efficient construction of DIA libraries, significantly enhancing the detection of AltProts. Our method achieved a 2-fold increase in the identification of AltProts and a 50% reduction in missing values compared to DDA. We conducted a comprehensive comparison of four AltProt databases, four DIA-library construction strategies, and three analytical software tools to establish an optimal workflow for AltProt analysis. Utilizing this workflow, we investigated the mouse heart development process and identified over 50 AltProts with differential expression between embryonic and adult heart tissues. Over 30 unannotated mouse AltProts were validated, including ASDURF, which played a crucial role in cardiac development. Our findings not only provide a practical workflow for MS-based AltProt analysis but also reveal novel AltProts with potential significance in biological functions.



INTRODUCTION

Over the past decades, 98% of the human genome has been presumed to be noncoding regions and remained largely unexplored.[1−4] However, emerging evidence suggests that many of these noncoding regions are translated, giving rise to what are known as alternative open reading frames (AltORFs). The proteins derived from these AltORFs, which have been termed alternative proteins (AltProts), are widely translated and perform unique functions in events such as respiratory chain reaction,[5] embryogenesis,[6] apoptosis,[7] and signal transduction.[8] The novelty and importance of AltProts underscores the urgent need for their further investigation with advanced methodologies.[9,10]

Although ribosome sequencing (Ribo-seq) has been powerful in predicting potential translation of AltProts, it does not provide direct evidence of the translation products. In contrast, mass spectrometry (MS) is capable of sequencing and quantifying AltProts with high confidence. Traditionally, the data-dependent acquisition (DDA) MS method has made great contribution to the constructing of the human proteome map,[11−16] the human protein Atlas,[17] etc., and has remained the mainstream method for detecting canonical proteins.[5,6,8,18−20] On the other hand, data-independent acquisition (DIA) has been demonstrated in recent years to outperform DDA in canonical protein analysis.[21−23] Although there has been a report on the application of DIA in AltProt analysis, in which Martinez et al. used GPF and DDA to generate an experiment-specific spectral library, systematic comparisons and method optimization for DIA analysis of AltProts remain limited, likely due to several technical challenges.[24] First, DIA identification largely depends on the quality of the mass spectral library, particularly the accuracy of the peptide fragmentation spectrum and the retention time (RT) contained within. Given the potential high sample/temporal/spatial specificity of AltProts,[25,26] there is significant uncertainty regarding the quantity and abundance of AltProts. Therefore, a sample-specific spectral library would be ideal, and the demand of instrument time and effort to construct this

Received: June 7, 2024
Revised: November 27, 2024
Accepted: November 27, 2024
Published: January 15, 2025

Downloaded via SUN YAT SEN (ZHONGSHAN) UNIV on September 22, 2025 at 09:21:27 (UTC).
See https://pubs.acs.org/sharingguidelines for options on how to legitimately share published articles.

ACS Publications

© 2025 The Authors. Published by American Chemical Society

sample-specific spectral library of AltProts is presumably much higher than that of the DDA method. Second, the use of DIA analysis to identify proteins from mixed spectra is more challenging, and the workflow is more complicated than DDA. Third, there are currently fewer search engines and less software available for DIA than for DDA. Theoretically, DIA could offer higher identification number and quantification accuracy of AltProts, because the DIA mode fragments all precursor ions instead of just the top abundant ions, generating mixed secondary spectra and thus providing more ion information for peptide identification and quantification.[27,28] However, its performance needs a systematic evaluation. Meanwhile, there is a pressing need for an optimal and effective DIA-based workflow customized for AltProts analysis.

In response to these needs, we systematically compared four widely used AltProt databases, four DIA-library building strategies, and three softwares to provide an optimal workflow. Our results indicated that DIA outperformed DDA in AltProt identification number and reproducibility. It was also worth noting that different library construction methods significantly influenced AltProts identification and yielded complementary discovery of AltProts. Next, we applied the DIA method to study functional AltProts in mouse heart development and identified over 50 differentially expressed AltProts. Among them, a representative ASDURF was validated to express a stable AltProt. The process involving ASDURF is likely to play a significant role in the development of the mouse heart.[29] Our work provided not only a reference to select appropriate DIA data processing method for AltProt studies but also novel AltProts that potentially regulated heart development and had important functions.

MATERIALS AND METHODS

Sample Collection. The C57BL/6 mice were purchased from Guangdong Medical Experimental Animal Center (Guangdong, China; License No: SCXK (YUE) 2018-0002). The embryonic and adult heart samples were collected from embryonic (E15.5) and adult (P42) C57BL/6 mice, respectively, and immediately frozen and preserved in liquid nitrogen. All animal experiments were authorized by the Hong Kong Polytechnic University Animal Subjects Ethics Sub-committee and conducted in accordance with the Institutional Guidelines and Animal Ordinance of the Department of Health guidelines.

Sample Preparation. Mouse heart samples were cut and homogenized with RIPA lysis buffer (50 mM Tris-HCl, 150 mM NaCl, 2 mM EDTA, 1% NP40, 1% SDC) containing 1× cocktail protein inhibitor. Liquid nitrogen was added to the homogenizer (Bertin Technologies, France), according to the manufacturer's manual to avoid protein degradation. The cell sample was ultrasonicated with the same lysis buffer and cocktail inhibitor for 3 min on ice. After extraction, the supernatant was collected after centrifugation at 16 000*g*, 4 °C for 25 min. BCA protein assay (Thermo Scientific) was applied to measure protein concentration. Dithiothreitol (DTT) at a final concentration of 10 mM was added first and placed at 37 °C for 45 min, followed by iodoacetamide (IAM) at a final concentration of 30 mM for 30 min in a dark room. The sp3 digestion experiment was performed using Sera-Mag beads (Cytiva), based on a previous study. Briefly, washed beads were added to the lysed sample, and then an equal volume of pure ethanol was added to bind the proteins and beads. Next, the sample was placed on a magnetic rack, the waste solution was discarded, and the beads were washed three times with 80% methanol. Then, 45 μL of 50 mM ammonium bicarbonate was added with Lys-C (enzyme to protein ratio 1:100) at 37 °C for 4 h. Lastly, trypsin (enzyme:protein ratio 1:40) was added at 37 °C overnight. Beads were washed with ddH_2O and the supernatant was collected for LC-MS analysis.

LC-MS Analysis. LC-MS analyses were performed using an Orbitrap Exploris 480 Mass Spectrometer (Thermo Fisher Scientific) coupled with an UltiMate 3000 HPLC system (Thermo Fisher Scientific). Peptides were injected into LC and separated by a 2-h gradient using the 25 cm Aurora column (IonOpticks, Australia) with mobile phases buffer A (99.9% H_2O, 0.1% FA) and buffer B (80% ACN, 0.1% FA). The 2-h gradient was (min, %B): 0,8; 3,8; 104,32; 113,90; 120,90 and utilized by three different MS modes. For data-dependent acquisition (DDA) mode, MS1 resolution was set at 120 000 with 50 ms max injection time (MIT) while MS2 was set at 30 000 with 50 ms MIT. The isolation window was set as 1.6 m/z, and the cycle time of data-dependent mode was set as 3 s. For the data-independent acquisition (DIA) mode, MS1 was set at 60 000 with 50 ms MIT and the precursor range starting from 400 to 1000. This range was evenly divided by 10 m/z into 60 MS2 windows, each with an overlap of 1 m/z. MS2 resolution was set as 30 000 at 32% HCD collision energy, while desired minimum points across the peak were set as 8. For gas-phase fractionation (GPF) mode, the recommended 6× GPF-DIA acquisition settings were applied based on previous research.[30] For PRM mode, the m/z of target peptides was set in the MS2 transition table with a 1.2 m/z isolation window and 80 ms MIT.

MS Raw Data Searching. Regarding the spectral library construction method for DIA analysis, four methods were designed. The first method involved collecting data from eight fractions and constructing a spectral library using the software MSFragger and FragPipe (v17.1)[31] (DDA-frac-library). In FragPipe, the tabs called "MSFragger", "Validation", and "Spec Lib" were utilized to generate the spectral library with default parameters. The second method involved collecting data from six GPF datasets, employing DIA-Umpire[32] and MSFragger for data searching and library construction (GPF-msf-library). In contrast to the first library construction strategy, this method utilized the "DIA-UMPIRE" tab. The third method also involved six GPF datasets, but it focused on searching against a fully predicted library using DIA-NN[33] to generate a sublibrary (GPF-diann-library). The *.speclib file was generated using default parameters of DIA-NN, which included a 1% false discovery rate (FDR), fixed modification of carbamidomethylation on cysteine residues, one missed cleavage, and trypsin protease mode. The fourth one directly used a fully predicted library (predicted-library). After obtaining the library, DIA-NN was used to analyze DIA data with the Match-between-run (MBR) search mode to reduce the library's redundancy.[34] For DDA, raw data were searched against different databases using Fragpipe (v17.1).[31] After MSFragger target-decoy searching with ± 10 ppm MS1 tolerance, ± 0.02 Da MS2 tolerance, Percolator[35] and ProteinProphet were used for PSM validation and Protein Inference. EasyPQP was used for the experimental spectral library generation. For PRM, raw data were analyzed by Skyline (v21.2.0.369).[36] After data searching, the identified peptides were classified as canonical peptides or AltProt peptides according to sequential mapping (Leucine and Isoleucine were treated as same) and BLAST.[37] All relevant raw data, fasta files, and DIA-NN search results have been

