Published online 17 December 2020

Nucleic Acids Research, 2021, Vol. 49, No. 5 ***e28***
doi: 10.1093/nar/gkaa1194

FAX-RIC enables robust profiling of dynamic RNP complex formation in multicellular organisms *in vivo*

Yongwoo Na[1,2], Hyunjoon Kim[1,2], Yeon Choi[1,2], Sanghee Shin[1,2], Jae Hun Jung[3], S. Chul Kwon[1,2], V. Narry Kim [1,2,*] and Jong-Seo Kim [1,2,*]

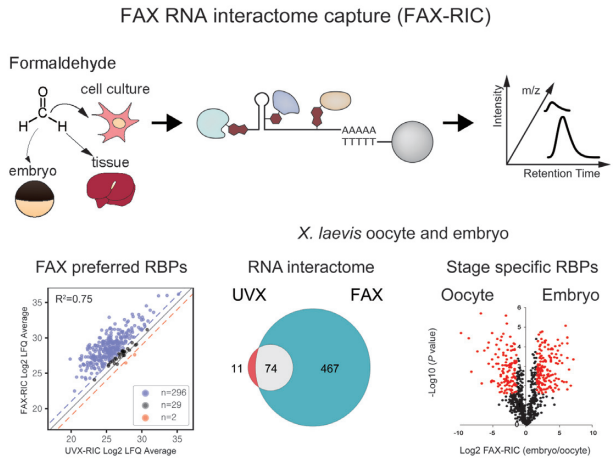
[1]Center for RNA Research, Institute for Basic Science, Seoul 08826, Korea, [2]School of Biological Sciences, Seoul National University, Seoul 08826, Korea and [3]Department of Applied Chemistry, Kyung Hee University, Yongin 17104, Korea

Received March 17, 2020; Revised November 11, 2020; Editorial Decision November 20, 2020; Accepted November 24, 2020

ABSTRACT

RNA–protein interaction is central to post-transcriptional gene regulation. Identification of RNA-binding proteins relies mainly on UV-induced crosslinking (UVX) followed by the enrichment of RNA–protein conjugates and LC-MS/MS analysis. However, UVX has limited applicability in tissues of multicellular organisms due to its low penetration depth. Here, we introduce formaldehyde crosslinking (FAX) as an alternative chemical crosslinking for RNA interactome capture (RIC). Mild FAX captures RNA–protein interaction with high specificity and efficiency in cell culture. Unlike UVX-RIC, FAX-RIC robustly detects proteins that bind to structured RNAs or uracil-poor RNAs (e.g. AGO1, STAU1, UPF1, NCBP2, EIF4E, YTHDF proteins and PABP), broadening the coverage. Applied to *Xenopus laevis* oocytes and embryos, FAX-RIC provided comprehensive and unbiased RNA interactome, revealing dynamic remodeling of RNA–protein complexes. Notably, translation machinery changes during oocyte-to-embryo transition, for instance, from canonical eIF4E to noncanonical eIF4E3. Furthermore, using *Mus musculus* liver, we demonstrate that FAX-RIC is applicable to mammalian tissue samples. Taken together, we report that FAX can extend the RNA interactome profiling into multicellular organisms.

GRAPHICAL ABSTRACT



INTRODUCTION

From synthesis to decay, messenger RNA (mRNA) is accompanied by RNA-binding proteins (RBPs) that govern mRNA export, translation, localization and decay (1). To understand the RNA–protein network, several methods have been developed to profile RBPs at the proteomic scale (2–6). RNA interactome capture (RIC) method is based on the oligo-dT bead pulldown and mass spectrometry of the crosslinked RNA–protein conjugates (2–4). Development and application of the RIC technique has significantly expanded the RBP repertoire by discovering hundreds of unorthodox RBPs without any known RNA-binding domains (RBDs) or RNA-related functions (2–4,7,8). Introduction of the RIC also led to the development of related methods to profile the interactome of poly(A)-less RNAs in a sequence-specific manner, and for newly synthesized RNAs that are metabolically labeled (5,6,9–11) (Supplementary Table S1). In most of such RNA interactome capture studies, crosslinking is induced by ultraviolet light (254 or 365 nm) irradiation of cells (2–9,11). However, UV crosslink-

*To whom correspondence should be addressed. Tel: +82 2 880 2698; Fax: +82 2 887 0244; Email: jongseokim@snu.ac.kr
Correspondence may also be addressed to V. Narry Kim. Tel: +82 2 880 9120; Fax: +82 2 887 0244; Email: narrykim@snu.ac.kr

ing (UVX) has an innate weakness owing to its highly limited penetration depth (12,13). With UVX-RIC, it is hard to profile the RNA interactome of large or opaque samples due to its low efficiency and the inevitable bias toward the molecules on the surface, as pointed out previously in a UVX-based study on *Drosophila* embryo (13). Thus, it is necessary to develop alternative crosslinking strategies to achieve the comprehensive and sensitive RIC in multicellular organisms.

Formaldehyde crosslinking (FAX) can be a promising alternative because of the high membrane permeability of formaldehyde (14). Although FAX has been used to characterize RNA–protein (RNP) complexes in a number of studies (10,15–19), the primary concern has been its selectivity (19,20) because it is generally thought that formaldehyde crosslinks promiscuously many biomolecules with various functional groups. However, the reported mechanisms of FAX indicate that formaldehyde is highly selective to crosslinking between nucleophiles such as amines via Schiff base formation and nucleophilic addition (21,22). The majority of RNA nucleobases retain exo-amino groups, and the RNA-binding motifs often contain amino groups such as lysine (7,23). Formaldehyde is a small chemical crosslinker with short molecular span which is thought to be ∼2 Å apart (21). Such knowledge suggests that amino acids closely located to the nucleobases of RNA can be preferentially crosslinked by formaldehyde treatment, compared to random and transient interactions with other macromolecules. Nevertheless, FAX has only been used sparsely in RNA biology and the use has also been largely restricted to the probing of structured or duplex RNA–protein interactions (15,24,25). Therefore, both the relative specificity and many of the potential benefits of FAX for probing *in vivo* RNA–protein interactions remained to be investigated through comprehensive system-wide analysis.

Here, we report that FAX can enable comprehensive and reliable RIC studies within the diverse biological systems, from cultured cells to *X. laevis* oocyte and embryo, and *M. musculus* liver. Systematic and quantitative comparison between RNA interactome profiles from the FAX-RIC and UVX-RIC in HeLa cells disclosed the distinct characteristics of two crosslinking methods and suggested the relatively high specificity of FAX-RIC. We further demonstrate the specificity of FAX to RNA binding domains, by developing a peptide-level FAX-RIC protocol. Using FAX-RIC, we profiled the changes in the RNA interactome landscape during oocyte-to-embryo transition (OET) for the first time *in vivo*. Furthermore, we tested the applicability of the FAX-RIC approach to mammalian tissues by utilizing mouse liver as a model. The findings of this study will significantly broaden our understanding of mRNP complex remodeling in multicellular organisms *in vivo*.

MATERIALS AND METHODS

HeLa cell culture

HeLa cells were cultured in DMEM (Welgene) supplemented with 9% fetal bovine serum (Welgene) and maintained in a humid incubator at 37°C in a 5% CO_2 environment to reach a density of 1×10^6 cells/ml. The HeLa cell line is a modified HeLa with TUT4 gene deletion.

Formaldehyde crosslinking (FAX) and RNA interactome capture (RIC) in HeLa Cell

Formaldehyde treatment condition was first optimized to achieve kinetically controlled crosslinking in HeLa cell. Formaldehyde in phosphate buffered saline (PBS), prepared with 16% formaldehyde (w/v), methanol-free (Thermo Scientific), with increasing concentration of 0.1, 0.3, 0.5, 0.7 and 1.0% were applied to HeLa cell for 5 min and oligo-dT beads (NE Biolabs) enriched protein amount from each condition was checked by SDS page gel running and silver staining. For respective replicate of RIC experiment, HeLa cells were plated on the dishes with surface area of 5×145 cm^2 to reach ∼90% cell confluence overnight, resulting in the cell number of $\sim 12.5 \times 10^7$ before crosslinking. FAX was done by directly applying formaldehyde in PBS to HeLa cells on plate. Briefly, cells were washed twice with PBS at room temperature (RT) and incubated with 0.5% formaldehyde in PBS for 5 min at RT. Formaldehyde solution was removed at 5 min and the cells were washed twice with ice-cold 200 mM Tris in PBS, with 30 s incubation time each, to quench the residual formaldehyde reaction. Cells were collected from the plate by scraping and washed twice with ice cold PBS by centrifuge. Cells were lysed in the lysis buffer (0.5% (w/v) lithium dodecyl sulfate (LDS), 500 mM lithium chloride (LiCl), 10 mM Tris (pH 7.5), 2 mM EDTA, and 5 mM dithiothreitol (DTT) (all Sigma)) and sheared by passing through 21-gauge needle syringe for ten times. Oligo-dT bead was added to the lysate and incubation were done with over and over rotation for 1 hour at 4°C. Beads were separated from the cell lysate using DynaMag (Thermo scientific) and washed twice in each buffer, the lysis buffer, low LDS lysis buffer (lysis buffer made with 0.1% (w/v) LDS), high salt buffer (lysis buffer made without LDS) and low salt buffer (lysis buffer made without LDS and 200 mM LiCl). Oligo-dT beads were then incubated with Turbo DNase (Thermo Scientific) in Turbo DNase buffer supplemented with 200 mM of LiCl, for 30 min with over and over rotation at RT. Beads were then washed twice with each wash buffer. Elution of the poly(A) RNA by heat was done twice in TE buffer, by incubation on Thermomixer C (Eppendorf) at 65°C for 3 min with mixing at 800 rpm. Peptide sample preparation for LC–MS/MS analysis were done with Microcon-30 kDa Centrifugal Filter Unit (Millipore), following the previously described protocol (3). Final peptide sample was desalted using Discovery DSC-18 SPE Tube (Supelco) and prepared for LC–MS/MS analysis.

Peptide-level FAX-RIC

Formaldehyde treatment and cell lysate preparation was done as described previously for protein-level FAX-RIC. HeLa cell lysate was then diluted with Tris EDTA (TE) buffer to make LDS 0.07% (w/v) and LiCl 70 mM. MS grade trypsin (Thermo scientific) was added to the cell lysate at protein weight ratio of 1:100 followed by incubation with over and over rotation for 8 h at RT. The cell lysate was then made to contain 0.5% LDS and 500 mM LiCl for the oligo-dT bead pulldown. Oligo-dT bead pulldown and washing steps were done as described previously for protein-level FAX-RIC. RNAs were eluted from the bead by adding 8 M urea in TE buffer, twice. Samples were then filtered

down four times with 8 M urea TE buffer and three times with TE buffer in 100 kDa Amicon filter (Millipore). The sample in TE buffer was incubated at 65°C on Thermomixer C overnight. The samples were subjected to the conventional trypsin digestion protocol for LC-MS/MS analysis and desalted using Millipore ZipTip with C18 resin.

