Large-scale identification of RBP-RNA interactions by RAPseq refines essentials of post-transcriptional gene regulation

Ionut Atanasoai¹, Sofia Papavasileiou¹, Natalie Preiß¹, Claudia Kutter¹

¹Department of Microbiology, Tumor and Cell Biology, Karolinska Institute, Science for Life Laboratory, Sweden

HIGHLIGHTS

- **RAPseq reveals *in vitro*-derived RBP-RNA interactomes**
- **the vertebrate-conserved HUR binding motif adapted to species-unique RNA features**
- **co-RAPseq and T7-RAPseq uncover binding cooperativity and modification dependencies**
- **non-canonical RBPs have specialized RNA interactomes**

SUMMARY

Over the past decade, thousands of putative human RNA binding proteins (RBPs) have been identified and increased the demand for specifying RNA binding capacities. Here, we developed RNA affinity purification followed by sequencing (RAPseq) that enables *in vitro* large-scale profiling of RBP binding to native RNAs. First, by employing RAPseq, we found that vertebrate HURs recognize a conserved RNA binding motif and bind predominantly to introns in zebrafish compared to 3'UTRs in human RNAs. Second, our dual RBP assays (co-RAPseq) uncovered cooperative RNA binding of HUR and PTBP1 within an optimal distance of 27 nucleotides. Third, we developed T7-RAPseq to discern m⁶A-dependent and -independent RNA binding sites of YTHDF1. Fourth, RAPseq of 26 novel non-canonical RBPs revealed specialized moonlighting interactions. Last, five pathological IGF2BP family variants exhibited different RNA binding patterns. Overall, our simple, scalable and versatile method enables to fast-forward RBP-related questions.

Graphical Abstract

Keywords

RNA binding proteins, RBP cooperativity, non-canonical RBPs, RBP mutants, RBP isoforms, native transcriptome, RNA modifications, RNA sequencing, RNA processing, vertebrate evolution

INTRODUCTION

RNA binding proteins (RBPs) decide the fate and function of all RNA species from synthesis to decay. Latest advances of crosslinking-based methods coupled with mass-spectrometry enabled the detection of RBP-RNA complexes and have extended the number of putative human RBPs to about 5,000 (Baltz et al., 2012; Castello et al., 2012, 2016; Queiroz et al., 2019; Trendel et al., 2019; Urdaneta et al., 2019). These can be grouped into canonical (c) or non-canonical (nc) RBPs depending on the presence or absence of a known RNA binding domain, respectively. Furthermore, technological developments allowing transcriptome-wide mapping of modified cellular RNA species uncovered extensive changes that impact binding of RBPs (Alarcón et al., 2015; Dominissini et al., 2012; Dominissini et al., 2016; Huang et al., 2018; Meyer, 2019; Wu et al., 2018; L.-S. Zhang et al., 2019).

To uncover mechanisms of RBP-mediated regulatory processes, it is crucial to identify RBP-bound RNA molecules, their exact RNA binding sites and interaction strengths. UV-crosslinking followed by immunoprecipitation (CLIP) (Ule et al., 2003) has revolutionized detecting RBP binding sites, and by coupling it to high-throughput RNA sequencing (RNAseq) has yielded many variations of the method, including HITS-CLIP (Licatalosi et al., 2008), PAR-CLIP (Hafner et al., 2010), iCLIP (König et al., 2010), irCLIP (Zarnegar et al., 2016) and eCLIP (Van Nostrand et al., 2016), among others. Although CLIPseq methods have enabled breakthroughs in the field, limitations remain due to the low UV-crosslinking efficiency favoring distinct amino acid-ribonucleotide interactions, antibody availability and extensive hands-on time to produce CLIPseq libraries resulting in quantitative errors and high experimental failure rates (Lin & Miles, 2019; Wheeler et al., 2018). In addition, large-scale *in vitro* technologies, such as RNA Bind-n-Seq (RBNS) (Lambert et al., 2014), RNAcompete (Ray et al., 2013) and HTR-SELEX (Jolma et al., 2020), were developed to profile binding of purified recombinant RBPs to random synthetic RNAs of 20 or 40 nucleotides (nt) and allow quantitative assessment of primary and secondary binding motifs. However, these methods cannot infer the impact of intrinsic RNA features present in native cellular transcripts on RNA binding affinities.