https://doi.org/10.1021/acs.analchem.4c02924
Anal. Chem. 2025, 97, 1517−1527

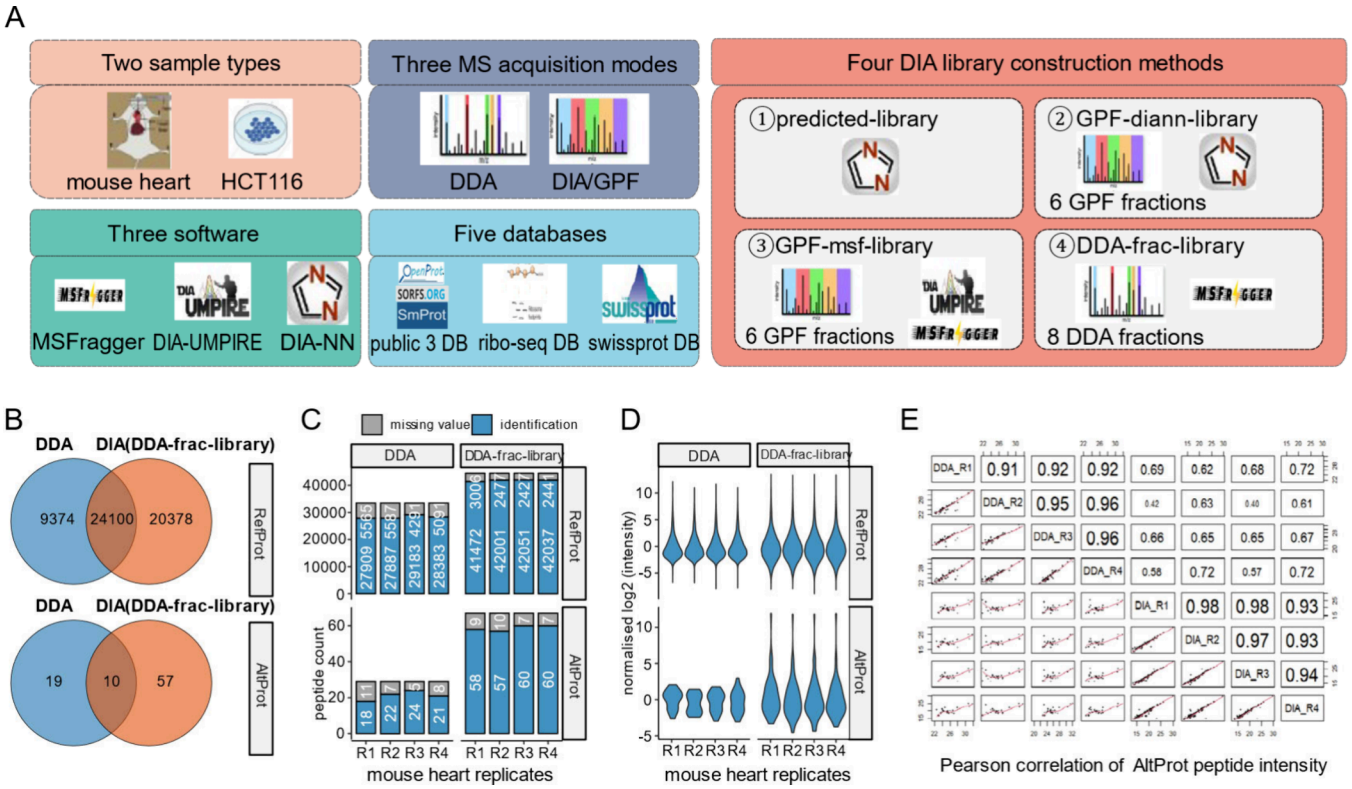


Figure 1. Comparison of different mass spectrometry methods in terms of AltProt analysis. (A) Schematic workflow of DDA and DIA with four different library construction methods. (B–D) Venn diagram, box plot and violin diagram of identified canonical and AltProt peptides of mouse heart from DDA and DIA, respectively. (E) Pearson correlation of AltProt peptide intensity from four DDA and four DIA mouse heart replicates.

uploaded and stored in the PRIDE database with Accessing ID PXD045956.

Differential Expression Analysis and Spectra Validation. After log 2 transformation, R package impute.knn (v1.66) was used to fill in the missing values, and those with more than 50% missing values were filtered out. However, if all missing values were present in the same biological group, such as adult or embryonic, these proteins were considered differential proteins. Differentially expressed proteins were then filtered using a 2-fold change, 0.05 *p*-value, and the Benjamini-Hochberg algorithm to correct the *p*-value. The PCA plot was performed with R package pcaMethods (v1.84) and GO analysis was performed with R package clusterProfiler (v4.0.5).[38]

Normalized spectral angle (SA) was used to validate spectra quality according to previous research.[39,40] Prosit[41] was used for spectra prediction to compare experimental spectra with predicted spectra and the high and low SA values represented the spectrum's relative quality.

Gene Conservation. UCSC Genome Browser[42] was employed to perform gene comparisons across species and gene conservation analyses, followed by the presentation of alignment sequence using Jalview (v2.11.2.4).[43]

Western Blot Analysis. The overexpression experiment was performed with transfection of plasmids containing AltProt cDNA sequences in the CHO cell line. The plasmids were constructed based on the PB510-P2A-EGFP-control-3 vector and ordered from Gene Universal (Delaware, USA).

Three micrograms (3 μg) of plasmid of each candidate was diluted in 6 μL of Promega ViaFect Transfection Reagent and added to a 6-well plate. After 48 h of transfection, samples were collected and lysed with RIPA lysis buffer. 20 μg proteins of cell lysed were separated by Tricine SDS-PAGE gel, followed by binding of antiflag antibodies and antirabbit IgG antibodies.

RESULTS AND DISCUSSION

The Design of Comparative Evaluation of Different MS-Based Workflows for AltProt Analysis. First, we designed a systematic scheme to evaluate commonly used DDA and DIA methods, in terms of AltProts detection. As illustrated in Figure 1A, two distinct sample types, the HCT116 cell line and the adult mouse heart tissue, were chosen to ensure an unbiased evaluation of methods. Subsequently, we compared three mass spectrometry scanning modes, DDA, DIA, and gas-phase fractionation (GPF),[44] with four technical replicates. Additionally, we collected eight fraction data and six GPF data from two samples to assist the DIA data analyses with various library construction methods. Three complementary databases were then selected to ensure a comprehensive coverage of AltProts and minimize false-negative discovery commonly caused by an incomplete reference database. The first database was a reviewed RefProt database from UniProt;[45] the second database was a combined AltProt database that included OpenProt,[46] sORFs.org,[47] and SmProt;[48] and the third one was our in-house Ribo-seq-based database.[49] Lastly, four different library construction methods (see the Methods section) were designed to evaluate the impact of several DIA library construction methods on protein identification. Briefly, the four library construction methods were as follows: the first approach is to build a predicted spectral library based on machine learning; the second and third approaches involve using DIA-NN and MSFragger, respectively, to search gas-phase fractionation data to generate

https://doi.org/10.1021/acs.analchem.4c02924
Anal. Chem. 2025, 97, 1517−1527

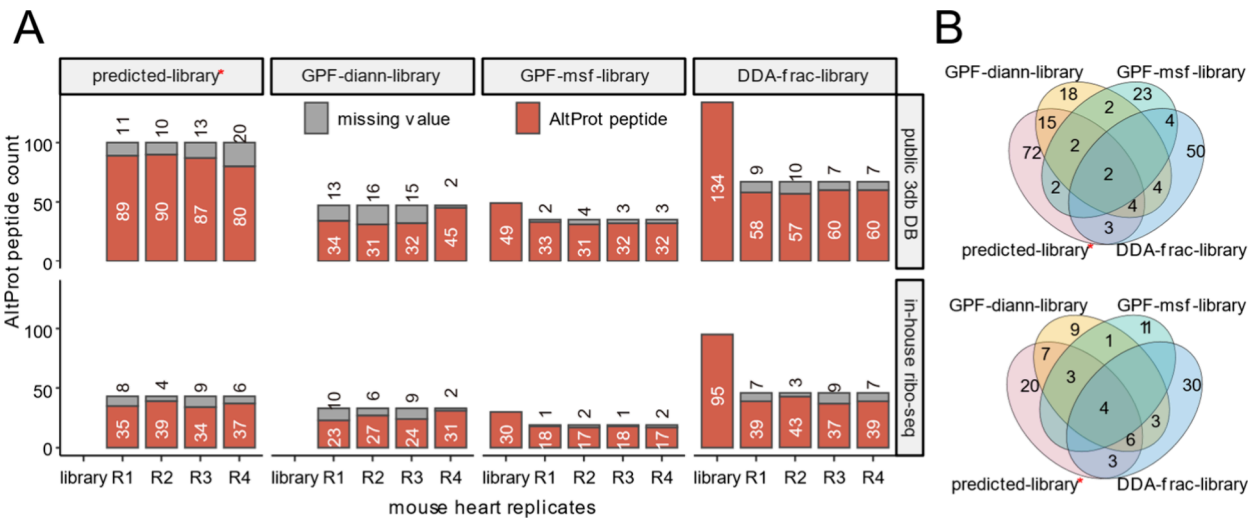


Figure 2. Diverse identifications by different strategies. (A) Boxplot of identified AltProt peptide count by four different library construction methods and three different databases within four mouse heart replicate samples. (B) Venn diagram of identified AltProt peptides by four different library construction methods from different databases.

spectral libraries, while the last approach is to construct a sample-specific spectral library from experimental fractionated samples. In summary, we conducted a comprehensive evaluation of the impact of various data processing parameters on the identification and quantification of canonical proteins and AltProts. The assessment involved the use of two biological samples (human HCT116 cell line and mouse heart), two mass spectrometry data acquisition modes (data-independent acquisition and data-dependent acquisition), three protein sequence databases (RefProt DB, AltProt DB, Ribo-seq DB), and four spectral library construction methods. By employing these multiple factors, it provided a comprehensive assessment of the impact of different data processing parameters on the identification and quantification of canonical proteins and AltProts.