RIC via UV crosslinking (UVX) in HeLa Cell

Cell culture condition and scale was identical to that described previously for the FAX-RIC. For UVX, cells were washed twice with PBS at RT and after removing the PBS, irradiated with 450 mJ (∼60 s) of UV light (254 nm), using the Spectrolinker XL-1500 UV crosslinker (Spectronics) on ice, as it was previously described (3). Following procedures for oligo-dT capture were identical to the above described procedure for the FAX-RIC.

Xenopus laevis oocyte and embryo preparation

Xenopus laevis oocytes were obtained from excised ovary of female *X. laevis* as previously described (26). *X. laevis* embryos collection were also done as described previously (27). Briefly, human chorionic gonadotropin was injected into a female frog 12 h before collecting eggs. The eggs were obtained in 1X Marc's Modified Ringer's (MMR) solution and *in vitro* fertilized using excised testes from a male frog.

RNA interactome profiling in *X. laevis* oocyte and embryo

X. laevis embryo collected as described above were washed three times with PBS and then treated with 2% formaldehyde in PBS for 10 min with gentle rotation. Formaldehyde reaction was quenched by treating the embryo with 200 mM tris in PBS for 5 min and then washed with ice cold PBS three times. Crosslinked *X. laevis* embryo was first lysed with high salt lysis buffer (0.5% lithium dodecyl sulfate, 1 M LiCl, 20 mM Tris (pH 7.5), 5 mM DTT, 1 mM EDTA), and then 8 M urea lysis buffer (8 M urea, 0.5% lithium dodecyl sulfate, 20 mM Tris (pH 7.5), 5 mM DTT, 1 mM EDTA) was added to adjust the concentration of urea and lithium chloride in sample lysate to become 4 M and 500 mM, respectively. We found that use of high salt lysis buffer and the urea in sample lysis buffer is crucial to the integrity of RNA for the preparation of *X. laevis* oocyte and embryo lysate (data not shown). RIC experiment in *X. laevis* oocyte and embryo were all performed in triplicate. Following procedures for RIC was identical to the above described procedure for the HeLa cell. FAX-RIC for *X. laevis* oocyte was identical to that of the protocol for the embryo. For UVX, prepared *X. laevis* oocytes and embryos were placed in minimal amount of PBS to cover the whole embryo on plate and irradiated with 500 mJ of UV light (254 nm) on ice for four times, with agitation of plate after each irradiation to turn around the embryo, resulting in total 2 J of UV light irradiation. Embryos were washed with PBS twice after UV light irradiation. Embryo lysis and oligo-dT capture procedure was identical to the above described procedure for *X. laevis* FAX-RIC.

TMT based quantitative profiling of global protein expression level in *X. laevis* oocyte

Protein samples in 0 and 8 h after the progesterone induced oocyte maturation were prepared as previously reported (28). Instrumental settings for LC–MS3 analysis and data analysis parameters were largely the same as the case of our previous work (29). The final protein quantification results were taken from the biological duplicate experiments.

RNA interactome profiling in mouse liver using peptide-level FAX-RIC

Mice liver samples were a kind gift from Hyun-Woo Rhee Lab (SNU). Animal experiments with C57BL/6 mice, all male and 8–10 weeks of age, were performed in accordance with the governmental and institutional laws and recommendations (Approval no. SNU-180521-2-3). Mouse liver tissue was cut by a lobe and submerged in the formaldehyde solution of 4% in PBS for 10 minutes with occasional shaking. Crosslinking reaction was quenched by submerging the tissue in 200 mM Tris (pH 7.0) in PBS buffer for 5 min. The tissue was weighed on scale and ∼20 mg of liver samples was used for each of three replicate dT pull down experiments and total RNA interactome extraction experiments via RNeasy column (Qiagen). The tissue was first lysed in 4 M GuSCN, 800 mM LiCl, 10 mM DTT, 5 mM EDTA and 20 mM HEPES (pH 7.5), and 3 volumes of ethanol was added to the sample followed by 1 h incubation at −20°C. Precipitated samples were centrifuged down at 16 000 g for 15 min and supernatant was removed followed by 70% ethanol wash twice at 8000 g. The sample was reconstituted with the 0.05% LDS, 100 mM LiCl, 10 mM DTT, 5 mM EDTA, and 20 mM Tris (pH 7.5) and trypsin was added 1:50 (w/w) and incubated at RT for 4 h. Prepared lysate was subjected to peptide-level FAX-RIC protocol as described above for HeLa cells. For total RNA interactome experiment, after the trypsin treatment twice the volume of RLT buffer was added to the sample and RNA sample preparation by RNeasy column was done as described in the product manual followed by heat incubation 65°C overnight and LC–MS/MS analysis.

LC–MS/MS analysis

Analytical capillary columns (100 cm × 75 μm i.d.) and trap columns (3 cm × 150 μm i.d) were packed in-house with 3 μm Jupiter C18 particles (Phenomenex, Torrance, CA). The long analytical column was placed in a column heater (Analytical Sales and Services, Pompton Plains, NJ) regulated to a temperature of 45°C. Ultimate 3000 nanoRSLC system (Thermo Scientific, Sunnyvale, CA) was operated at a flow rate of 350 nl/min over 2 h with linear gradient ranging from 95% solvent A (water with 0.1% formic acid) to 40% of solvent B (acetonitrile with 0.1% formic acid). The enriched samples were analyzed on an Orbitrap Fusion Lumos mass spectrometer (Thermo Scientific) equipped with an in-house customized nanoelectrospray ion source. Precursor ions were acquired (m/z 300–1500 at 120k resolving power and the isolation of precursor for MS/MS analysis was performed with a 1.4 Th. Higher-energy collisional dissociation (HCD) with 30% collision energy was used for

sequencing with a target value of 5E4 ions determined by automatic gain control. Resolving power for acquired MS2 spectra was set to 30k at m/z 200 with 150 ms maximum injection time.

Protein identification

MS raw data files were processed with MaxQuant (version 1.5.3.30) (30). Enzyme specificity was set to trypsin/P and a maximum of two missed cleavages were allowed. Cysteine carbamidomethylation and methionine oxidation were selected as fixed and variable modifications, respectively. The derived peak list was searched using the built-in Andromeda search engine in MaxQuant against the human UniProt database (version 2/2018) or '*X. laevis* protein (Xenbase)' fasta file (2/14/2018 version) downloaded from Xenbase website. Initial maximal allowed mass tolerance was set to 20 ppm for peptide masses, followed by 6 ppm in the main search, and 0.5 Da for fragment ion masses. The minimum peptide length was set to six amino acid residues, and three labelled amino-acid residues were allowed. A false discovery rate (FDR) <0.01 was required at both the protein-level and the peptide-level. Label free quantification was turned on and applied to each experimental group separately, except for the comparison between *X. laevis* oocyte and embryo FAX.

Statistical analysis for label-free quantification (LFQ)

Statistical analysis for defining RNA interactome and quantitative comparison between each experiments were done with PERSEUS software (31). For defining the RNA interactome, protein groups with LFQ value in two or more replicates in each experiment were used for statistical analysis. Missing LFQ values were imputed with a normal distribution shifted by −4 and sharpened with a standard deviation factor of 0.3. Student's *t*-test was performed to test if any $\log_2$ fold-change ratio is different from 0 and protein groups with Benjamini Hochberg FDR <0.01 were considered as RNA interactome. Same statistical analysis procedure was used to test for differentially captured protein groups between individual RNA interactome. Protein groups with Benjamini Hochberg FDR <0.05 were considered to be significantly enriched in certain RNA interactome.

Gene ontology (GO) and domain analysis

GO and PFAM domain annotations of the identified proteins were retrieved using ENSEMBL Biomart and the Ensembl Human release 81 (GRCh38.p3). GO term enrichment analysis was done using the DAVID tool (32,33). Human orthologues of *X. laevis* proteins were retrieved from UniProt database by matching the gene name of *X. laevis* proteins with that of human proteins.

SDS–PAGE and western blotting

Input cell lysate and oligo-dT enriched samples were all treated with benzonase and RNase A/T1 and sonicated for 15 min by Bioruptor (COSMO BIO). Protein samples were separated on Bolt 4–12% Bis–Tris Plus Gel (Thermo Fisher Scientific) and transferred to polyvinylidene fluoride membrane (GE Healthcare). Western blotting was performed as previously described (34). The following antibodies were used for western blotting at 1:1000 dilution in PBS containing 1% skim milk and 0.1% Tween 20 (USB): anti-AGO1 (Cell signaling), anti-EIF4E (Cell signaling), anti-PABPC1 (Gift from Dr. Dreyfuss' lab), anti-Tubulin (Abcam), DDX19B (Novus Bio), GAPDH (Santa Cruz), and ENO1 (Proteintech Group).

qPCR analysis

Oligo-dT pulled down of RNAs were all treated with proteinase K in proteinase K buffer (0.5% SDS, 100 mM NaCl, 1 mM EDTA and 50 mM Tris (pH 7.5) at 65°C for 4 h. For total RNAs, HeLa cell lysate in lysis buffer for oligo-dT pull down was treated with the proteinase K at 65°C for 4 h. The RNAs were then purified by the Trizol reagent (Invitrogen) and reverse-transcribed with 5× RT Master Mix (Takara) and the RNA levels were measured with SYBR Green assays (Applied biosystems) with primers against 18S rRNA (forward: GAAACTGCGAATGGCTCATT AAA, reverse: CACAGTTATCCAAGTGGGAGAGG), EEF2 (f: AACTTCACGGTAGACCAGATCC, r: TCGT CCTTCCGGGTATCAGTG), and TS (f: GGCAGAAT ACAGAGATATGGAATCAGA, r: TCGTCAGGGTTG GTTTTGATG).

RESULTS

FAX-RIC profiles known RBPs with high specificity in HeLa cells

In order to find the optimal FAX condition for RNA–protein interactions, we screened a series of mild FAX conditions in HeLa cells and the RNP was pulled down with oligo-dT (Figure 1A). The total amount of the precipitated proteins was significantly greater than that of UVX-RIC even at concentrations as low as 0.1–0.5% formaldehyde (w/v) for 5 min (Figure 1B). The protein profiles were relatively constant across the tested FAX conditions in comparison with UVX-captured protein profiles (Figure 1B). Through the comparative qPCR analysis of the input and oligo-dT enriched RNA samples, we found that FAX did not impair the RNA pull down efficiency nor the specificity (Supplementary Figure S1).

We performed western blot analysis for the initial assessment of the relative specificity and the enhanced efficiency of FAX-RIC to the representative RBPs along with a negative control protein, Tubulin A (Figure 1C). Moreover, when the RBPs were eluted using RNase A/T1 treatment, instead of heat treatment, to prevent potential reversal of formaldehyde crosslinking, the western blot bands did not shift upward (Supplementary Figure S2). This suggested that their enrichment via FAX-RIC is dependent on RNA–protein crosslinking rather than protein-protein crosslinking.