To enable quantitative assessments of RBP binding sites in cellular transcriptomes, we developed RAPseq, an RBP-centric affinity purification-based method. RAPseq allows to study *in vitro* RBP-RNA interactomes. We applied RAPseq to (i) describe the evolution of HUR as an essential RBP across multiple vertebrates by cross-species RAPseq assays, (ii) measure transcriptome-wide cooperativity of HUR and PTBP1 through co-RAPseq assays, (iii) clarify the role of YTHDF1 as an m⁶A-reader by comparing cellular (RAPseq) and *in vitro* generated (T7-RAPseq) transcriptomes, (iv) profile transcriptomes bound by 26 ncRBPs and (v) assess the impact of disease-associated RBP variants in RNA binding activities on a transcriptome-wide scale. Altogether, RAPseq proves to be a simple, scalable and multiplexable method that can be easily implemented to deliver robust and highly reproducible RBP-RNA interactomes across various applications.

RESULTS

RAPseq enables interrogating RBP-RNA interactions based on RNA structure-, sequence- and modification-specific binding modalities

RAPseq is a binding assay that measures interactions between a recombinant protein and native RNA using next generation sequencing (NGS) as a readout (**Figure 1A**). For generating the RNA substrate, we chemically fragmented RNA isolated from cells and tissues to circa 35 nt products (**Figure S1A**). The subsequent RNA 3'-end dephosphorylation and 5'-end phosphorylation enabled ligation of NGS compatible adapters (**Figure 1A**). For producing the recombinant RBP, we cloned the respective cDNA fused to HaloTag (Los et al., 2008) into a cell-free expression vector (**Figure 1A**). Tagging Halo at the RBP C-terminus ensured assaying

- purification and RNA binding assay. Next-generation sequencing serves as read-out using unique molecular identifiers (UMI) (right).
- (B)** Dotplot displays binding of HUR to non-targeting (NT) and targeting (T1 to T3) RNA regions (x-axis) of fragmented (left) and full-length (right) RNA from HepG2 cells. RAP-qPCR (left) was performed in three replicates (colored dots), means (black horizontal lines) are shown. RNA binding is displayed as fold change ($\Delta\Delta Ct$) of HUR over HaloTag control (y-axis). RAPseq coverage tracks (right) represent HUR binding to fragmented RNA (in reads per million, RPM). Genomic locations of the corresponding target regions (T1-3) are specified. Tracks for two HUR replicates (green) and two RAPseq controls (HaloTag and RNA input, grey) are shown.
- (C-E)** Genome tracks demonstrate binding of **(C)** HUR to ARE and GRE motifs, **(D)** IRP1 to the IRE and **(E)** YTHDF1 to a modified nucleotide (m^6A) within KDELR2, FTL and HNRNPA0 mRNAs, respectively, in HepG2 cells. Genomic locations with scale bar indicate the length of the genomic region in bases and gene features (black rectangle: exon and UTR, grey line: intron, arrow: direction of transcription) (x-axis) and normalized read density (RPM, y-axis) are shown. Vertical yellow lines highlight bound RNA elements. Tracks for two RBP replicates (green) and m^6A -specific RNA immunoprecipitation (IP, yellow) over input control are presented.

RAPseq confirms known RBP binding specificities to RNA

To benchmark RAPseq, we focused on five cRBPs (HNRNPA1, HNRNPC, PTBP1, RBFOX2 and YBX3) that recognize RNA through sequence-specific interactions. We profiled their RBP-RNA interactomes using HepG2 RNA and compared our results to publicly available *in cellulo* eCLIPseq data (Van Nostrand et al., 2020) and *in vitro* RBP-RNA interaction assays (RNA compete, RBNS and HTR-SELEX).