DIA Outperformed DDA in Identification and Quantification of AltProts. To evaluate the effectiveness of DIA proteomics with DDA proteomics in terms of AltProt analysis, we selected the traditional DDA-assisted spectral library construction method, named "DDA-frac-library", for DIA data analysis and compared it with traditional TopN DDA proteomics. Our results showed that DIA outperformed DDA both qualitatively and quantitatively in AltProt analysis (Figure 1C). DIA identified 1.48 times more peptides than DDA in mouse heart samples, with an average of 41 890 and 28 340 peptides, respectively. In addition, the missing values of canonical peptides identified by DIA (5.8%) were 2.64 times lower than those of DDA (15.3%). Meanwhile, for the AltProt peptides, DIA and DDA identified 22 peptides and 59 peptides on average (Figures 1B and 1C), respectively, and the number of AltProt peptides in DIA was 2.67 times more than DDA. As shown in Figure 1C, the missing values of DIA and DDA were 26.4% and 12.3%, respectively, the former being 2.14 times greater than the latter.

Apart from identification, we also observed that DIA had a broader quantitative range and better quantitative correlation than did DDA (Figures 1D and 1E). The broader distribution area of low abundance proteins detected by DIA implied its superior quantitative performance for these proteins in comparison with DDA. A similar trend was observed in HCT116 data for both canonical and AltProt peptides (Figure S1). Collectively, the mass spectrometry data from the mouse heart and HCT116 provided direct evidence of the superiority of the DIA method in the identification of canonical and AltProt peptides.

Different Library Construction Methods Influence AltProt Identification. There is growing evidence that DIA is more effective in identifying AltProt peptides; however, the identification of these peptides is influenced by the library construction method used, and the influence of various library construction methods on AltProt peptide identification should be examined further. Currently, the usual fractionation-based DIA procedure relies on an experimental library to aid in the analysis of mixed spectra. However, with the development of deep learning, predicted peptide MS2 spectra, and retention time (RT) can be directly used for DIA analysis, which can reduce the cost of experiments and increase identification number.[41,50,51] To evaluate the DIA data analysis method using the predicted library, we calculated the false discovery rate (FDR) with a pseudotargets FDR estimation strategy and LFQbench analysis were performed. Our observation was consistent with previous report that the experimental FDR was smaller than the set FDR, which demonstrated the accountability of the spectral library.(Figure S4A).[52] The LFQbench[28] plot illustrated the qualitative and relative quantitative accuracy of the prediction library. (Figure S4B). Both pseudotargets FDR and LFQbench analysis implied that the FDR of the prediction library construction method was controllable and the identification results were specific and sensitive.

Next, in order to examine the effect of different library construction methods on AltProt identification (see the Methods section for library construction), adult mouse heart data were searched with four different libraries against three different databases. The data suggested the fully predicted library (predicted-library) and sublibrary generated by GPF (GPF-diann-library) led to relatively high identification of 50K canonical peptides and 100 AltProt peptides, while the experimental fraction library (DDA-frac-library) only identified 41K canonical peptides and 67 AltProt peptides (Figure 2A and Figure S2A). In terms of AltProt peptide identification, the predicted library method also yielded the highest number of

https://doi.org/10.1021/acs.analchem.4c02924
Anal. Chem. 2025, 97, 1517−1527

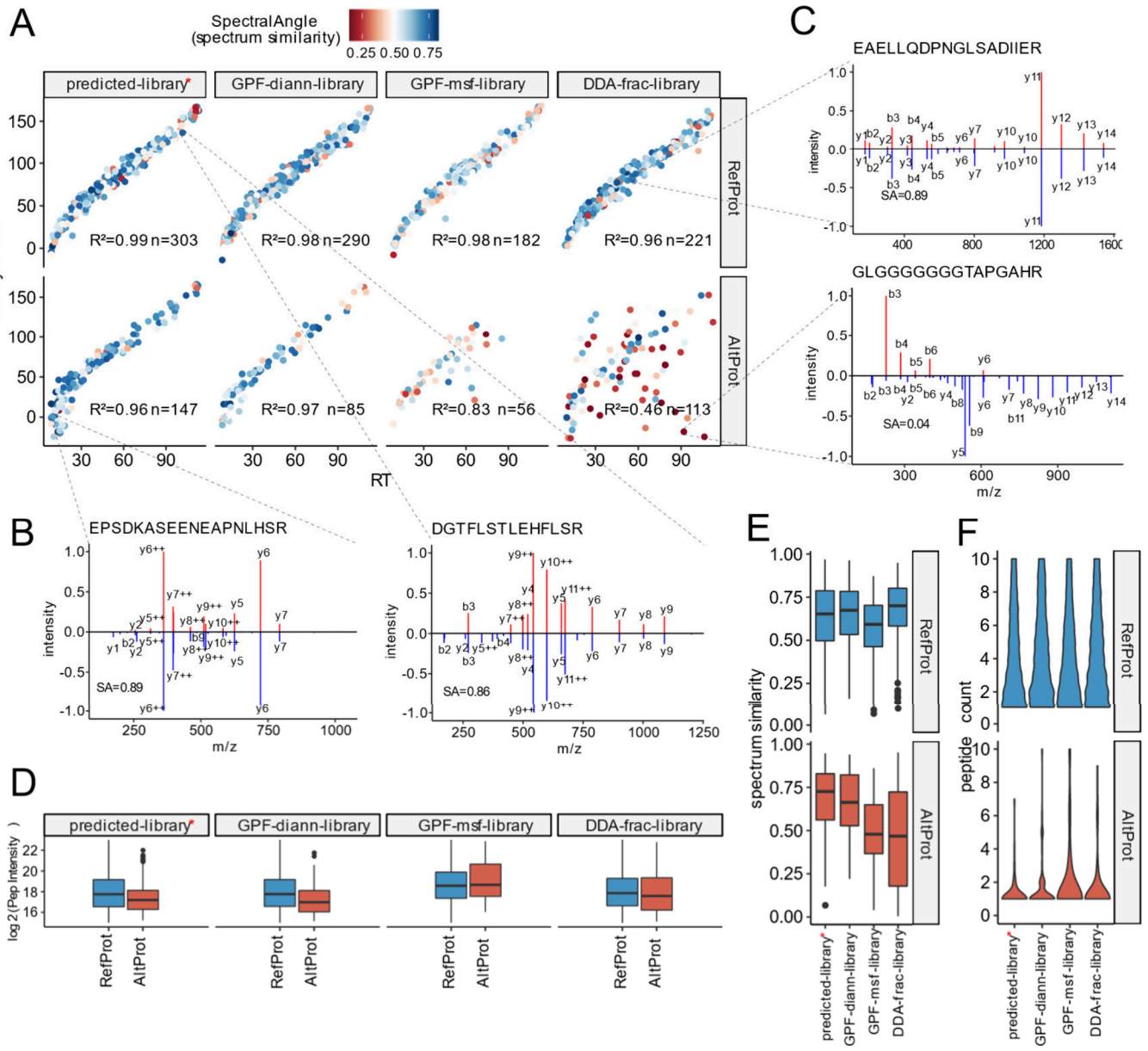


Figure 3. Evaluation of identifications from different strategies. (A) Correlation plots of the endogenous and predicted retention time of the canonical and AltProt and peptides, R square was calculated by Pearson method and SA value was indicated by colored dots. (B, C) Four examples of comparing spectra between endogenous and predicted peptides. (D) Distribution plot of peptide intensity identified by four strategies. (E) SA distribution plot of three types of peptides identified by four strategies. (F) Distribution of the number of peptide identifications per protein by four strategies.