We carried out quantitative proteomic profiling of FAX-captured proteins (Supplementary Figure S3). A total of 912 proteins passed the LFQ intensity-based quantitative filtering criteria against the no crosslinking (NoX) control

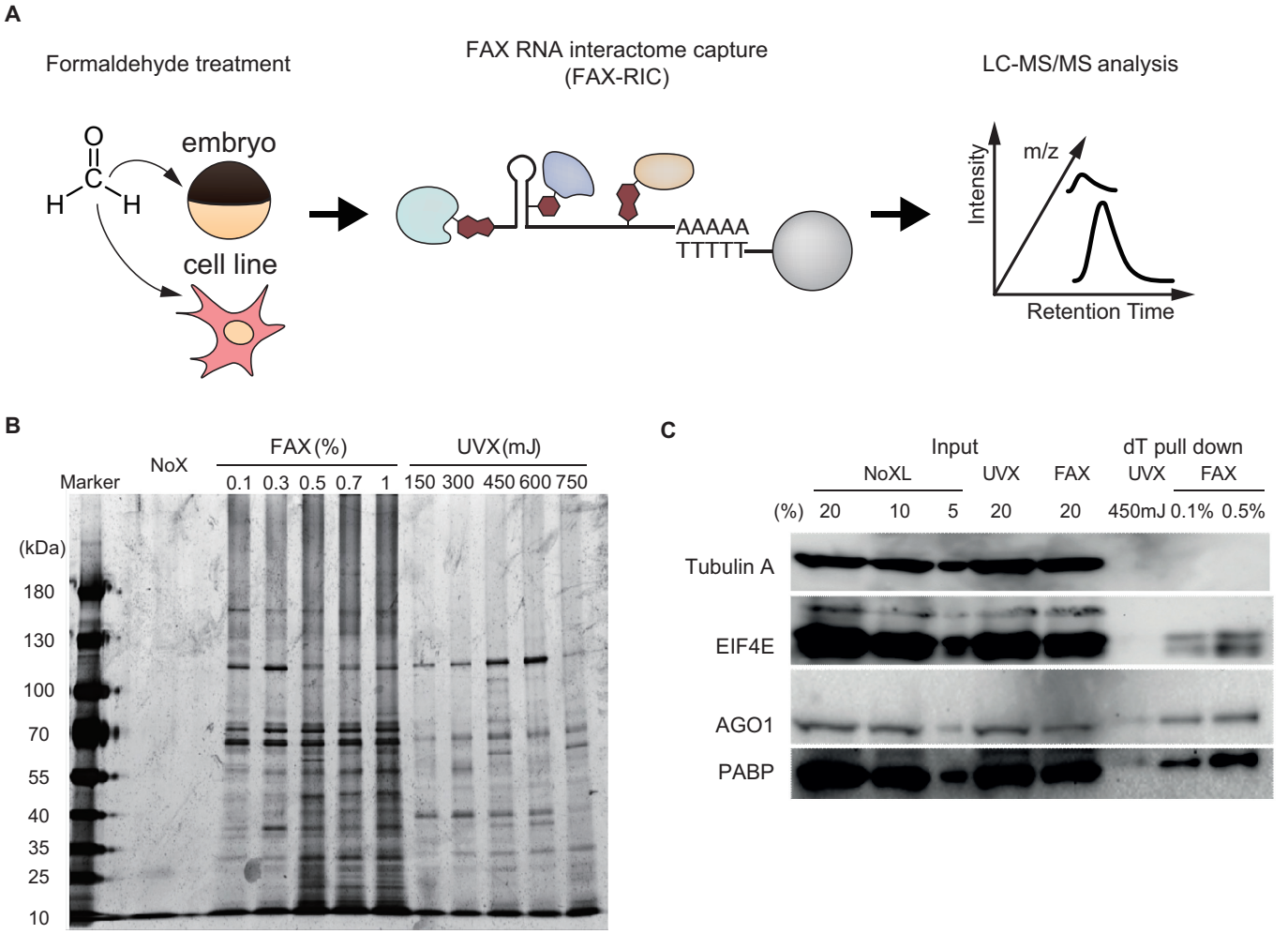


Figure 1. Development of FAX-RIC and its specificity for the representative RBPs in HeLa cell. (**A**) Schematic outline of the FAX based RNA interactome capture (FAX-RIC) method. Biological samples are treated with formaldehyde solution in PBS at conditions that are separately optimized to form covalent bond between proximal RNA–protein interactions. RNA crosslinked proteins are enriched through oligo-dT pulldown of poly(A) tailed RNA. Profile of enriched protein samples are obtained via LC-MS/MS analysis. (**B**) SDS-PAGE and silver staining of the oligo-dT pulldown samples from the lysate of HeLa cells that are treated with indicated formaldehyde and UV light crosslinking conditions. (**C**) Western blot analysis for representative RBPs (EIF4E, AGO1 and PABP) and a negative control protein (Tubulin A).

at a <0.01 false discovery rate (FDR). These proteins are regarded as our 'FAX RNA interactome' in HeLa cells (Figure 2A). We found that >94% (861 proteins) of the FAX RNA interactome have been reported as RBPs in previous RBP profiling studies (3–7) (Figure 2B). The proportions of RBPs with known RBDs (3) (Figure 2C) or those detected by RBDmap (2,5,6) (Figure 2D) were comparable between our FAX RNA interactome and the combined list of the UVX RNA interactomes (3,4,7), which we referred to as 'REF-UVX', demonstrating that FAX-RIC can capture both conventional RBPs and unorthodox RBPs.

Peptide-level FAX-RIC

Unlike UV-induced crosslink, formaldehyde-induced crosslink can be reversed at high temperature (19). Taking advantage of this property, we modified FAX-RIC by treating cell lysates with trypsin before oligo-dT pulldown (Supplementary Figure S4A). Subsequently, the peptide fragments directly linked to RNA were purified via oligo-dT pulldown, size selective filtration, and de-crosslinking by heat. The non-crosslinked peptides are depleted via stringent washing and filtration. This method allows us to identify direct RNA-binding regions of RBPs.

The results from peptide-level FAX-RIC were highly reproducible between replicates (Supplementary Figure S4B), resulting in 382 significantly enriched proteins (compared to the NoX control; <0.05 FDR) (Supplementary Figure S4C). Over 80% of RNA interactome obtained by peptide-level FAX-RIC overlapped with that of the protein-level FAX-RIC. We compared the protein profile from both FAX-RICs to that from RBDmap (23) (Supplementary Figure S4D). RBDmap determines the potential RBDs through the identification of the peptides which are adjacent to the RNA-crosslinked sites by UVX. Our peptide-level FAX-RIC and the UVX-based RBDmap identified ∼200 RBPs exclusively. Incomplete overlap of the protein

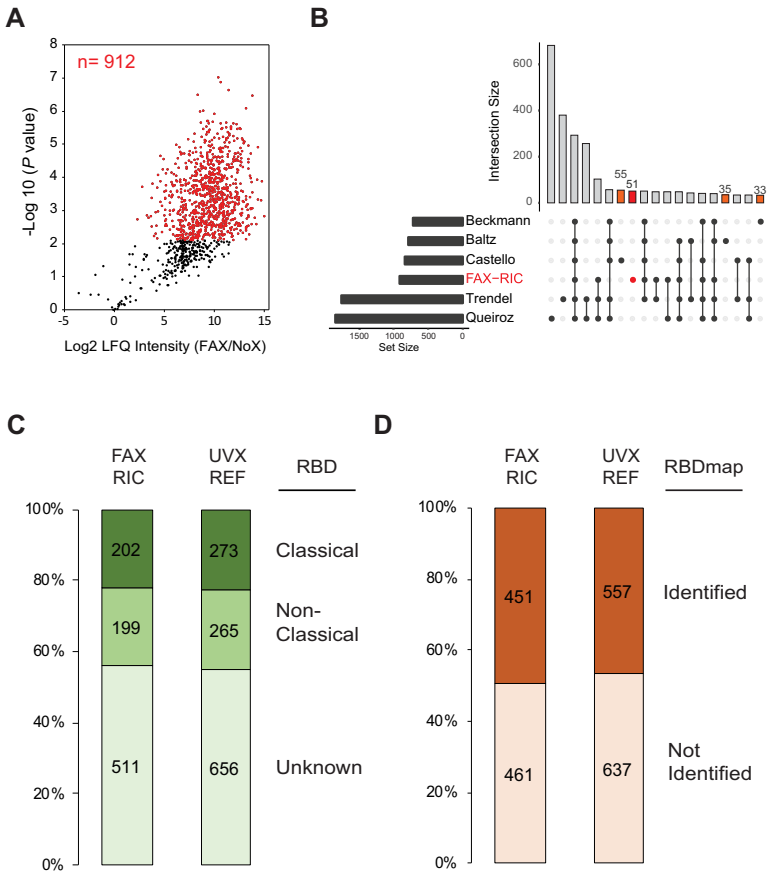


Figure 2. FAX-RIC profile known human RBPs with high specificity. (**A**) Defining the high confidence FAX RNA interactome in HeLa cell. Volcano plot displaying the fold-change of average LFQ intensity (FAX-RIC/NoX-RIC) (x-axis) and the $-\log_{10}$ Student's *t*-test *P* value (y-axis) for all the proteins quantified in at least two out of three replicate FAX-RIC experiment. Proteins with $\log_2$ fold-change >1 and statistically significant enrichment over the NoX-RIC (*P* value < 0.01, Student's t–test, adjusted by Benjamini–Hochberg method) are highlighted in red. (**B**) UpSet plot for the number of proteins that are identified in indicated group of RBP profiles obtained from our FAX-RIC in HeLa cell and previous UVX based RBP profiling studies (3–7). Each bar on the plot represent the number of proteins that were identified in single or multiple RNA interactome profiles denoted by the black dots below for respective studies whose name and the number of identified proteins are indicated on the left column. For example, the first and second bar on the plot represent the number of RBPs that were exclusive to Trendel *et al.* (5) and Queiroz *et al.* (6) study, respectively. The third bar represent the number of RBPs that were common to all six RNA interactome profile, indicated by the 6 black dots joined by a solid line. Number of the proteins that are exclusive to FAX-RIC and three representative RIC experiments (3,4,7) are highlighted in red or orange, respectively. (**C**) Number and proportion of proteins annotated with the known RBDs, either 'classical' or 'non-classical' as defined previously (3). 'UVX-REF' include all the proteins identified in three representative RIC experiments (3,4,7). (**D**) Composition of the proteins with or without RNA interacting region defined in previous RBDmap based studies (2,5,6).

profile from the two methods suggest that the use of two RBD profiling methods using different crosslinking method may complement each other for the deeper coverage of RNA binding region in RBPs. Compared with the protein-level FAX-RIC, peptide-level FAX-RIC showed a higher proportion of peptides mapped within known RBDs or RNA-interacting regions experimentally defined by the RBDmap approach (23) (Supplementary Figure S4E). Notably, the peptide-level FAX-RIC offers higher resolution than RBDmap (Supplementary Figure S5). The identified peptides are often found within classical RBDs such as CSD, dsrm, KH and RRM domains, but they also reveal novel RNA–protein interaction sites within other structural features such as intrinsically disordered regions (Supplementary Figure S5A–D). These results provide a further basis for the FAX specificity to proximal and stable RNA–protein interactions within the known RBDs, and demonstrate a potential of peptide-level FAX-RIC for the discovery and comparative analysis of the RNA-binding motifs within the RBPs.