We observed high similarity in the bound motifs for the five RBPs across the diverse interaction assays (**Figure 2A** and **Figure S2A**). To assess the quantitative nature of RAPseq, we calculated a peak binding score, which is defined as the average enrichment of reads over our two negative controls weighted by the average replicate normalized read counts (STAR Methods). When determining the peak binding score as a function of uracil content at each individual HNRNPC binding site, our RAPseq data showed that binding of HNRNPC to its RNA targets reached a plateau at an uracil content of $\geq 50\%$ per target. This trend was also visible when considering only fold changes demonstrating that RAPseq can produce typical biochemical binding curves. In contrast, this assessment was not detectable in eCLIP data (**Figure 2B**). For RBFOX2, we noticed a slight deviation from the GCAUG motif identified in the other assays. We found that the second position in the motif showed an equal abundance of C and A. This GAAUG motif has been recently reported (Begg et al., 2020). A more generalizable RBFOX2 motif, GNWYG (N corresponds to any, W to A/U, Y to C/U nucleotide), identified using lenient criteria (STAR Methods) was present in almost 90% of all binding sites compared to 44% for G(C/A)AUG. We observed that binding sites containing only the canonical GCAUG motif have significantly higher peak binding scores compared to the other 13 possible 5-mers confirming GCAUG as the motif with the highest binding affinity (**Figure S2B**). All GNWYG 5-mers were specific for RBFOX2 compared to our other RAPseq assays (**Figure S2C**). Peak binding scores for the majority (86%, 12/14) of the individual RBFOX2 5-mer motif increased upon consideration of an additional 5'U (**Figure S2D**) verifying the previously reported increase in binding affinity (Lambert et al., 2014). Binding sites containing two 5-mers had higher peak binding scores than sites with only one 5-mer (**Figure S2D**). We also confirmed the generalizable GNWYG motif when inspecting eCLIP enrichments as a function of the number of GNWYG (excluding GCAUG) motifs per peak

(**Figure S2E**). Like RBFOX2, YBX3 peak binding scores increased in the presence of additional 5' adenosines with respect to the core CAHC motif (**Figure S2F**).

We next compared the concordance between RAPseq and eCLIPseq regarding motif location and read enrichment of the five RBPs. In RAPseq, the motif locations were consistent with read densities around the center of peak coordinates resembling profiles from transcription factor chromatin immunoprecipitation (ChIP) assays. This allowed the use of standardized peak calling algorithms (STAR Methods). In contrast, CLIPseq methods require assay-dependent customized algorithms (Hafner et al., 2021) due to the shift between read and motif density locations (**Figure 2C-D**). RAPseq exhibited higher reproducibility and more significant binding events than eCLIPseq (Spearman rank correlation for RAPseq $\rho=0.93-0.94$ and for eCLIPseq $\rho=0.58-0.92$, Wilcoxon rank sum test for eCLIPseq and RAPseq $p<0.01$) (**Figure 2E-F**). When combining peaks on the gene level, we noticed that the majority (56% to 83%) of RBP bound genes in eCLIPseq overlapped with RAPseq. More genes were bound by RBPs in our RAPseq assays when compared to eCLIPseq. However, binding of most of the transcripts *in cellulo* (eCLIPseq) can be attributed solely to the native RBP binding properties (RAPseq) (**Figure 2G**). In addition, transcripts bound *in vitro* were significantly stronger bound *in cellulo* (Wilcoxon rank sum test, $p<0.01$) (**Figure 2G**). Overall, RAPseq profiles quantitatively RBP-binding specificities and RBP-RNA interactomes.

Figure 2. The robustness of RAPseq is confirmed by benchmarking against state-of-the-art methods

- (A) Position weight matrices display *de novo* identified motifs of five RBPs in RAPseq and eCLIPseq assays and published motif models from three *in vitro* assays (RNAcompete, RBNS and HTR-SELEX).
- (B) LOESS fitted smoothed regression lines (left) and boxplots (right) of HNRNPC RAPseq (green, top) and eCLIPseq (grey, bottom) assays show HNRNPC binding in terms of peak binding scores (left) and fold changes (right) (y-axes) as a function of uracil content (x-axes) in a 40 nt bin centered around the peak summit (for RAPseq, top) and around the highest motif density location (for eCLIPseq, bottom).
- (C) Line and polygon peak metaplots compare the motif locations (polygons) with respect to the average normalized read enrichment (RPM, lines) for RAPseq (green) and eCLIPseq (grey) for five RBPs. Line plots show the average RPM for all peaks plotted as a fraction of the maximum mean RPM for two experimental replicates (top, y-axes). The bin size