AltProt identification. And the limited overlap in the identification of AltProt peptides across different library construction methods (Figure 2B) suggested that the use of different library construction methods was more likely to result in the discovery of distinct AltProt peptides. A similar phenomenon was also observed in the HCT116 dataset (Figures S2B–S2D). And the DIA spectral library construction method influenced the identification of AltProt peptides more dramatically than on canonical peptides. Several explanations for this phenomenon include the redundancy and inaccuracy of the database entries. As illustrated in Figure S9A, the number of entries in the three public databases ranges from 240 086 to 503 779, which may result in an increased number of false negatives. Furthermore, regarding the identification of canonical proteins, the four different library construction strategies were able to identify more than one-third of the proteins, whereas this proportion was <5% for identifying AltProt, as demonstrated in Figure 2B and Figure S2D. This discrepancy suggested that the AltProt databases utilized in this study may not be as comprehensive as the UniProt database, implying a potential limitation in the inaccuracy of the AltProt library, compared with the UniProt database.

Besides the quantity, we also assessed the quality of each identification by comparing each spectrum and the corresponding retention time (RT) with theoretical values predicted with Prosit. We chose all AltProt peptides and randomly selected similar amounts of canonical peptides to predict MS2 spectra and RT. By calculating the spectral angle (SA), the SA and RT correlation was plotted (Figure 3A). The chromatic representation of the data points indicated the quality of the

https://doi.org/10.1021/acs.analchem.4c02924
Anal. Chem. 2025, 97, 1517−1527

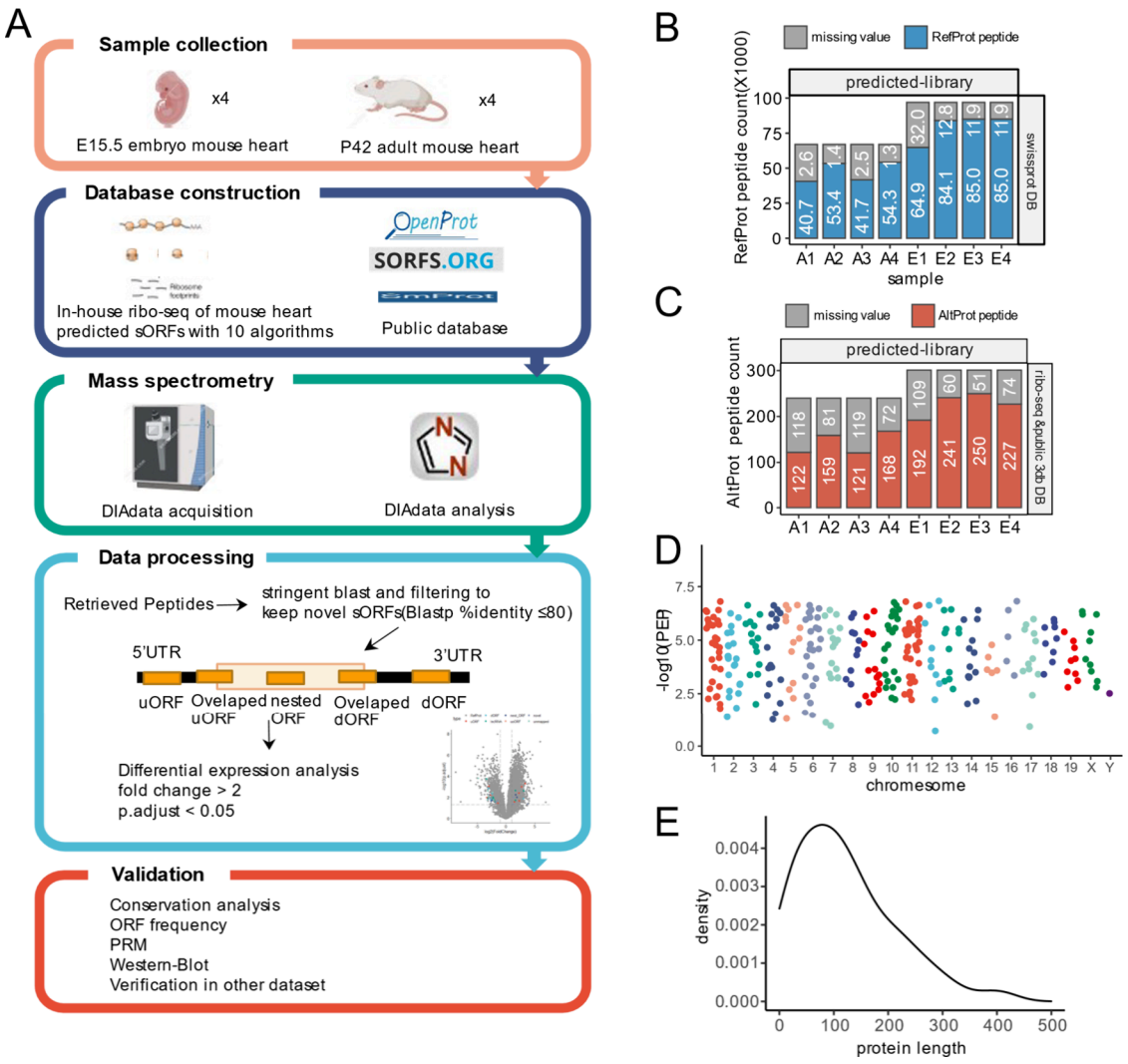


Figure 4. Application of DIA in mouse heart development. (A) Workflow of AltProt exploration in embryonic and adult mouse heart samples. (B, C) Count of canonical and AltProt peptides identified in eight samples by the predicted library. (D) Manhattan plot of gene chromosome location of AltProt peptides. (E) Protein length distribution of AltProt including uORF, overlapped uORF, AltProt ORF, overlapped dORF, and dORF.

spectrum, and the spatial distribution of the points represented the correlation of RT. As illustrated in Figures 3B and 3C for instance, the AltProt peptide GLGGGGGGGTAPGAHR, which was identified by the DDA-frac-library, exhibited a spectral similarity of only 0.04 to the predicted spectrum, whereas the AltProt peptide EPSDKASEENEAPNLHSR, which was derived from the predicted library, had an SA value of 0.89 (Figures 3B and 3C). Additionally, we observed that the peptide GLGGGGGGGTAPGAHR, which had a lower SA value, deviated from the RT correlation. Based on RT and SA, we had reason to believe that the identification of EPSDKASEENEAPNLHSR was more reliable than that of GLGGGGGGGTAPGAHR.

In general, the results illustrated that the SA values of the two libraries from MSFragger (GPF-msf-library and DDA-frac-library) were relatively low and the SA value of the AltProt peptide on average was ∼0.5 and the msbooster was not applied in the data searching process step of FragPipe. In contrast, the SA values obtained by the other two methods were in the range of 0.70−0.75, which was close to the SA value derived from the canonical method (Figure 3E). Moreover, the RT correlation of AltProt peptides derived from two MSFragger libraries was also inferior. The poor spectral similarity and RT correlation implied a higher false discovery rate of AltProt peptides (Figure S10). Intriguingly, in the identification of canonical peptides, the SA and RT obtained by different library-building methods matched comparatively to the predicted spectra, indicating that the FDR of canonical peptides was relatively controlled. Obviously, the usage of experimental DDA libraries has the potential to result in more false AltProt peptide identifications. Potential false discoveries of AltProt were discovered during the DDA database search process and could be propagated into the DIA searching results. Such false discoveries could also lead to perturbations in the FDR calculation during DIA database searching and eventually reduce the quantity of identified AltProts. On the contrary, the predicted strategy was constructed spectral library purely based on protein sequences from the database in an unbiased manner, which avoided steps that could introduce false discoveries in the spectral library.

Therefore, two mitigations are recommended to reduce the impact of false discoveries. One may utilize algorithms such as deeprescore[40] and msbooster[53] to optimize library construction procedures. Another option is to manually verify the