Quantitative comparison of UVX- and FAX-RIC

To further characterize the merits and efficacy of FAX as an alternative *in vivo* crosslinking method in terms of capture specificity and efficiency, we performed a quantitative comparison of the UVX- and FAX-captured proteins based on the LFQ intensity. For fair comparison, our UVX RNA interactome (657 proteins) was generated according to the conventional UVX-RIC protocol at an identical sampling scale to that of the FAX-RIC protocol (Supplementary Figure S6A). FAX-RIC enabled more comprehensive profiling, generating ∼40% more proteins than UVX-RIC did (Figure 3A). Consistent with our previous analysis on the FAX-RIC's specificity to the known RBPs (Figure 2C, D), we found that both UVX- and FAX-RIC had similar quantitative enrichment rate (LFQ intensity from the RIC di-

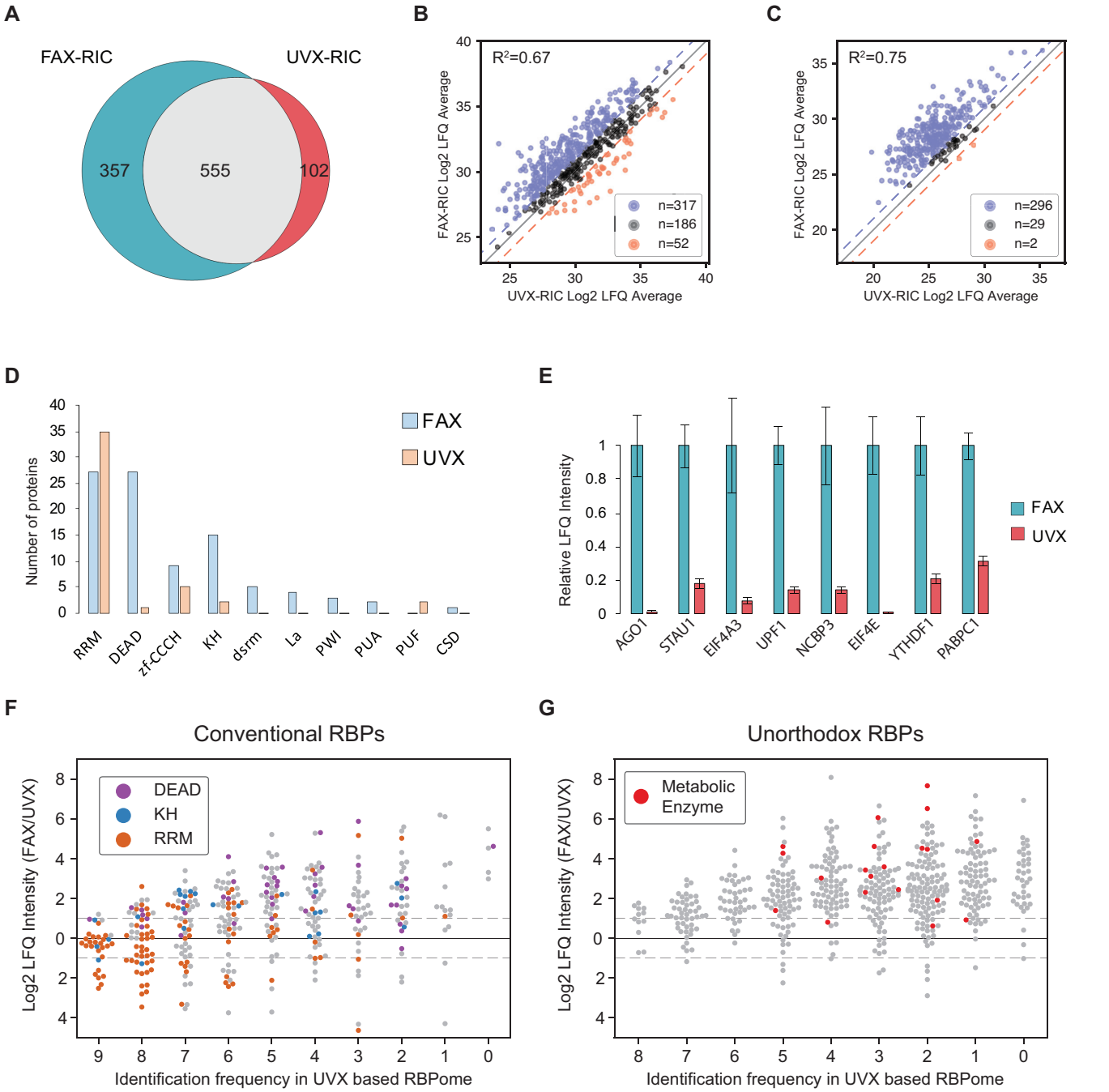


Figure 3. FAX-RIC enable more comprehensive profiling of RBPs through the enhanced capture efficiency. (**A**) Overlap between the UVX- and FAX-RIC based high confidence RNA interactome profiles. (B, C) Scatter plot of average LFQ intensity between UVX- and FAX-RIC experiments, drawn for RBPs, common to UVX and FAX RNA interactome (**B**) and exclusive to FAX RNA interactome (**C**). (**D**) Number of proteins with significantly greater LFQ intensity (>1 $\log_2$ fold-change) in UVX- or FAX-RIC, annotated with indicated classical and non-classical RBDs. (**E**) Relative LFQ intensities from the UVX- and FAX-RIC for the representative RBPs. Error bars represent mean standard error from three independent experiments. (F, G) Average LFQ intensity fold-change (FAX/UVX) for the RBPs, annotated with either 'classical' or 'non-classical RBDs' (**F**) and not annotated with such well characterized RBDs (**G**). RBPs were grouped by the identification frequency in total of nine UVX based RBP profiling studies (2,5,6).

vided by the relative protein intensity in the total proteome) for the RBPs with known RBDs or those defined via 'RBDmap' (2,5,6) (Supplementary Figure S6B, C). No significant change in the relative amount of ribosomal proteins obtained by UVX- and FAX-RIC, suggest that there was no significant increase in the rRNA binding proteome contamination in the FAX-RIC and thus they had comparable specificity to the poly(A)+ RNA interactome (Supplementary Figure S6B). This quantitative comparison showed that FAX-RIC profiled a larger number of RBPs with specificity comparable to that of UVX-RIC.

Next, we expanded the quantitative comparison to the whole identified RNA interactome, i.e., including the unorthodox RBPs. In the 555 common RNA interactome proteins, 57% of the proteins were more efficiently captured (>2-fold) by FAX-RIC, whereas only 9.3% of proteins showed better efficiency in UVX-RIC (Figure 3B). As expected, over 90% of the FAX-exclusively determined RNA interactome showed a higher intensity value over that obtained by UVX-RIC (Figure 3C). Thus, these RBPs may not have been detected by the UVX-RIC due to the low efficiency of UVX. Of note, we found a strong correlation between protein intensity signals obtained via UVX- and FAX-RIC. Taken together, our quantitative analyses demonstrated that FAX-RIC could profile RNA interactome robustly with significantly enhanced capture efficiency.

Our system-wide comparison between UVX- and FAX-RIC generated a quantitative UVX- or FAX-preferred RBP list (> 2-fold relative ratio; 82 or 613 proteins for UVX- or FAX-preferred RBPs, respectively; Supplementary Table S2). Using this protein list, we further tried to dissect the molecular characteristics of both methods based on the known RBD information of those RBPs, since the different crosslinking mechanism of each method should be preferential to distinct modes of RNA–protein interactions. Notably, we found a strong overrepresentation of the RRM domain in UVX-preferred RBPs (Figure 3D) that was consistent with the known preference of UVX on uracil-aromatic amino acid residues (35), a major contributor to the affinity between RRM domains and RNAs (36,37). On the contrary, FAX-preferred RBPs included a broad range of canonical RBPs annotated with the classical RBDs, such as RRM, KH, DEAD, La and PWI domains (Figure 3D). Furthermore, we found that the representative RBPs with preferences for distinct RNA sequences or structures were profiled with significantly greater protein intensity by FAX-RIC. These included RBPs associated with RNA duplexes (e.g. AGO1 and STAU1), helicases (e.g. EIF4A3 and UPF1), and uracil-poor RNA sequences such as the mRNA cap (e.g. NCBP3 and EIF4E), N6-methyladenosine (e.g. YTHDF1/2/3), and poly-adenosine (e.g. PABP1/2 and ZC3H14) (Figure 3E and Supplementary Table S2). It is important to note that the RBPs such as AGO1 and EIF4E were conspicuously missing from many of the previous interactome lists, as well as our UVX RNA interactome list (Supplementary Table S2).

We classified our RNA interactome into two groups; conventional RBPs annotated with the known RBDs and unorthodox RBPs, not annotated with the known RBDs. The relative capture efficiencies of FAX over UVX in terms of LFQ intensity were represented via one-dimensional scatter plotting (i.e., beeswarm plot) versus identification frequency among the nine UVX-based RBP profiling studies (2–6) in human cancer cell lines for conventional (Figure 3F, Supplementary Table S3) and unorthodox RBPs (Figure 3G, Supplementary Table S3). The plots showed that less frequently identified RBPs in UVX-RIC studies were more favorably captured by FAX-RIC in both conventional and unorthodox RBPs. Most notably, RBPs with an RRM domain were identified with the most predominant frequency, consistent with the quantitatively UVX-preferred character of those proteins in our data (Figure 3D). In contrast, RBPs with other classical RBDs such as the DEAD and KH domains were identified with relatively low frequency (Figure 3F). Strong overrepresentation of RRM domain in the UVX-based studies highlight the strong bias of UVX method toward a particular RBD (Figure 3F). Presumably owing to the bias of UVX, the significant underrepresentation of the FAX-preferred RBPs with the well-defined RBD annotation such as KH and DEAD in the previous UVX-based RBP profiling studies (Figure 3F) suggests that enhancing crosslinking efficiency using FAX can be a key solution for the robust mapping of such RBPs.

Among the well-known unorthodox RBPs are the metabolic enzymes (2,38). Interestingly, we found that nearly all metabolic enzymes identified in UVX-RIC (2,39), including SHMT2, GAPDH, and ENO1, showed strongly enhanced capture efficiency in FAX-RIC (Figure 3G). A number of these unorthodox RBPs were found to bind RNA through their nicotinamide adenine dinucleotide binding domains (38,39), indicating that those metabolic enzymes could retain affinity to RNA adenine nucleotides. Our results thus suggested that these RBPs may have been significantly underrepresented in previous UVX-RIC studies owing to the strong nucleobase bias of UVX. We further validated the RNA-crosslinking dependent enrichment of the representative FAX-RIC exclusive RBPs, DDX19B, ENO1, and, GAPDH, through the WB analysis (Supplementary Figure S2B). Near exclusive detection of the protein signal for all three proteins in the FAX-RIC is consistent with our quantitative proteomics analysis and their absence in all of the seven representative UVX based RNA interactome profiles of the human cell lines (2). In summary, our system-wide quantitative analysis between FAX- and UVX-RIC revealed that FAX-RIC can enable robust RNA interactome mapping with broader and more balanced coverage of RBPs that are both conventional and unorthodox.