- for the line plots is 1 nt. Polygon plots display the density of RBP binding motif locations (Figure 2A) at bound sites for RAPseq (green), eCLIPseq (grey) and the same motifs over unbound control sites (yellow). The x-axes show the distance from the peak summit. Ticks are drawn at -100, -50, 0, 50 and 100 nt from the respective RBP peak summit.
- (D) Illustrative model explains the read enrichment at (RAPseq, left) and next to the motif (eCLIPseq, right) due to intrinsic experimental differences when capturing RBP-RNA interactions.
 - (E) Scatter plots show the correlations in binding enrichments ($\log_2$ of peak fold changes) between two replicates (x, y-axes) for all RAPseq binding events (left column) and subsampled RAPseq binding events (middle column) to match the number of eCLIP binding events (right column) for five RBPs (Spearman's rank correlation coefficients).
 - (F) Comparison of the experimental steps in RAP (left) and eCLIP protocols. Dotplot correlates the mean of two replicates for five RAPseq (green) and eCLIPseq (grey) experiments. Asterisks indicate p -values (Spearman rank correlation, $**p < 0.01$).
 - (G) Two-way Venn diagrams (left) intersect the number of genes with identified binding sites for five RBPs in RAPseq (green) and eCLIPseq (grey). The intersected area is colored (yellow) and the number of genes per intersection is displayed. Violin plots (right) of eCLIPseq demonstrate gene enrichments (in $\log_2$ scale) common in RAPseq and eCLIPseq (yellow) and eCLIPseq only genes (grey). Asterisks represent p -values (one-tailed Wilcoxon rank sum test, $**p < 0.01$, $***p < 0.001$). The gene enrichments are computed as the sum of all peak enrichments (binding score) per gene.

HUR binding to uracil triplets is conserved but has species-specific transcript-processing preferences

Species radiation underlines rapid changes in the regulation of the transcriptome requiring adaptation of RBP-RNA binding. RAPseq enables cross-species comparisons between RBP orthologs and RNA substrates. Since HUR is the most divergent paralog of the ELAV gene family that evolved through gene duplication in vertebrates (Good, 1995), we quantified and compared HUR binding specificities in vertebrate evolution. We inspected HUR orthologs from human, mouse, opossum, chicken, frog and fish. All orthologs had a similar length and the amino acid sequence identity changed gradually from human to zebrafish (**Figure 3A**). Except for the frog ortholog, all amino acid changes resided outside the RNA interacting beta sheets of the three RNA recognition motifs (RRMs) suggesting no or only few differences in RNA binding (**Figure S3A**). Despite two amino acid differences in the RRM1 and RRM3 beta sheets of the frog HUR, RNA contacts should be maintained according to crystallographic models showing specific interactions of all three RRM with uracil triplets to the human structure (Ripin et al., 2019; H. Wang et al., 2013) (**Figure S3A**). To validate that binding specificity is conserved, we first produced HUR proteins of the tetrapod orthologs and performed RAPseq assays with the same human RNA substrate (**Figure 3B**). We found that tetrapod HURs bound to uracil triplets (UUU), which were significantly enriched in more than 90% of all ortholog binding events (**Figure 3C** and **S3A**). The uracil triplet was localized centrally at the peak summit over background control sites (**Figure 3D**) (STAR Methods). The fold changes were proportional to the number of triplets per binding site implying tetrapod-conserved RRM cooperativity in substrate recognition (**Figure 3E**). However, we noticed differences in global enrichment profiles. The frog HUR had the lowest enrichments compared to the other tetrapod HURs suggesting that some of the amino acid changes caused differences in RNA binding affinities (Wilcoxon rank sum test, $p < 0.001$) (**Figure 3F**).