https://doi.org/10.1021/acs.analchem.4c02924
Anal. Chem. 2025, 97, 1517−1527

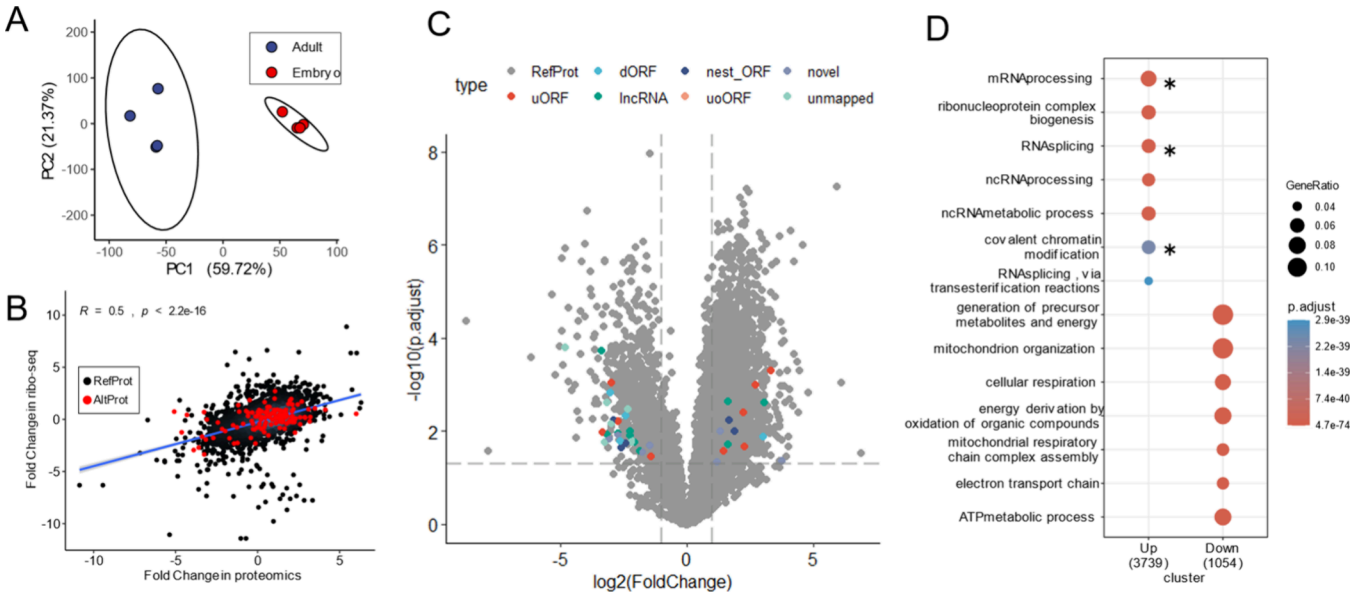


Figure 5. Different expression analyses. (A) PCA plot of eight samples, where A represents adult mouse heart and E represents embryonic mouse heart. (B) Protein fold change correlation between Ribo-seq and mass spectrometry technique. (C) Volcano plot and protein type distribution of identified differentially expressed proteins. (D) GO terms (biological process) enriched in up- and down-regulated genes in embryonic development. The asterisks represent that certain proteins under this GO pathway have corresponding alternative proteins present.

spectra of specific AltProts following the data search processes with DIA or DDA, particularly for AltProt studies, as a cutoff of 1% FDR cannot ensure accurate identifications.

In addition to the quality of the spectra, the intensities of the AltProt peptides identified by the fully predicted library were significantly lower than those of the canonical peptides, with an average intensity that was just half that of the canonical spectrum (Figure 3D). The average peptide count number per AltProt was typically one (Figure 3F). These findings suggested that the lower abundance and smaller number of peptides from one AltProt could contribute to poor spectrum quality, resulting in different analysis methods leading to different AltProt identifications, thus explaining why AltProt detection was challenging.

Differentially Expressed AltProts in the Mouse Heart Development Were Detected by Using an Optimized DIA Method. To further delve into the practical biological applications of the DIA method in AltProt exploration studies, we utilized the same 10 processing software in tandem with the previous Ribo-seq data to create a specific database to aid mass spectrometry analysis.[54] To circumvent the potential shortfall in AltProt identification due to the diversity of library search methods, we persistently employed the four library-building methods and three different databases for concurrent analysis (Figure 4A) for identifying functional AltProts in mouse heart development.

Taking the Ribo-seq database as an example, we identified approximately 50K−80K peptides in both adult and embryonic mouse heart samples and subsequently discovered nearly 300 AltProt peptides after mapping and blast filtering (Figures 4B and 4C). We found that the N_extension held the top spot in terms of sORF type (Figure S5C), a finding that aligned with Na et al.'s observation of the N_extension having the highest count.[55] These mass spectral results underscored the dynamic nature of genome translation and the incompleteness of the conventional protein database. Additionally, as shown in Figure 4D, AltProts were identified across various chromosomes, and these AltProts were often less than 200 amino acids in length. This revealed the widespread presence of AltProts in the human genome, warranting further investigation.

As illustrated in Figure 5A, we performed a PCA analysis that revealed a substantial divergence between the proteomes of adult mouse hearts and those of embryonic mouse hearts. Additionally, we assessed the correlation between the fold change in the mass spectrometry-based proteome and the fold change in the Ribo-seq-based transcriptome and got the result of a p-value of less than 2.2×10^{-16} (Figure 5B). These analyses lend credence to the relative reliability of our proteomic study. Upon applying a 2-fold change and a 0.05 adjusted p-value as filtering criteria, 14 and 41 differentially expressed AltProts in the Ribo-seq database and 3db were identified, respectively (Figure 5C) in addition to thousands of canonical differential proteins. GO pathways (Figure 5D) revealed that, in the biological process, embryonic mouse hearts were predominantly involved in RNA-related pathways, while proteins of adult mouse hearts were predominantly involved in metabolism and energy production, a finding that aligned perfectly with the biological phenotype. Within the GO pathways (please refer to Table S3), we not only identified the uORF of Cd2bp2, the lncRNA of Ash1l, and the dORF of Smyd2, whose reference proteins were involved in pathways such as mRNA processing and covalent chromatin modification, but also detected the uORFs of Noct and Drd4. The RefProt of these two AltProts were in the circadian rhythm pathway of GO. This suggested potential interactions and underlying roles between AltProts and their corresponding RefProts.

Validation of Expression of ASDURF and Other AltProts. We compiled a comprehensive list of differentially expressed AltProts derived from two databases and four library construction methodologies (see Table S3). Further validation of these AltProts was conducted through PRM, Western blot, conservation analysis, and cross-referencing with other datasets.

https://doi.org/10.1021/acs.analchem.4c02924
Anal. Chem. 2025, 97, 1517−1527

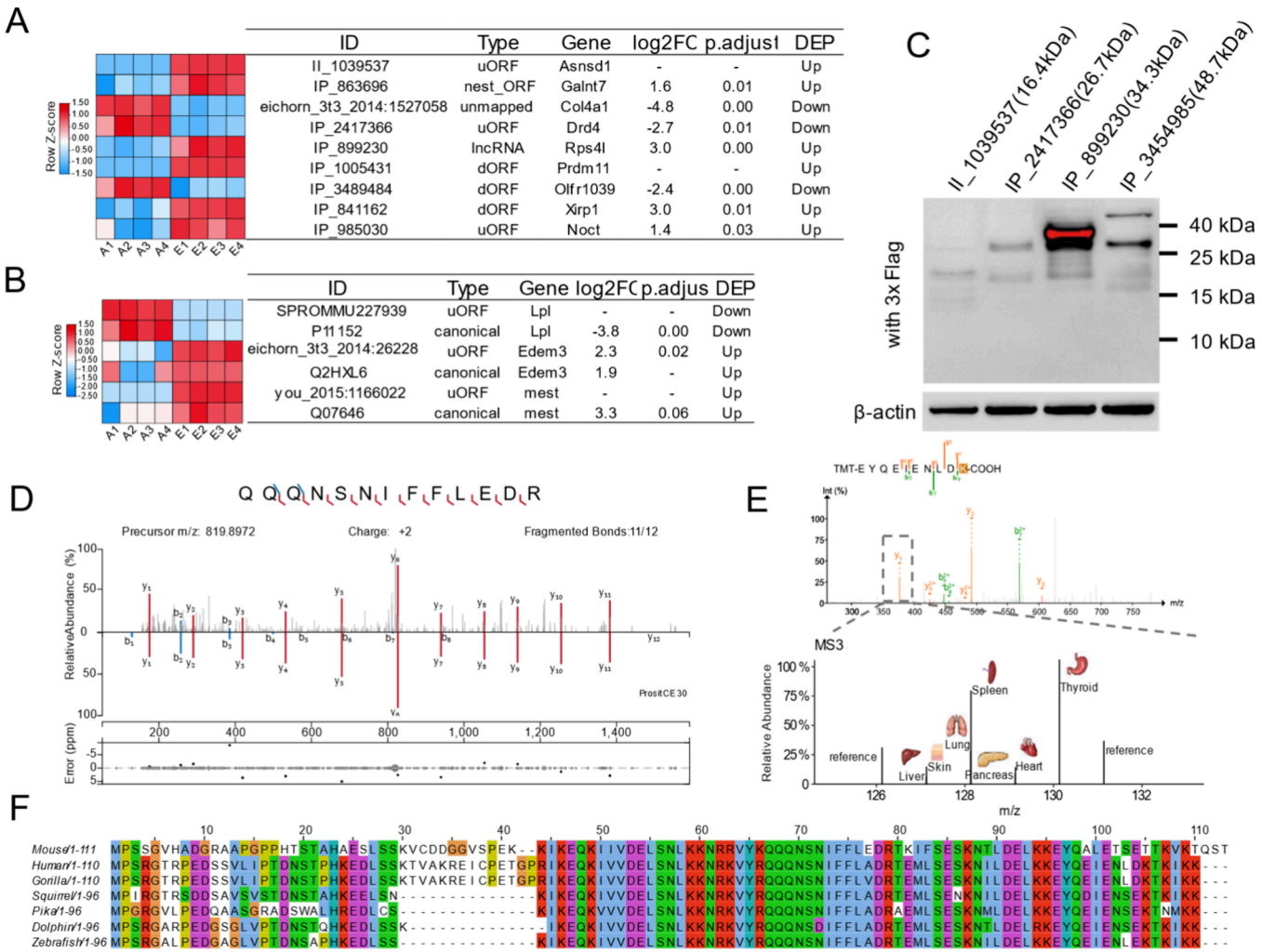


Figure 6. Validation of ASDURF and other AltProts. (A) Heatmap of nine differently expressed AltProt examples. (B) Heatmap of three uORFs with corresponding main ORF. (C) Overexpression validation of four selected AltProts by Western Blot experiment. The labeled molecular weight had been added with the 3.3 kDa of flag tag size. (D) Comparison between predicted and endogenous spectra of ASDURF. (E) MS2 and MS3 spectra of AltProt peptide EYQEIENLDK from ASDURF. The y1 ion was triggered for the MS3 event and each TMT peak represented the protein abundance existed in different human organs. (F) Protein sequence alignment plot of ASDURF in seven species.