We extended our comparative analysis to a more comprehensive reference of UVX RNA interactome in HeLa cells (3) (referred to as REF-RIC), which was determined with greater input cell amount (∼10 times) and more in-depth profiling manner for LC-MS/MS analysis (Supplementary Figure S7A). Despite the significant differences in the experimental settings, the correlation between the protein LFQ intensities from the REF- and UVX-RIC experiments were high, suggesting that UVX-RIC largely reproduced the relative protein intensity value obtained by the REF-RIC (Supplementary Figure S7A, B). We also found that majority of the REF-RIC exclusive RBPs had higher adjusted *P*-value from the DEseq analysis, and the relatively low average LFQ values, compared to those of the com-

monly identified RBPs, suggesting that UVX-RIC reliably identified most of the high-confidence RBPs reported by the REF-RIC study (Supplementary Figure S7C-D). The size of REF and FAX RNA interactome were comparable, and yet there was a significant difference in the relative protein intensity of the individual RBPs (Supplementary Figure S7E, F). Enhanced profiling of the metabolic enzymes by FAX-RIC is also consistent with our previous analysis and result (Figure 3G, Supplementary Figure S2B). Furthermore, overrepresentation of the distinct classical RBDs in the set of REF- (RRM) or FAX-RIC preferred RBPs (DEAD, KH and dsrm) also closely recapitulate our previous result on the same analysis (Figure 3D, Supplementary Figure S7G). Taken together, our comparative analyses using previously reported UVX-RIC data further validated the similarities and differences between the UVX- and FAX-RIC approaches.

FAX-RIC enables comprehensive and unbiased RNA interactome profiling in multicellular organisms *in vivo* (*X. laevis* oocytes and embryos)

The most evident limitation of UVX is its highly limited penetration depth into opaque biological systems (12,13). Due to this innate limitation of UVX, to our knowledge, no previous study has shown whether one can achieve sensitive and comprehensive profiling of the RNA interactome in multicellular organisms. Owing to the good membrane permeability of formaldehyde (14), FAX-RIC readily overcomes this limitation of conventional UVX-RIC, enabling robust RNA interactome profiling in multicellular tissues and organisms. Thus, we performed a comparative analysis of UVX- and FAX-RIC in *X. laevis* oocytes (stage VI) and embryos (stage 8–9), which are large (1–2 mm in diameter), opaque, and partially pigmented. We found that at a newly optimized FAX condition for *X. laevis* samples (2% formaldehyde (w/v) for 10 min), FAX-RIC captured a significant amount of proteins whose profile was distinct from that of both the input and NoX control (Supplementary Figure S8). On the contrary, the protein staining pattern of UVX-RIC experiments done with relatively high UV-irradiation energy (2 J) was indistinguishable from that of the NoX control samples (Supplementary Figure S8).

FAX-RIC and UVX-RIC experiments were carried out in triplicate using 50 oocytes or embryos per replicate along with the NoX control. The LC-MS/MS results of FAX-RIC experiments were highly reproducible between the replicate experiments in both oocytes and embryos (Supplementary Figure S9). FAX-RIC then identified 693 and 541 RNA interactome of *X. laevis* oocytes and embryos, respectively, at <0.01 FDR (Figure 4A-B and Supplementary Table S4). We found that collectively ∼80% of these FAX RNA interactome with obvious human orthologues in oocytes or embryos were reported as RBPs with relevant records such as 'RNA-binding' Gene Ontology (GO) terms or known RBDs (40) (Figure 4C). These results clearly demonstrated that the aforementioned advantage and specificity of FAX-RIC in HeLa cells could be readily reproduced even in large and opaque samples, enabling the comprehensive profiling of *in vivo* RNA interactome landscapes.

In contrast, UVX-RIC identified no more than 94 and 85 proteins as RNA interactome at the same oocyte and embryo scales, respectively (Figure 4A-B), most of which were also covered by the FAX-RIC results (Figure 4D-E). We found that the RBPs localized to nucleus in *X. laevis* oocytes, which were defined by Wuhr *et al.* (41), were significantly underrepresented compared to those of embryos in UVX-RIC (Figure 4F, left). This observation illustrates the inability of UVX-RIC to form detectable amounts of crosslinks for the RNP complexes localized to the oocyte nucleus. During the OET, these complexes become more evenly distributed throughout the egg (42) and thus become more available to be captured and identified via UVX-RIC in embryo. As a result, comparative analysis of UVX-RIC experiments can result in inaccurate conclusions for some of such stage-specifically localized RBPs, e.g. overestimation of their embryo stage-specific RNA-binding activities. In contrast, FAX-RIC unbiasedly identified a larger number of stage-specifically localized RBPs from both oocytes and embryos (Figure 4F, right). These results illustrated how the use of FAX-RIC was critical to the comprehensive and unbiased profiling of the RNA interactome landscape in physiologically distinct multicellular organisms and tissues *in vivo*.

FAX-RIC reveals the landscape of mRNP remodeling in *X. laevis* oocyte-to-embryo transition

Quantitative comparison of the RNA interactome profiles from FAX-RIC can characterize RNA interactome landscape transformation in the *X. laevis* OET. There were 295 FAX-captured proteins with significant changes in their LFQ intensity level at <0.05 FDR and >1.5 $\log_2$ fold-change between oocytes and embryos (Figure 5A and Supplementary Table S5). The differences in the two RNA interactomes may reflect the changes in protein expression level and/or those in RNA-binding activity of these RBPs. Thus, we investigated how influential the protein expression changes were to the captured protein changes during the transition (Figure 5B and Supplementary Table S5). We found that a relatively small number of differentially captured RBPs represented stage-specific protein expression patterns (≥1.5 in $\log_2$ protein fold-change), while the majority of differentially captured RBPs were largely stable in their expression levels, strongly indicating the alterations in their RNA-binding activity (Figure 5B).

Among those RBPs with an oocyte-specific expression pattern, Cpeb1, Caprin2, Eif4enif1, Zar1 and Patl1 were highly enriched via FAX-RIC experiments and all of them, with the notable exception of Caprin2, are known to have essential regulatory functions in *X. laevis* OET (42–45) (Figure 5C). Interestingly, CAPRIN2 and TDRKH were reported as the only two proteins with significant downregulation during human oocyte maturation *in vitro* (46), and both proteins were consistently observed in our data (Figure 5C), suggesting that the regulatory mechanism and/or importance of these RBPs in OET may be conserved from *X. laevis* to humans. In contrast, we found relatively few RBPs with an embryo-specific expression pattern (Figure 5B), consistent with the previous report that the early embryonic proteome of *X. laevis* largely consists of maternally

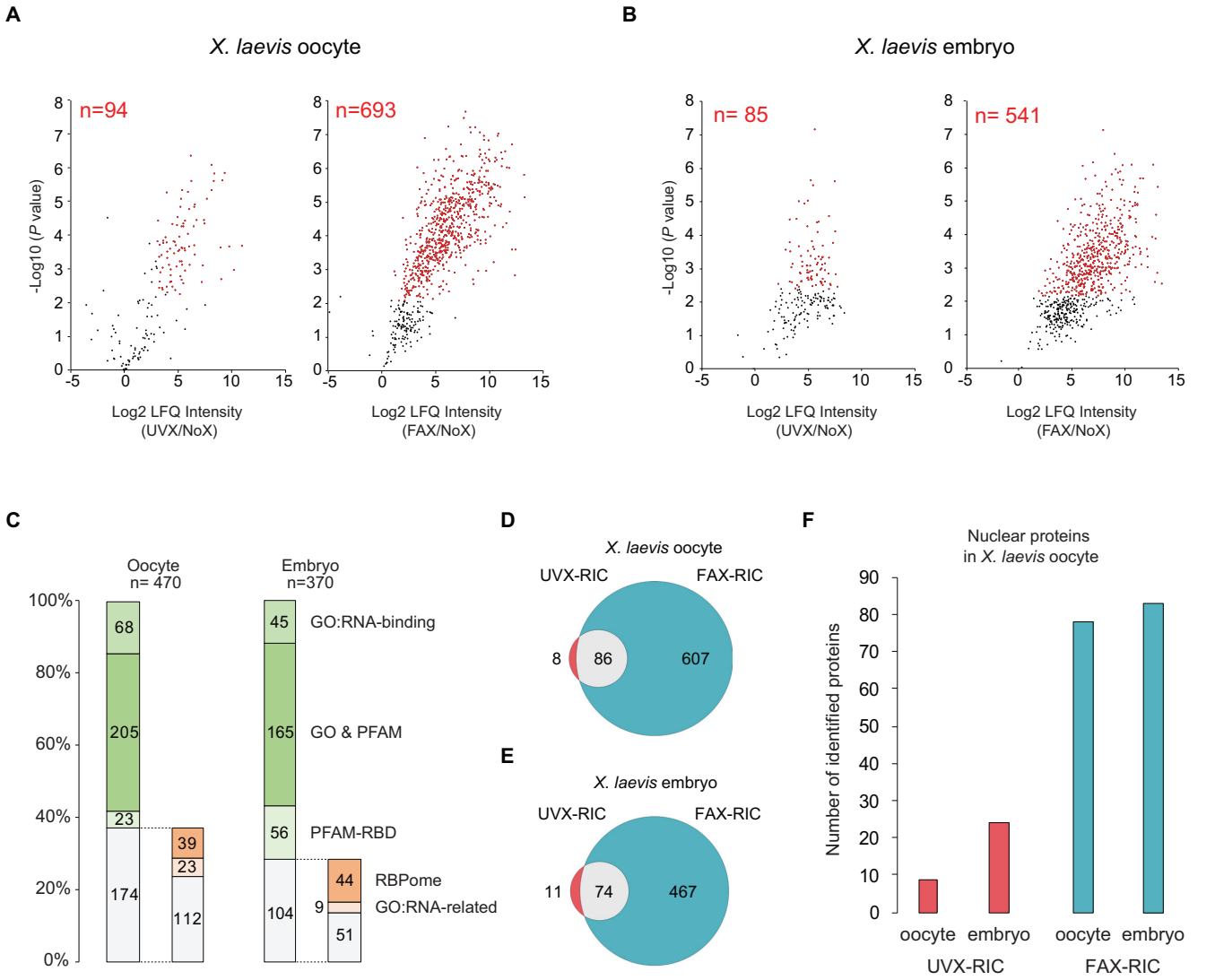


Figure 4. FAX-RIC enable both comprehensive and unbiased RNA interactome profiling in Xenopus laevis oocyte and embryo. (A, B) Defining the high confidence UVX and FAX RNA interactome in X. laevis oocyte (stage VI) (**A**) and embryo (stage 8–9) (**B**). Volcano plots displaying the $\log_2$ fold-change of average LFQ intensity (x-axis) and the $-\log_{10}$ Student's *t*-test *P* value (y-axis) for all the proteins quantified in at least two out of three replicate UVX- or FAX-RIC experiments. Proteins with $\log_2$ fold-change >1 and statistically significant enrichment over the NoX-RIC experiments (*P* value < 0.01, Student's *t*-test, adjusted by Benjamini–Hochberg method) are highlighted in red. (**C**) Proportion and number of human orthologous proteins in X. laevis RNA interactome that are annotated with GO:'RNA-binding', previously defined (42) 'RNA-related GO' and the 'RBD', or identified as RBPs in previous RBP profiling studies (2,5,6) RBPome. (D, E) Overlap between UVX-RIC and FAX-RIC RNA interactome in oocyte (**D**) and embryo (**E**). (**F**) Number of the *X. laevis* oocyte nucleus enriched proteins, defined by having >0.5 protein amount ratio in *X. laevis* oocyte nucleus compared to the cytoplasm (31), identified in two or more replicate UVX- and FAX-RIC experiments in oocyte or embryo.

deposited proteins (47). Nevertheless, among those RBPs, Ago4 and Khsrp proteins whose embryo-specific expressions and functions were reported previously (48,49) (Figure 5C).