We grouped HUR binding by orthologous species intersections (**Figure S3B**) and averaged the peak binding scores of the tetrapod HURs for all RNA binding events to quantify the extent of binding similarities across all orthologs (**Figure S3C**). Average peak binding

scores were higher in the tetrapod-conserved than species-specific RNA binding events (**Figure 3G**). To assess the degree of evolutionary conservation of HUR binding sites, we used phyloP scores obtained from 100 vertebrate genome alignments. We found that ortholog-shared and -unique RNA binding sites were highly constrained compared to adjacent RNA sequences used as control (**Figure 3H**) (STAR Methods). Remarkably, the underlying sequence was under stronger selection in the common than the ortholog-unique RNA binding sites (**Figure 3H**) and followed the same trend as the average peak binding scores (**Figure 3G**) and uracil triplet count (**Figure 3I**). This suggested positive selection of uracil triplet expansion serving as HUR binding sites in the vertebrate lineage.

Given that the zebrafish HUR is most diverse across the vertebrate lineage, we profiled zebrafish and human HUR proteins using zebrafish RNA as substrate (**Figure 3B**). We found that the zebrafish and human HURs recognized uracil triplets in the zebrafish RNA with similar specificity and cooperativity (**Figure S3D-F**). However, the zebrafish HUR had significantly lower enrichments than the human HUR (**Figure S3G**). By inspecting HUR binding events across the transcript body, we noticed a higher fraction of intronic sites in the zebrafish transcriptome compared to human, where most HUR sites reside in 3'UTRs (**Figure 3J**). To determine pathways that could be affected by this divergence, we examined the 94 Gene Ontology (GO) terms commonly bound in human and zebrafish transcriptomes. In this common set, the binding of zebrafish HUR occurred at a nearly even proportional frequency to zebrafish intronic (47%) and 3'UTR (53%) sequences whereas human HUR binding to human substrate was more prevalent in human 3' UTRs (75%) relative to introns (25%) (**Figure S3H**). We next inspected the proportion of HUR binding sites within introns and 3'UTRs for each of the 94 GO terms (**Figure S3I**) and discovered the highest divergence in gene products mediating cellular responses to insulin in which 80% of HUR binding sites occurred in introns in zebrafish compared to 20% in humans (**Figure S3J**). Given that HUR is the only paralog with both nuclear and cytoplasmic functions in vertebrates and the original strictly nuclear function of ELAV proteins in invertebrates (Good, 1995), our results suggest an evolutionary shift in HUR function.

In conclusion, our cross-species combinatorial RAPseq experiments revealed vertebrate-conserved HUR binding to uracil triplets. Vertebrate HUR binding sites expanded selectively and affinities gradually increased from zebrafish to human. This allowed HUR to adapt predominant post-transcriptional roles from mRNA processing in zebrafish to mRNA stability in human (**Figure 3K**).

Figure 3. RNA binding specificity of HUR is conserved in vertebrates but acquired functional differences

- (A) Orthologous vertebrate HUR protein length and RNA recognition motifs (RRM1-3) (green boxes) are shown (color-coded by species). Amino acid (aa) changes (black arrows) for each ortholog compared to the human reference are highlighted.
- (B) Schematic representation of the performed RAPseq assays using various vertebrate HUR proteins (colored) and either the human (left) or zebrafish RNA substrate (right).
- (C) Table lists the percentage and corresponding DREME E-value of uracil triplet motifs in the HUR-bound transcriptome (± 25 nt around the peak summit).