The PRM results demonstrated that 23 AltProts from the 3db database and 6 AltProts from the Ribo-seq database successfully passed manual verification (see Figures 6A, 6B, and S6B). We selected four AltProts for further translational confirmation via overexpression Western Blot experiments (Figure 6C). Four distinct bands of AltProts were observed in the Western plot, indicating their translationality. Among the differentially expressed AltProts validated by PRM, a few previously reported AltProts, including ASDURF (encoded by uORF),[29] ASH1L (encoded by lncRNA),[56] and RPS4L (encoded by lncRNA)[24,57] were detected in our study, which demonstrated the sensitivity of our method. Besides, we identified three uORFs that corresponded to the canonical protein of their respective annotated ORF. These three pairs of AltProts and canonical proteins exhibited identical expression trends in mouse heart development. This may be due to the transcription and translation mechanisms, or there may be some unidentified regulatory mechanism between uORF and main ORF. For instance, Lipoprotein lipase (LPL) was highly expressed in adult mouse hearts, while ER degradation-enhancing alpha-mannosidase-like protein (Edem) and Mesoderm-specific transcript homologue protein (mest) were highly expressed in embryonic hearts (Figure 6B). Their coexpression patterns with uORFs hint at potential unidentified regulatory mechanisms, although further evidence is required to elucidate their interactions.

Lastly, we discovered the AltProt ASNSD1 Upstream Open Reading Frame (ASDURF) in our dataset. Previous reports have suggested that this AltProt is involved in the construction of the PAQosome,[58,59] which plays a crucial role in various fundamental cellular activities, including protein synthesis, ribosome biogenesis, transcription, and splicing.[60] Furthermore, it has been reported that phenotypically it can enhance the survival of medulloblastoma cells.[61] Our PRM spectra confirmed the presence and elevated expression of ASDURF in mouse embryonic hearts (Figure 6D), ASDURF has been identified as a previously uncharacterized component of the PAQosome complex,[58] which is involved in the biogenesis of various protein complexes, transcription, splicing, and other cellular processes.[60] Moreover, in addition to ASDURF, we observed that other subunits of the PAQosome, including RPAP3, PIH1D1, PFDN2, and RUVBL1, are similarly highly expressed in mouse embryonic hearts. This observation suggests that the PAQosome, comprising components such

https://doi.org/10.1021/acs.analchem.4c02924
Anal. Chem. 2025, 97, 1517−1527

as ASDURF, plays a significant role in functions, including protein synthesis and transcriptional regulation during cardiac development. To further validate this, we cross-referenced ASDURF with public human MS data[62] and found ASDURF was identified in several different organs, especially in the thyroid (Figure 6E). Moreover, conservation analysis revealed that AltProt was highly conserved across several species (Figure 6F). Although ASDURF is currently classified as a reviewed human protein in the UniProt database, it is not yet classified as such in mice. The mass spectrometry findings presented herein provide the first compelling evidence of the presence of ASDURF in the mouse proteome.

CONCLUSIONS

We have innovatively and systematically assessed the performance of DDA and DIA in the identification of both canonical and AltProt peptides. This was achieved through comprehensive comparisons across diverse samples, various mass spectrometry data-type databases, library construction methodologies, and mass spectrometry analysis software. Our findings underscored the obvious superiority of DIA over DDA in both qualitative and quantitative dimensions. Particularly in the identification of AltProt peptides, the DIA method outperformed DDA by a factor of 2 while also halving the number of missing values. This suggests that the DIA method holds significant potential for enhancing the identification of both canonical and AltProt peptides.

However, our study also has revealed that the choice of library construction method exerts a substantial influence on the final identification results. Different library construction methods can lead to variations in peptide identification, especially in the case of AltProt peptides. Notably, AltProt peptides derived from the experimental fraction library construction method were found to have a higher false discovery rate, compared to those obtained through the fully predicted library method.

In our application of DIA to a mouse heart development model, we identified nearly 50 differentially expressed AltProts through rigorous filtering. Following multiple validation techniques, we discovered ASDURF and three uORFs, three of which have the potential to regulate the host gene. Furthermore, we provided the first evidence of ASDURF in the mouse proteome, thereby substantiating the existence of AltProts and hinting at their potentially significant role in the proteome.

We therefore posit that our comparative analysis of previous methods and practical biological applications can serve as a valuable guide for selecting DIA analysis methods and refining the spectral library, paving the way for future exploratory studies on AltProts.

LIMITATIONS

In this Article, we conducted a systematic evaluation of detecting AltProts using mass spectrometry techniques including DDA, DIA, library construction methods, and databases. In this study, we did not adopt an AltProt enrichment strategies. We believe that current enrichment strategies, such as SDS-PAGE in-gel digestion,[63] multiple-protease,[64] and multidimensional separation,[65] should have some enrichment effects. However, our focus in this Article was primarily on the mass spectrometry DDA/DIA mode and library construction methods. Different AltProt enrichment methods should be evaluated in the future for seamless integration with DIA. With regard to databases, we collected as many potential AltProt entries as possible from public data. While the number of potential AltProts has increased, so has the prevalence of false positives in the database. Currently, we do not have more effective methods to reduce library redundancy. Although DIA-NN can perform two-step searches to reduce spectral library redundancy, a high-quality AltProt library can theoretically better assist in AltProt identification, which is an area that we should explore further. Furthermore, regarding various library construction methodologies, our data indicated that the experimental library method based on DDA fractions may exhibit a higher rate of false positives, a conclusion derived from comparisons to the Prosit-predicted spectral library. Although the synthesis of peptides for PRM validation of AltProt spectra could more effectively confirm the existence of AltProts, we did not implement this approach in the present study. We acknowledge that conducting PRM experiments with several hundred synthesized peptides presents significant challenges.

Lastly, regarding the potential functional AltProts identified in mouse heart development, many AltProts require further detailed validation. For example, confocal microscopy experiments are needed to determine the localization of AltProts. For the uORF of LPL, Edem3, and mest, subsequent coimmunoprecipitation experiments should be conducted to assess whether there is indeed a physical interaction between the uORF and the main ORF.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.analchem.4c02924.

Construction procedure of AltProt database; additional evaluation of MS method using HCT116 sample; LFQbench analysis of Public MS data; comparison of AltProts across public databases and additional validation of AltProt peptides (Figures S1−S6, S8−S10) and associated references (PDF)

Peptide list data from DDA vs DIA comparison dataset (Table S1) (XLSX)

Peptide list data from four DIA library construction methods dataset (Table S2) (XLSX)

Protein list data of different expressed AltProts in mouse heart development (Table S3) (XLSX)

Experimental spectrum validation using PRM and corresponding Prosit predicted spectrum (Figure S7) (PDF)

AUTHOR INFORMATION

Corresponding Author

Qian Zhao − *Department of Applied Biology and Chemical Technology, State Key Laboratory of Chemical Biology and Drug Discovery, Hong Kong Polytechnic University, Hong Kong 999077, China;* orcid.org/0000-0003-2244-6516; Email: q.zhao@polyu.edu.hk

Authors

Yuanliang Zhang − *Department of Applied Biology and Chemical Technology, State Key Laboratory of Chemical Biology and Drug Discovery, Hong Kong Polytechnic University, Hong Kong 999077, China*

Ying Yang − *Department of Applied Biology and Chemical Technology, State Key Laboratory of Chemical Biology and Drug Discovery, Hong Kong Polytechnic University, Hong Kong 999077, China*
Kecheng Li − *Department of Applied Biology and Chemical Technology, State Key Laboratory of Chemical Biology and Drug Discovery, Hong Kong Polytechnic University, Hong Kong 999077, China*
Lei Chen − *Department of Applied Biology and Chemical Technology, State Key Laboratory of Chemical Biology and Drug Discovery, Hong Kong Polytechnic University, Hong Kong 999077, China*
Yang Yang − *Department of Applied Biology and Chemical Technology, State Key Laboratory of Chemical Biology and Drug Discovery, Hong Kong Polytechnic University, Hong Kong 999077, China*
Chenxi Yang − *Department of Applied Biology and Chemical Technology, State Key Laboratory of Chemical Biology and Drug Discovery, Hong Kong Polytechnic University, Hong Kong 999077, China*
Zhi Xie − *State Key Laboratory of Ophthalmology, Zhongshan Ophthalmic Center, Sun Yat-sen University, Guangzhou 510060, China*
Hongwei Wang − *State Key Laboratory of Ophthalmology, Zhongshan Ophthalmic Center, Sun Yat-sen University, Guangzhou 510060, China*

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.analchem.4c02924

Notes

The authors declare no competing financial interest.