As aforementioned, over 141 differentially captured RBPs showed no significant change in their protein expression level during OET (yellow dots in Figure 5B) and could be considered dynamic RBPs whose RNA-binding activities were significantly changed during *X. laevis* OET. Translational repression and subsequent activation of the specific target mRNAs (42,50) are one of the most distinctive events in mRNA biology of oocyte and embryo. Accordingly, GO term analysis of our dynamic RBPs identified the 'translational initiation' as the most significantly enriched GO term (Figure 5D). It has been known that translational activation of mRNAs during OET occurs through the first dissolution of translation-repressive oocyte-specific mRNP complexes containing Eif4e and the subsequent formation of the canonical Eif4e complex for translation initiation (42,50). Most intriguingly, however, we found that the FAX-captured levels of Eif4e were significantly downregulated in OET while those of other noncanonical translational pathway related RBPs such as Eif4g2, Denr, and Eif4e3 were upregulated (Figure 5E). It is noteworthy that in mice, Eif4g2 proteins are essential for the development of early embryos (51) and the differentiation of embryonic stem cells (52).

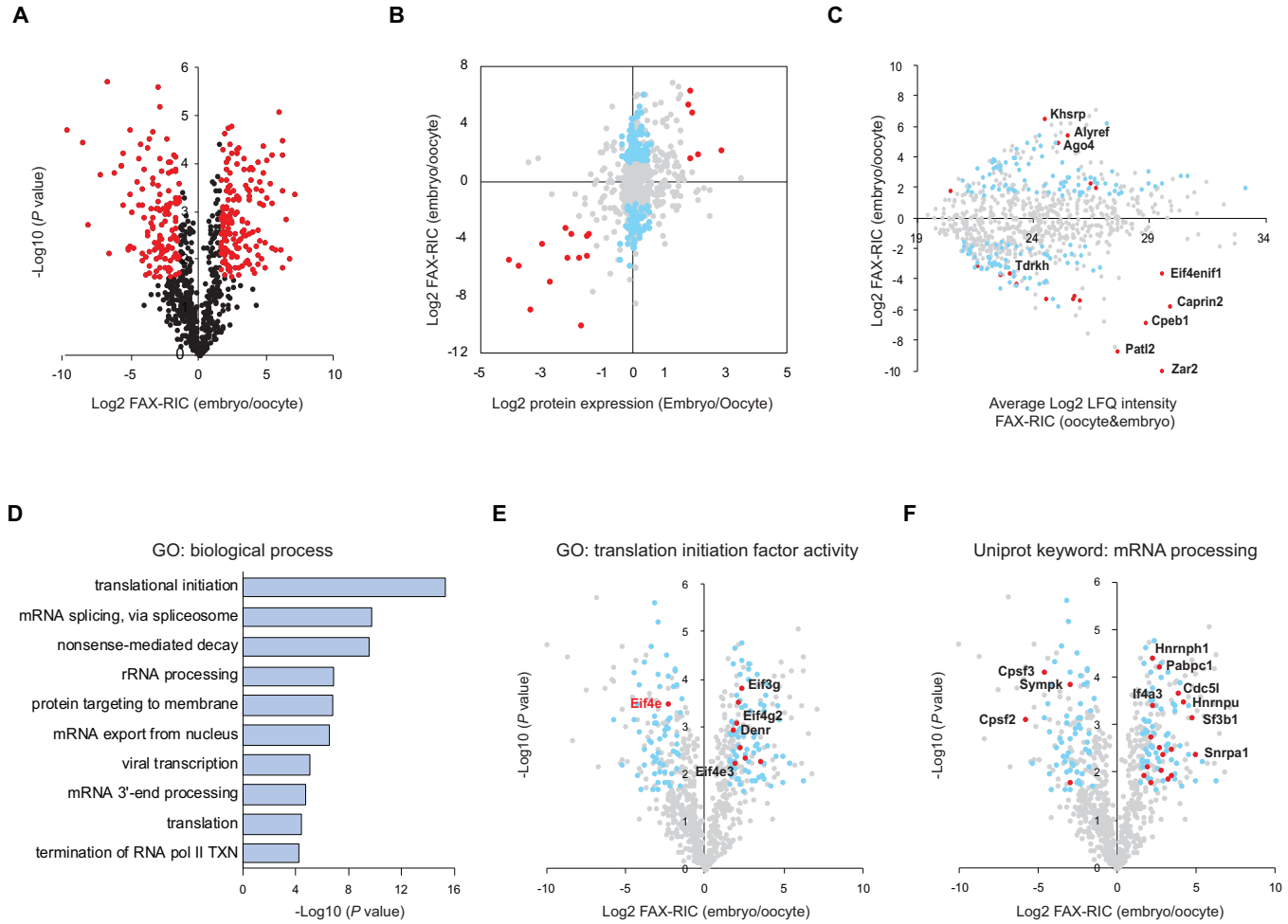


Figure 5. Transformation of mRNP complex landscape in X. laevis oocyte-to-embryo transition (OET). (**A**) Defining the differentially captured RBPs in oocyte or embryo stage FAX-RIC. Volcano plot displaying the $\log_2$ fold-change of average LFQ intensity and the $-\log_{10}$ Student's *t*-test *P* value (y-axis) for all the proteins identified as FAX RNA interactome in *X. laevis* are shown as black dots. Proteins with $\log_2$ fold-change >1.5 and had statistically significant enrichment in oocyte or embryo FAX-RIC experiments (*P* value < 0.05, Student's *t*-test, adjusted by Benjamini–Hochberg method) are highlighted in red. (**B**) FAX-RIC enrichment level change in X. laevis OET and the respective change in total protein expression level. Scatter plot displaying the sum of the $\log_2$ protein expression level changes in oocyte maturation and early embryo development, previously reported by our group (55) (x-axis) and the respective change in average FAX-RIC LFQ intensity level (y-axis). All the identified proteins were shown by grey dots. Proteins whose FAX-RIC captured protein amount change can be explained by their respective change in total protein abundance during OET are highlighted in red. The 'dynamic RBPs', whose FAX-RIC enriched protein amounts are significantly changed while their respective total protein abundance were changed with $\log_2$ fold-change <0.5 are highlighted in light blue. (**C**) Scatter plot displaying average LFQ intensity from the FAX-RIC experiments in oocyte and embryo FAX-RIC (x-axis) and the change in FAX-RIC enrichment level between oocyte and embryo FAX-RIC (y-axis). RBPs are marked with light blue or red, as described in (**B**). The protein names are inserted for the targets with most significant change and/or enrichment level in FAX-RIC, along with the Tdrkh whose human homologue had similar change during the maturation of human oocyte (46). (**D**) Most significantly enriched biological process GO terms in the 'dynamic RBPs', as defined in (B). (**E**) Volcano plot same as (A) but the 'dynamic RBPs' are highlighted with light blue and the dynamic RBPs annotated with 'GO: translation initiation' are highlighted with red. The protein Eif4e is highlighted for its unexpected change. (**F**) Same as (**E**) but RBPs annotated with the 'UniProt keyword: mRNA processing' are highlighted with red.

These results indicate that upon the dissociation of oocyte-specific Eif4e-containing mRNP complexes, noncanonical translation via Eif4e-independent mechanisms, either cap-independent (Eif4g2 (53,54) and Denr (55)) or alternatively cap-dependent (Eif4e3 (56)), are activated in part along with the canonical translation pathway during the OET.

Furthermore, we found that most of the RBPs related to 'RNA processing', most of which were splicing factors, had a significantly increased FAX-RIC levels in embryos (Figure 5F). This result is consistent with the fact that zygotic genome transcription activation in early embryos (57) likely requires the involvement of RNA processing factors to process the newly produced zygotic pre-mRNAs. This result is also in line with the conclusions drawn from a similar study profiling 'dynamic RBPs' in the drosophila maternal to zygotic transition (13). Intriguingly, notable exceptions occurred for Cpsf2, Cpsf3 and Sympk, all of which are cleavage and polyadenylation related factors (58) (Figure 5F). In oocytes, these RBPs are known to form oocyte-specific mRNP complexes with polyadenylated mRNAs in the cytoplasm (58). Unlike oocytes, these RBPs are known to interact with newly transcribed mRNAs prior to polyadenylation in the embryos (58) and most other somatic cells (59), and therefore should not be identified as part of

the RNA interactome via oligo-dT pulldown. Accordingly, these RBPs were notably absent from our RNA interactome in HeLa cells and all of the reviewed human RNA interactome studies (2), despite the ubiquitous expression of those RBPs in human cancer cell lines (60). Our results thus suggest that a significant majority of these RBPs become part of the embryo stage-specific RNP complexes at or before the embryo stage. Of note, such dynamic RBPs could not be revealed by gene expression analysis alone but by the integrated analysis of the RNA interactome and global proteomics. Collectively, all of these results demonstrate that the FAX-RIC method can enable robust mapping of physiologically distinct changes in RNP complex formations in multicellular organisms *in vivo.*

FAX-RIC based RNA interactome profiling in mouse liver

We further demonstrate that FAX can enable RNA interactome profiling in the mammalian tissue samples. For this, we utilized *M. musculus* liver samples. Analyzing the profile of the nucleic acids obtained from the liver samples, treated with 2% or 4% formaldehyde (10 min), we found distinct and near complete change in migration pattern of genomic DNAs and RNAs at 4% FAX condition (Supplementary Figure S10A). Furthermore, consistent with the known characteristics of protein crosslinked RNA (5), most of the FAX liver RNAs were found at the interphase of acid guanidinium thiocyanate-phenol chloroform extraction, suggesting that extensive RNA–protein crosslinking had occurred. On the contrary, the profiles of the extracted liver RNAs in both negative control and UVX (2 J) were largely indistinguishable from each other (Supplementary Figure S10B). Taken together, we found that the FAX (at 4% 10 min) can both penetrate and crosslink mouse liver RNA interactome with high efficiency.

To reduce the potential interference of protein-to-protein crosslinking products at such strong FAX condition, we applied our peptide-level FAX-RIC strategy, in which the peptides that are not directly crosslinked to RNAs are removed via trypsin digestion and washing. This strategy ensures the specificity of the RIC to direct RNA–protein interactions. Of note, we adapted the GITC based lysis buffer to prevent RNA degradation (61), instead of the conventional LDS lysis buffer, which can be critical to animal tissue samples, especially the nuclease-rich mammalian organs such as the liver (61). After tissue lysis, GITC removal, and trypsin digestion, we could effectively retrieve the RNA-peptide conjugates using oilgo-dT beads or silica columns (62), for the profiling of poly(A)+ RNA interactome or total RNA interactome, respectively (Supplementary Figure S11A).