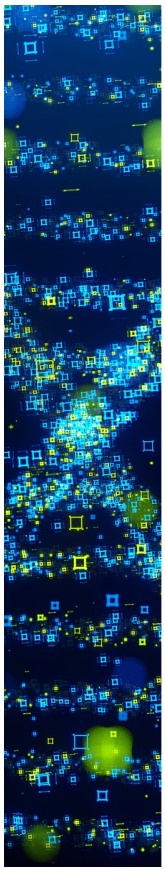
■ ACKNOWLEDGMENTS

We acknowledge the funding support provided by the Research Grants Council-GRF 15305821, CRF Equipment C5033-19E, and RGC-RIF R5050-18. We also acknowledge the support from the Laboratory for Synthetic Chemistry and Chemical Biology Limited (LSCCB) and the Centre for Eye and Vision Research (CEVR) under the Health@InnoHK Programme launched by ITC, HKSAR. We extend our thanks to Prof. Mankin Wong for his supervision and appreciate the technical support from the State Key Laboratory of Chemical Biology and Drug Discovery at PolyU, as well as the PolyU Research Facilities in Chemical and Environmental Analysis (UCEA) and Life Sciences (ULS).

■ REFERENCES

(1) Vitsios, D.; Dhindsa, R. S.; Middleton, L.; Gussow, A. B.; Petrovski, S. *Nat. Commun.* **2021**, *12* (1), 1−14.
(2) Statello, L.; Guo, C.-J.; Chen, L.-L.; Huarte, M. *Nat. Rev. Mol. Cell Biol.* **2021**, *22* (2), 96−118.
(3) Lu, S.; Zhang, J.; Lian, X.; Sun, L.; Meng, K.; Chen, Y.; Sun, Z.; Yin, X.; Li, Y.; Zhao, J.; et al. *Nucleic Acids Res.* **2019**, *47* (15), 8111−8125.
(4) Kung, J. T.; Colognori, D.; Lee, J. T. *Genetics* **2013**, *193* (3), 651−69.
(5) Zhang, S.; Reljić, B.; Liang, C.; Kerouanton, B.; Francisco, J. C.; Peh, J. H.; Mary, C.; Jagannathan, N. S.; Olexiouk, V.; Tang, C.; et al. *Nat. Commun.* **2020**, *11* (1), 1312.
(6) Pauli, A.; Norris, M. L.; Valen, E.; Chew, G. L.; Gagnon, J. A.; Zimmerman, S.; Mitchell, A.; Ma, J.; Dubrulle, J.; Reyon, D.; et al. *Science* **2014**, *343* (6172), No. 1248636.
(7) Guo, B.; Zhai, D.; Cabezas, E.; Welsh, K.; Nouraini, S.; Satterthwait, A. C.; Reed, J. C. *Nature* **2003**, *423* (6938), 456−61.
(8) Matsumoto, A.; Pasut, A.; Matsumoto, M.; Yamashita, R.; Fung, J.; Monteleone, E.; Saghatelian, A.; Nakayama, K. I.; Clohessy, J. G.; Pandolfi, P. P. *Nature* **2017**, *541* (7636), 228−232.
(9) Sandmann, C. L.; Schulz, J. F.; Ruiz-Orera, J.; Kirchner, M.; Ziehm, M.; Adami, E.; Marczenke, M.; Christ, A.; Liebe, N.; Greiner, J.; et al. *Mol. Cell* **2023**, *83* (6), 994−1011.e18.
(10) Morgado-Palacin, L.; Brown, J. A.; Martinez, T. F.; Garcia-Pedrero, J. M.; Forouhar, F.; Quinn, S. A.; Reglero, C.; Vaughan, J.; Heydary, Y. H.; Donaldson, C. *Nat. Commun.* **2023**, *14* (1), 1328.
(11) Kim, M.-S.; Pinto, S. M.; Getnet, D.; Nirujogi, R. S.; Manda, S. S.; Chaerkady, R.; Madugundu, A. K.; Kelkar, D. S.; Isserlin, R.; Jain, S.; et al. *Nature* **2014**, *509* (7502), 575−581.
(12) Thul, P. J.; Åkesson, L.; Wiking, M.; Mahdessian, D.; Geladaki, A.; Ait Blal, H.; Alm, T.; Asplund, A.; Björk, L.; Breckels, L. M.; et al. *Science* **2017**, *356* (6340), No. eaal3321.
(13) Adhikari, S.; Nice, E. C.; Deutsch, E. W.; Lane, L.; Omenn, G. S.; Pennington, S. R.; Paik, Y.-K.; Overall, C. M.; Corrales, F. J.; Cristea, I. M.; et al. *Nat. Commun.* **2020**, *11* (1), 1−16.
(14) Zhang, Y.; Lin, Z.; Hao, P.; Hou, K.; Sui, Y.; Zhang, K.; He, Y.; Li, H.; Yang, H.; Liu, S.; Ren, Y. *J. Proteome Res.* **2018**, *17* (12), 4152−4159.
(15) Zhang, Y.; Lin, Z.; Tan, Y.; Bu, F.; Hao, P.; Zhang, K.; Yang, H.; Liu, S.; Ren, Y. *J. Proteome Res.* **2020**, *19* (1), 401−408.
(16) Zhang, Y.; Zhang, K.; Bu, F.; Hao, P.; Yang, H.; Liu, S.; Ren, Y. *J. Proteome Res.* **2020**, *19* (12), 4857−4866.
(17) Uhlén, M.; Fagerberg, L.; Hallström, B. M.; Lindskog, C.; Oksvold, P.; Mardinoglu, A.; Sivertsson, Å.; Kampf, C.; Sjöstedt, E.; Asplund, A.; et al. *Science* **2015**, *347* (6220), No. 1260419.
(18) Wang, B.; Wang, Z.; Pan, N.; Huang, J.; Wan, C. *International journal of molecular sciences* **2021**, *22* (11), 5476.
(19) Ma, J.; Ward, C. C.; Jungreis, I.; Slavoff, S. A.; Schwaid, A. G.; Neveu, J.; Budnik, B. A.; Kellis, M.; Saghatelian, A. *J. Proteome Res.* **2014**, *13* (3), 1757−1765.
(20) Zhu, Y.; Orre, L. M.; Johansson, H. J.; Huss, M.; Boekel, J.; Vesterlund, M.; Fernandez-Woodbridge, A.; Branca, R. M.; Lehtiö, J. *Nat. Commun.* **2018**, *9* (1), 1−14.
(21) Egertson, J. D; Kuehn, A.; Merrihew, G. E; Bateman, N. W; MacLean, B. X; Ting, Y. S; Canterbury, J. D; Marsh, D. M; Kellmann, M.; Zabrouskov, V.; Wu, C. C; MacCoss, M. J *Nat. Methods* **2013**, *10* (8), 744−746.
(22) Barkovits, K.; Pacharra, S.; Pfeiffer, K.; Steinbach, S.; Eisenacher, M.; Marcus, K.; Uszkoreit, J. *Molecular & Cellular Proteomics* **2020**, *19* (1), 181−197.
(23) Müller, F.; Kolbowski, L.; Bernhardt, O. M.; Reiter, L.; Rappsilber, J. *Molecular & Cellular Proteomics* **2019**, *18* (4), 786−795.
(24) Martinez, T. F.; Lyons-Abbott, S.; Bookout, A. L.; De Souza, E. V.; Donaldson, C.; Vaughan, J. M.; Lau, C.; Abramov, A.; Baquero, A. F.; Baquero, K.; et al. *Cell Metab* **2023**, *35* (1), 166−183.e11.
(25) Wang, H.; Wang, Y.; Yang, J.; Zhao, Q.; Tang, N.; Chen, C.; Li, H.; Cheng, C.; Xie, M.; Yang, Y.; Xie, Z. *Nucleic Acids Res.* **2021**, *49* (11), 6165−6180.
(26) Martinez, T. F.; Chu, Q.; Donaldson, C.; Tan, D.; Shokhirev, M. N.; Saghatelian, A. *Nat. Chem. Biol.* **2020**, *16* (4), 458−468.
(27) Gillet, L. C.; Navarro, P.; Tate, S.; Röst, H.; Selevsek, N.; Reiter, L.; Bonner, R.; Aebersold, R. *Mol. Cell Proteomics* **2012**, *11* (6), No. O111.016717.
(28) Navarro, P.; Kuharev, J.; Gillet, L. C.; Bernhardt, O. M.; MacLean, B.; Röst, H. L.; Tate, S. A.; Tsou, C.-C.; Reiter, L.; Distler, U.; et al. *Nat. Biotechnol.* **2016**, *34* (11), 1130−1136.
(29) Cloutier, P.; Poitras, C.; Faubert, D.; Bouchard, A.; Blanchette, M.; Gauthier, M. S.; Coulombe, B. *J. Proteome Res.* **2020**, *19* (1), 18−27.
(30) Pino, L. K.; Just, S. C.; MacCoss, M. J.; Searle, B. C. *Molecular & Cellular Proteomics* **2020**, *19* (7), 1088−1103.
(31) Kong, A. T.; Leprevost, F. V.; Avtonomov, D. M.; Mellacheruvu, D.; Nesvizhskii, A. I. *Nat. Methods* **2017**, *14* (5), 513−520.