Poly(A)+ RNA interactome from mouse liver was relatively small but largely consisted of the RBPs annotated with known RBDs (Supplementary Figure S11B and S11D, Table S6), suggesting that the peptide-level FAX-RIC protocol can profile the tissue RNA interactome with high specificity. Among the notable RBPs observed were A1CF, NONO and CELF1, all of which have known function in post transcriptional regulation of mRNAs in mouse liver (63–65) (Supplementary Table S6). On the other hand, the total RNA interactome capture identified 761 significantly enriched proteins (Supplementary Figure S11C, Table S6), the majority of which are the known RBPs (Supplementary Figure S11D). Significant increase in the relative protein intensity of the representative mRNA and rRNA binding proteins from the poly(A) and total RIC experiments, respectively, further suggested the methods' specificity to the target RNA interactome (Supplementary Figure S11E, Table S6). In sum, we introduced a new versatile strategy to enable RNA interactome profiling of the mammalian tissue samples *in vivo* and showed the feasibility of the strategy to mouse liver.

DISCUSSION

UVX has long been regarded as the gold standard of *in vivo* crosslinking method and highly specific for RNA–protein interaction, in particular due to its 'zero-length' characteristic to the interaction. However, it has been reported that UVX can irreversibly crosslink protein-protein interactions as well (66,67), indicating that a certain degree of false positive RBPs will be profiled via protein-protein crosslinking even from UVX-RIC experiment as a similar concern exists with FAX. Moreover, UVX can induce RNA damaging or fragmentation (7,68), which can reduce the RNA recovery of oligo-dT pulldown and thereby decrease the overall quantity of captured proteins. These drawbacks of UVX have been largely neglected in the field of RNA biology and it would be appropriate to re-evaluate the value of UVX and FAX for the researches of RNA–protein interactions.

In this study, we report that FAX can capture RNA–protein interaction with high specificity and efficiency not only in cultured cells but also in multicellular organisms. Through the first system-wide and quantitative comparison of two *in vivo* crosslinking methods (FAX versus UVX), we solidly demonstrated that FAX-RIC was more efficient and as highly specific as UVX-RIC for the mapping of the *in vivo* RNA–protein interactions, while possessing advantages over UVX, particularly in opaque samples such as *X. laevis* oocytes and embryos. Furthermore, for the first time we also performed the systematic *de novo* analyses of RNA interactome during OET using the FAX-RIC in addition to global proteome profiling data. The majority of the differentially enriched RBPs had no significant change in protein expression level, underscoring the importance of the RIC method in discovering functionally regulated RNA–protein interactions *in vivo*. The significant changes in our RBP profiles were clearly reflective of the known changes in RNP complex functions during early animal development (13,69) and disclosed some of the under-evaluated components of RNP complexes, such as those associated with noncanonical translational pathway.

Despite significant expansion of our knowledge in the RBP repertoire, many aspects of context-dependent RNA–protein interactions, e.g., such 'dynamic RBPs' in OET, demand to be further elucidated especially in human physiology and diseases. FAX-RIC can allow for the profiling of the context-dependent 'dynamic RBPs' in various human tissues in combination with global proteome profiling or/and post-translational modification (PTM) proteomics data. Such integrated approach would serve as a powerful platform for discovering novel key regulatory RBPs and

PTMs of such RBPs (i.e. RBP-code). We thus expect that FAX-RIC would significantly broaden our understanding of the dynamic mRNP formation in multicellular organisms and human tissues *in vivo* as we demonstrate with mouse liver tissues.

DATA AVAILABILITY

All the original LC-MS/MS datasets and related identification files used to support this paper have been deposited to the ProteomeXchange Consortium (http://proteomecentral.proteomexchange.org) via the PRIDE partner repository with the dataset identifier <PXD010046>, and for mouse liver RNA interactome <PXD022023>.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

FUNDING

Institute for Basic Science from the Ministry of Science and ICT of Korea [IBS-R008-D1 to Y.N., H.K., Y.C., S.S., S.C.K., V.N.K., J.-S.K.]; BK 21 Research Fellowships from the Ministry of Education of Korea [to Y.N., Y.C., S.S.]. Funding for open access charge: College of Natural Sciences, Seoul National University.
Conflict of interest statement. None declared.

REFERENCES

1. Dreyfuss,G., Kim,V.N. and Kataoka,N. (2002) Messenger-RNA-binding proteins and the messages they carry. *Nat. Rev. Mol. Cell Biol.*, **3**, 195–205.
2. Hentze,M.W., Castello,A., Schwarzl,T. and Preiss,T. (2018) A brave new world of RNA-binding proteins. *Nat. Rev. Mol. Cell Biol.*, **19**, 327.
3. Castello,A., Fischer,B., Eichelbaum,K., Horos,R., Beckmann,B.M., Strein,C., Davey,N.E., Humphreys,D.T., Preiss,T., Steinmetz,L.M. *et al.* (2012) Insights into RNA biology from an atlas of mammalian mRNA-binding proteins. *Cell*, **149**, 1393–1406.
4. Baltz,A.G., Munschauer,M., Schwanhausser,B., Vasile,A., Murakawa,Y., Schueler,M., Youngs,N., Penfold-Brown,D., Drew,K., Milek,M. *et al.* (2012) The mRNA-bound proteome and its global occupancy profile on protein-coding transcripts. *Mol. Cell*, **46**, 674–690.
5. Trendel,J., Schwarzl,T., Horos,R., Prakash,A., Bateman,A., Hentze,M.W. and Krijgsveld,J. (2019) The human RNA-binding proteome and its dynamics during translational arrest. *Cell*, **176**, 391–403.
6. Queiroz,R.M.L., Smith,T., Villanueva,E., Marti-Solano,M., Monti,M., Pizzinga,M., Mirea,D.-M., Ramakrishna,M., Harvey,R.F., Dezi,V. *et al.* (2019) Comprehensive identification of RNA–protein interactions in any organism using orthogonal organic phase separation (OOPS). *Nat. Biotechnol.*, **37**, 169–178.
7. Beckmann,B.M., Horos,R., Fischer,B., Castello,A., Eichelbaum,K., Alleaume,A.M., Schwarzl,T., Curk,T., Foehr,S., Huber,W. *et al.* (2015) The RNA-binding proteomes from yeast to man harbour conserved enigmRBPs. *Nat. Commun.*, **6**, 10127.
8. Perez-Perri,J.I., Rogell,B., Schwarzl,T., Stein,F., Zhou,Y., Rettel,M., Brosig,A. and Hentze,M.W. (2018) Discovery of RNA-binding proteins and characterization of their dynamic responses by enhanced RNA interactome capture. *Nat. Commun.*, **9**, 4408.
9. McHugh,C.A., Chen,C.K., Chow,A., Surka,C.F., Tran,C., McDonel,P., Pandya-Jones,A., Blanco,M., Burghard,C., Moradian,A. *et al.* (2015) The Xist lncRNA interacts directly with SHARP to silence transcription through HDAC3. *Nature*, **521**, 232–236.
10. Chu,C., Zhang,Q.C., da Rocha,S.T., Flynn,R.A., Bharadwaj,M., Calabrese,J.M., Magnuson,T., Heard,E. and Chang,H.Y. (2015) Systematic discovery of Xist RNA binding proteins. *Cell*, **161**, 404–416.
11. Bao,X., Guo,X., Yin,M., Tariq,M., Lai,Y., Kanwal,S., Zhou,J., Li,N., Lv,Y., Pulido-Quetglas,C. *et al.* (2018) Capturing the interactome of newly transcribed RNA. *Nat. Methods*, **15**, 213–220.
12. Elinson,R.P. and Pasceri,P. (1989) Two UV-sensitive targets in dorsoanterior specification of frog embryos. *Development*, **106**, 511–518.
13. Sysoev,V.O., Fischer,B., Frese,C.K., Gupta,I., Krijgsveld,J., Hentze,M.W., Castello,A. and Ephrussi,A. (2016) Global changes of the RNA-bound proteome during the maternal-to-zygotic transition in Drosophila. *Nat. Commun.*, **7**, 12128.
14. Thavarajah,R., Mudimbaimannar,V.K., Elizabeth,J., Rao,U.K. and Ranganathan,K. (2012) Chemical and physical basics of routine formaldehyde fixation. *J. Oral Maxillofac. Pathol.*, **16**, 400–405.
15. Kim,B., Jeong,K. and Kim,V.N. (2017) Genome-wide mapping of DROSHA cleavage sites on primary microRNAs and noncanonical substrates. *Mol. Cell*,**66**, 258–269.
16. Vasudevan,S. and Steitz,J.A. (2007) AU-rich-element-mediated upregulation of translation by FXR1 and Argonaute 2. *Cell*, **128**, 1105–1118.
17. Yong,J., Kasim,M., Bachorik,J.L., Wan,L. and Dreyfuss,G. (2010) Gemin5 delivers snRNA precursors to the SMN complex for snRNP biogenesis. *Mol. Cell*, **38**, 551–562.
18. Knoener,R.A., Becker,J.T., Scalf,M., Sherer,N.M. and Smith,L.M. (2017) Elucidating the in vivo interactome of HIV-1 RNA by hybridization capture and mass spectrometry. *Sci. Rep.*, **7**, 16965.
19. Panhale,A., Richter,F.M., Ramírez,F., Shvedunova,M., Manke,T., Mittler,G. and Akhtar,A. (2019) CAPRI enables comparison of evolutionarily conserved RNA interacting regions. *Nat. Commun.*, **10**, 2682.
20. Wheeler,E.C., Van Nostrand,E.L. and Yeo,G.W. (2018) Advances and challenges in the detection of transcriptome-wide protein-RNA interactions. *Wiley Interdiscip. Rev. RNA*, **9**, e1436.
21. Hoffman,E.A., Frey,B.L., Smith,L.M. and Auble,D.T. (2015) Formaldehyde crosslinking: a tool for the study of chromatin complexes. *J. Biol. Chem.*, **290**, 26404–26411.
22. Gavrilov,A., Razin,S.V. and Cavalli,G. (2015) In vivo formaldehyde cross-linking: it is time for black box analysis. *Brief. Funct. Genomics*, **14**, 163–165.
23. Castello,A., Fischer,B., Frese,C.K., Horos,R., Alleaume,A.M., Foehr,S., Curk,T., Krijgsveld,J. and Hentze,M.W. (2016) Comprehensive identification of RNA-binding domains in human cells. *Mol. Cell*, **63**, 696–710.
24. Singh,G., Kucukural,A., Cenik,C., Leszyk,J.D., Shaffer,S.A., Weng,Z. and Moore,M.J. (2012) The cellular EJC interactome reveals higher-order mRNP structure and an EJC-SR protein nexus. *Cell*, **151**, 750–764.
25. Kim,Y., Lee,J.H., Park,J.-E., Cho,J., Yi,H. and Kim,V.N. (2014) PKR is activated by cellular dsRNAs during mitosis and acts as a mitotic regulator. *Genes Dev.*, **28**, 1310–1322.
26. Sive,H.L., Grainger,R.M. and Harland,R.M. (2010) Isolation of Xenopus oocytes. *Cold Spring Harb. Protoc.*, **2010**, pdb.prot5534.
27. Sive,H.L., Grainger,R.M. and Harland,R.M. (2007) Xenopus laevis in vitro fertilization and natural mating methods. *Cold Spring Harb. Protoc.*, **2007**, pdb.prot4737.
28. Peuchen,E.H., Cox,O.F., Sun,L., Hebert,A.S., Coon,J.J., Champion,M.M., Dovichi,N.J. and Huber,P.W. (2017) Phosphorylation dynamics dominate the regulated proteome during early xenopus development. *Sci. Rep.*, **7**, 15647.
29. Jung,J., Jeong,K., Choi,Y., Kim,S.A., Kim,H., Lee,J.W., Kim,V.N., Kim,K.P. and Kim,J.S. (2019) Deuterium-free, three-plexed peptide diethylation for highly accurate quantitative proteomics. *J. Proteome Res.*, **18**, 1078–1087.
30. Tyanova,S., Temu,T. and Cox,J. (2016) The MaxQuant computational platform for mass spectrometry-based shotgun proteomics. *Nat. Protoc.*, **11**, 2301.
31. Tyanova,S., Temu,T., Sinitcyn,P., Carlson,A., Hein,M.Y., Geiger,T., Mann,M. and Cox,J. (2016) The Perseus computational platform for comprehensive analysis of (prote)omics data. *Nat. Methods*, **13**, 731.

32. Huang da,W., Sherman,B.T. and Lempicki,R.A. (2009) Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. *Nucleic Acids Res.*, **37**, 1–13.
33. Huang da,W., Sherman,B.T. and Lempicki,R.A. (2009) Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat. Protoc.*, **4**, 44–57.
34. Lim,J., Ha,M., Chang,H., Kwon,S.C., Simanshu,D.K., Patel,D.J. and Kim,V.N. (2014) Uridylation by TUT4 and TUT7 Marks mRNA for Degradation. *Cell*, **159**, 1365–1376.
35. Hockensmith,J.W., Kubasek,W.L., Vorachek,W.R. and von Hippel,P.H. (1986) Laser cross-linking of nucleic acids to proteins. Methodology and first applications to the phage T4 DNA replication system. *J. Biol. Chem.*, **261**, 3512–3518.
36. Lunde,B.M., Moore,C. and Varani,G. (2007) RNA-binding proteins: modular design for efficient function. *Nat. Rev. Mol. Cell Biol.*, **8**, 479–490.
37. Maris,C., Dominguez,C. and Allain,F.H. (2005) The RNA recognition motif, a plastic RNA-binding platform to regulate post-transcriptional gene expression. *FEBS J.*, **272**, 2118–2131.
38. Castello,A., Hentze,M.W. and Preiss,T. (2015) Metabolic enzymes enjoying new partnerships as RNA-Binding proteins. *Trends Endocrinol. Metab.*, **26**, 746–757.
39. Liao,Y., Castello,A., Fischer,B., Leicht,S., Foehr,S., Frese,C.K., Ragan,C., Kurscheid,S., Pagler,E., Yang,H. *et al.* (2016) The cardiomyocyte RNA-binding proteome: links to intermediary metabolism and heart disease. *Cell Rep.*, **16**, 1456–1469.
40. Gerstberger,S., Hafner,M. and Tuschl,T. (2014) A census of human RNA-binding proteins. *Nat. Rev. Genet.*, **15**, 829–845.
41. Wuhr,M., Guttler,T., Peshkin,L., McAlister,G.C., Sonnett,M., Ishihara,K., Groen,A.C., Presler,M., Erickson,B.K., Mitchison,T.J. *et al.* (2015) The nuclear proteome of a vertebrate. *Curr. Biol.*, **25**, 2663–2671.
42. Radford,H.E., Meijer,H.A. and de Moor,C.H. (2008) Translational control by cytoplasmic polyadenylation in Xenopus oocytes. *Biochim. Biophys. Acta*, **1779**, 217–229.
43. Nakamura,Y., Tanaka,K.J., Miyauchi,M., Huang,L., Tsujimoto,M. and Matsumoto,K. (2010) Translational repression by the oocyte-specific protein P100 in Xenopus. *Dev. Biol.*, **344**, 272–283.
44. Standart,N. and Minshall,N. (2008) Translational control in early development: CPEB, P-bodies and germinal granules. *Biochem. Soc. Trans.*, **36**, 671–676.
45. Wu,X., Wang,P., Brown,C.A., Zilinski,C.A. and Matzuk,M.M. (2003) Zygote arrest 1 (Zar1) is an evolutionarily conserved gene expressed in vertebrate ovaries. *Biol. Reprod.*, **69**, 861–867.
46. Virant-Klun,I., Leicht,S., Hughes,C. and Krijgsveld,J. (2016) Identification of maturation-specific proteins by single-cell proteomics of human oocytes. *Mol. Cell. Proteomics*, **15**, 2616–2627.
47. Peshkin,L., Wuhr,M., Pearl,E., Haas,W., Freeman,R.M. Jr., Gerhart,J.C., Klein,A.M., Horb,M., Gygi,S.P. and Kirschner,M.W. (2015) On the relationship of protein and mRNA dynamics in vertebrate embryonic development. *Dev. Cell*, **35**, 383–394.
48. Lund,E., Sheets,M.D., Imboden,S.B. and Dahlberg,J.E. (2011) Limiting Ago protein restricts RNAi and microRNA biogenesis during early development in Xenopus laevis. *Genes Dev.*, **25**, 1121–1131.
49. Kroll,T.T., Zhao,W.M., Jiang,C. and Huber,P.W. (2002) A homolog of FBP2/KSRP binds to localized mRNAs in Xenopus oocytes. *Development*, **129**, 5609–5619.
50. Richter,J.D. (2007) CPEB: a life in translation. *Trends Biochem. Sci.*, **32**, 279–285.
51. Yamanaka,S., Zhang,X.Y., Maeda,M., Miura,K., Wang,S., Farese,R.V. Jr., Iwao,H. and Innerarity,T.L. (2000) Essential role of NAT1/p97/DAP5 in embryonic differentiation and the retinoic acid pathway. *EMBO J.*, **19**, 5533–5541.
52. Sugiyama,H., Takahashi,K., Yamamoto,T., Iwasaki,M., Narita,M., Nakamura,M., Rand,T.A., Nakagawa,M., Watanabe,A. and Yamanaka,S. (2017) Nat1promotes translation of specific proteins that induce differentiation of mouse embryonic stem cells. *Proc. Natl. Acad. Sci. U.S.A.*, **114**, 340–345.
53. Shatsky,I.N., Terenin,I.M., Smirnova,V.V. and Andreev,D.E. (2018) Cap-independent translation: what's in a name? *Trends Biochem. Sci.*, **43**, 882–895.
54. Takahashi,K., Maruyama,M., Tokuzawa,Y., Murakami,M., Oda,Y., Yoshikane,N., Makabe,K.W., Ichisaka,T. and Yamanaka,S. (2005) Evolutionarily conserved non-AUG translation initiation in NAT1/p97/DAP5 (EIF4G2). *Genomics*, **85**, 360–371.
55. Skabkin,M.A., Skabkina,O.V., Dhote,V., Komar,A.A., Hellen,C.U.T. and Pestova,T.V. (2010) Activities of Ligatin and MCT-1/DENR in eukaryotic translation initiation and ribosomal recycling. *Genes Dev.*, **24**, 1787–1801.
56. Volpon,L., Osborne,M.J., Culjkovic-Kraljacic,B. and Borden,K.L.B. (2013) eIF4E3, a new actor in mRNA metabolism and tumor suppression. *Cell Cycle*, **12**, 1159–1160.
57. Tadros,W. and Lipshitz,H.D. (2009) The maternal-to-zygotic transition: a play in two acts. *Development*, **136**, 3033–3042.
58. Charlesworth,A., Meijer,H.A. and de Moor,C.H. (2013) Specificity factors in cytoplasmic polyadenylation. *Wiley Interdiscip. Rev. RNA*, **4**, 437–461.
59. Gruber,A.R., Martin,G., Keller,W. and Zavolan,M. (2014) Means to an end: mechanisms of alternative polyadenylation of messenger RNA precursors. *Wiley Interdiscip. Rev. RNA*, **5**, 183–196.
60. Thul,P.J. and Lindskog,C. (2018) The human protein atlas: a spatial map of the human proteome. *Protein Sci.*, **27**, 233–244.
61. Chomczynski,P., Wilfinger,W. and Mackey,K. (2017) Single-step method of total RNA isolation by guanidine–phenol extraction. In: *eLS*, John Wiley & Sons, Ltd (Ed.).
62. Asencio,C., Chatterjee,A. and Hentze,M.W. (2018) Silica-based solid-phase extraction of cross-linked nucleic acid–bound proteins. *Life Sci. Alliance*, **1**, e201800088.
63. Nikolaou,K.C., Vatandaslar,H., Meyer,C., Schmid,M.W., Tuschl,T. and Stoffel,M. (2019) The RNA-binding protein A1CF regulates hepatic fructose and glycerol metabolism via alternative RNA splicing. *Cell Rep.*, **29**, 283–300.
64. Benegiamo,G., Mure,L.S., Erikson,G., Le,H.D., Moriggi,E., Brown,S.A. and Panda,S. (2018) The RNA-binding protein NONO coordinates hepatic adaptation to feeding. *Cell Metab.*, **27**, 404–418.
65. Lewis,K., Valanejad,L., Cast,A., Wright,M., Wei,C., Iakova,P., Stock,L., Karns,R., Timchenko,L. and Timchenko,N. (2017) RNA binding protein CUGBP1 inhibits liver cancer in a phosphorylation-dependent manner. *Mol. Cell. Biol.*, **37**, e00128-17.
66. Itri,F., Monti,D.M., Della Ventura,B., Vinciguerra,R., Chino,M., Gesuele,F., Lombardi,A., Velotta,R., Altucci,C., Birolo,L. *et al.* (2016) Femtosecond UV-laser pulses to unveil protein-protein interactions in living cells. *Cell. Mol. Life Sci.*, **73**, 637–648.
67. Leo,G., Altucci,C., Bourgoin-Voillard,S., Gravagnuolo,A.M., Esposito,R., Marino,G., Costello,C.E., Velotta,R. and Birolo,L. (2013) Ultraviolet laser-induced cross-linking in peptides. *Rapid Commun. Mass Spectrom.*, **27**, 1660–1668.
68. Kladwang,W., Hum,J. and Das,R. (2012) Ultraviolet shadowing of RNA can cause significant chemical damage in seconds. *Sci. Rep.*, **2**, 517.
69. Richter,J.D. and Lasko,P. (2011) Translational control in oocyte development. *Cold Spring Harb. Perspect. Biol.*, **3**, a002758.