(32) Tsou, C.-C.; Avtonomov, D.; Larsen, B.; Tucholska, M.; Choi, H.; Gingras, A.-C.; Nesvizhskii, A. I. *Nat. Methods* **2015**, *12* (3), 258−264.
(33) Demichev, V.; Messner, C. B.; Vernardis, S. I.; Lilley, K. S.; Ralser, M. *Nat. Methods* **2020**, *17* (1), 41−44.
(34) Ge, W.; Liang, X.; Zhang, F.; Hu, Y.; Xu, L.; Xiang, N.; Sun, R.; Liu, W.; Xue, Z.; Yi, X.; et al. *J. Proteome Res.* **2021**, *20* (12), 5392−5401.
(35) Käll, L.; Canterbury, J. D.; Weston, J.; Noble, W. S.; MacCoss, M. J. *Nat. Methods* **2007**, *4* (11), 923−925.
(36) MacLean, B.; Tomazela, D. M.; Shulman, N.; Chambers, M.; Finney, G. L.; Frewen, B.; Kern, R.; Tabb, D. L.; Liebler, D. C.; MacCoss, M. J. *Bioinformatics* **2010**, *26* (7), 966−968.
(37) Boratyn, G. M.; Camacho, C.; Cooper, P. S.; Coulouris, G.; Fong, A.; Ma, N.; Madden, T. L.; Matten, W. T.; McGinnis, S. D.; Merezhuk, Y.; et al. *Nucleic Acids Res.* **2013**, *41* (W1), W29−W33.
(38) Wu, T.; Hu, E.; Xu, S.; Chen, M.; Guo, P.; Dai, Z.; Feng, T.; Zhou, L.; Tang, W.; Zhan, L.; et al. *Innovation* **2021**, *2* (3), No. 100141.
(39) Toprak, U. H.; Gillet, L. C.; Maiolica, A.; Navarro, P.; Leitner, A.; Aebersold, R. *Molecular & Cellular Proteomics* **2014**, *13* (8), 2056−2071.
(40) Li, K.; Jain, A.; Malovannaya, A.; Wen, B.; Zhang, B. *Proteomics* **2020**, *20* (21−22), No. 1900334.
(41) Gessulat, S.; Schmidt, T.; Zolg, D. P.; Samaras, P.; Schnatbaum, K.; Zerweck, J.; Knaute, T.; Rechenberger, J.; Delanghe, B.; Huhmer, A.; et al. *Nat. Methods* **2019**, *16* (6), 509−518.
(42) Rosenbloom, K. R.; Armstrong, J.; Barber, G. P.; Casper, J.; Clawson, H.; Diekhans, M.; Dreszer, T. R.; Fujita, P. A.; Guruvadoo, L.; Haeussler, M.; et al. *Nucleic Acids Res.* **2015**, *43* (D1), D670−D681.
(43) Waterhouse, A. M.; Procter, J. B.; Martin, D. M.; Clamp, M.; Barton, G. J. *Bioinformatics* **2009**, *25* (9), 1189−1191.
(44) Searle, B. C.; Pino, L. K.; Egertson, J. D.; Ting, Y. S.; Lawrence, R. T.; MacLean, B. X.; Villén, J.; MacCoss, M. J. *Nat. Commun.* **2018**, *9* (1), 1−12.
(45) Bateman, A. (The UniProt Consortium). UniProt: a worldwide hub of protein knowledge. *Nucleic Acids Res.* **2019**, *47* (D1), D506−D515 .
(46) Brunet, M. A.; Lucier, J.-F.; Levesque, M.; Leblanc, S.; Jacques, J.-F.; Al-Saedi, H. R.; Guilloy, N.; Grenier, F.; Avino, M.; Fournier, I.; et al. *Nucleic Acids Res.* **2021**, *49* (D1), D380−D388.
(47) Olexiouk, V.; Van Criekinge, W.; Menschaert, G. *Nucleic acids research* **2018**, *46* (D1), D497−D502.
(48) Hao, Y.; Zhang, L.; Niu, Y.; Cai, T.; Luo, J.; He, S.; Zhang, B.; Zhang, D.; Qin, Y.; Yang, F.; Chen, R. *Brief. Bioinform.* **2017**, *19* (4), 636−643.
(49) Wang, H.; Wang, Y.; Yang, J.; Zhao, Q.; Tang, N.; Chen, C.; Li, H.; Cheng, C.; Xie, M.; Yang, Y.; Xie, Z. *Nucleic Acids Res.* **2021**, *49* (11), 6165−6180.
(50) Yang, Y.; Liu, X.; Shen, C.; Lin, Y.; Yang, P.; Qiao, L. *Nat. Commun.* **2020**, *11* (1), 1−11.
(51) Searle, B. C.; Swearingen, K. E.; Barnes, C. A.; Schmidt, T.; Gessulat, S.; Küster, B.; Wilhelm, M. *Nat. Commun.* **2020**, *11* (1), 1−10.
(52) Lu, Y. Y.; Bilmes, J.; Rodriguez-Mias, R. A.; Villén, J.; Noble, W. S. *Bioinformatics* **2021**, *37* (Supplement_1), i434−i442.
(53) Yang, K. L.; Yu, F.; Teo, G. C.; Li, K.; Demichev, V.; Ralser, M.; Nesvizhskii, A. I. *Nat. Commun.* **2023**, *14* (1), 4539.
(54) Yang, Y.; Wang, H.; Zhang, Y.; Chen, L.; Chen, G.; Bao, Z.; Yang, Y.; Xie, Z.; Zhao, Q. *Mol. Cell Proteomics* **2023**, *22* (1), No. 100480.
(55) Na, C. H.; Barbhuiya, M. A.; Kim, M.-S.; Verbruggen, S.; Eacker, S. M.; Pletnikova, O.; Troncoso, J. C.; Halushka, M. K.; Menschaert, G.; Overall, C. M.; Pandey, A. *Genome Res.* **2018**, *28* (1), 25−36.
(56) Sun, L.; Wang, W.; Han, C.; Huang, W.; Sun, Y.; Fang, K.; Zeng, Z.; Yang, Q.; Pan, Q.; Chen, T.; Luo, X.; Chen, Y. *Mol. Cell* **2021**, *81* (21), 4493−4508.e9.
(57) Li, Y.; Zhang, J.; Sun, H.; Chen, Y.; Li, W.; Yu, X.; Zhao, X.; Zhang, L.; Yang, J.; Xin, W.; Jiang, Y.; Wang, G.; Shi, W.; Zhu, D. *Mol. Ther* **2021**, *29* (4), 1411−1424.
(58) Cloutier, P.; Poitras, C.; Faubert, D.; Bouchard, A.; Blanchette, M.; Gauthier, M.-S.; Coulombe, B. *J. Proteome Res.* **2020**, *19* (1), 18−27.
(59) Coulombe, B.; Cloutier, P.; Pinard, M.; Forget, D.; Poitras, C.; Gauthier, M.-S. *FASEB J.* **2020**, *34* (S1), 1−1.
(60) Gauthier, M. S.; Cloutier, P.; Coulombe, B. *Adv. Exp. Med. Biol.* **2018**, *1106*, 25−36.
(61) Hofman, D. A.; Ruiz-Orera, J.; Yannuzzi, I.; Murugesan, R.; Brown, A.; Clauser, K. R.; Condurat, A. L.; van Dinter, J. T.; Engels, S. A. G.; Goodale, A.; et al. *Mol. Cell* **2024**, *84* (2), 261−276.e18.
(62) Jiang, L.; Wang, M.; Lin, S.; Jian, R.; Li, X.; Chan, J.; Dong, G.; Fang, H.; Robinson, A. E.; Snyder, M. P. *Cell* **2020**, *183* (1), 269−283.e19.
(63) Kaulich, P. T.; Cassidy, L.; Weidenbach, K.; Schmitz, R. A.; Tholey, A. *Proteomics* **2020**, *20* (19−20), No. e2000084.
(64) Kaulich, P. T.; Cassidy, L.; Bartel, J.; Schmitz, R. A.; Tholey, A. *J. Proteome Res.* **2021**, *20* (5), 2895−2903.
(65) Cassidy, L.; Helbig, A. O.; Kaulich, P. T.; Weidenbach, K.; Schmitz, R. A.; Tholey, A. *J. Proteomics* **2021**, *230*, No. 103988.



CAS BIOFINDER DISCOVERY PLATFORM™
STOP DIGGING THROUGH DATA —START MAKING DISCOVERIES
CAS BioFinder helps you find the right biological insights in seconds
Start your search
CAS
A division of the American Chemical Society

https://doi.org/10.1021/acs.analchem.4c02924
Anal. Chem. 2025, 97, 1517−1527