- (D) Line plot (smoothing band width 2 nt) shows frequency of uracil triplets $\pm$ 100 nt at HURs peak summits (green, color-code by species) and control sites residing 1000 nt distant from the peak summit (grey, color-code by species).
 - (E) Binding site enrichments (peak fold change, y-axis) are shown as a function of the frequency of uracil triplets (x-axis) for all HUR orthologs. LOESS fitted smoothed regression lines with 95% confidence interval display the averaged enrichments (Figure S3C) across all (top left) and for each individual ortholog (bottom panels). Boxplots (top right) show enrichment distributions (orthologue average fold change) of each uracil triplet bin.
 - (F) Boxplots show peak fold change per ortholog (color-coded by species). Asterisks represent *p*-values (one-tailed Wilcoxon rank sum test, ****p*<0.001).
 - (G-I) Plots represent (G) average HUR binding (scored as log₂ transformed average binding score), (H) degree of sequence conservation of HUR bound sites compared to adjacent unbound control sites (according to 100 vertebrate phyloP scores) and (I) proportional frequency (0-5) of uracil triplet motifs within the bound sites per orthologous species intersections (1-5) (Figure S3B).
 - (J) Pie chart shows the percentage of HUR binding sites for a given RNA feature (color-coded) in the combinatorial human and zebrafish cross-species RAPseq assays.
 - (K) Proposed model summarizes the co-evolutionary dependencies between HUR protein and transcriptomes in vertebrates. Top: From zebrafish to human, HUR binding transitioned from mRNA processing to mRNA stability (inverse triangles). Middle: During evolution, the size of the actively transcribed genome doubled from zebrafish to human while HUR binding sites increased 5-fold (line plots). Bottom: HUR binding specificity for uracil triplets remained unchanged (motif in the background) but HUR binding affinities changed across vertebrates (colored species cartoon).
- Hs, human; Mm, mouse; Md, opossum; Gg, chicken; Xt, frog; and Dr, zebrafish.

HUR and PTBP1 cooperate post-transcriptionally by keeping a specific distance between binding sites

Previous reports have shown that HUR and PTBP1 act cooperatively and competitively to regulate mRNA processing (Galbán et al., 2008; Shwetha et al., 2015). In our assays, HUR and PTBP1 sites overlapped with a bimodal distance distribution separated at 30 nt. Proximal sites (0-30 nt, *n*=797) could result in competitive binding events, whereas distal peaks (31-100 nt, *n*=422) indicated co-binding (**Figure 4A**). We statistically inferred the outcome of competitive events and quantified that 17% and 22% of the proximal peaks could be significantly attributed to one or the other RBP (fold change ≥ 2 , FDR-adjusted *p*≤0.05) (**Figure 4B**).

To test cooperative binding properties of the two RBPs, we developed co-RAPseq, in which RNA binding is assayed in the presence of both proteins at the same time (**Figure 4C**). We hypothesized that in contrast to the single RBP assays, cooperative events determined by co-RAPseq lead to wider binding sites (peak width) due to the concomitant presence of the two proteins on the same RNA substrate (**Figure 4D**). While the peak width of the single RBP assays was 8 nt (HUR) and 9 nt (PTBP1) above median library size, co-RAPseq assays with both factors resulted in significantly wider 37 and 24 nt peak distributions, respectively (one-tailed Wilcoxon rank sum test, *p*<0.001) (**Figure 4E**). By measuring binding site enrichments in the co-RBP assay as a function of the distance between HUR and PTBP1 motifs, we determined that cooperativity was optimal at a 27 nt distance (**Figure 4F**). This optimal distance was neither resolvable in the single HUR and PTBP1 RAPseq assays nor when mapping co-RAPseq enrichment scores to the distances between motifs at control unbound sites (**Figure S4A-C**).

Of the 6,885 genes for which we detected binding events in either single or co-RAPseq assays, 873 were only present in the co-RAPseq assay (**Figure S4D**), suggesting that under

steady-state conditions the post-transcriptional regulation of these gene products depended on the co-activity of both factors. In the single RAPseq assays, 86% (2,874/3,324) of the HUR and 59% (2,855/4,862) of the PTBP1-bound genes overlap with the co-RAPseq assay (**Figure S4D**). Of those genes, 19% (536/2,874 for HUR) and 22% (636/2,855 of PTBP1) were significantly stronger bound when both RBPs acted together (**Figure S4E**). In comparison to all 6,885 HUR- and PTBP1-bound genes, the uniquely and differentially bound genes in the co-RAPseq assay were functionally enriched in innate immune response pathways (**Figure S4F**) as previously observed for each factor individually (Christodoulou-Vafeiadou et al., 2018; Geng et al., 2021; Rothamel et al., 2021; Sasanuma et al., 2019; Sueyoshi et al., 2018). Remarkably, the strongest cooperative sites resided within introns of these pathways signifying cooperative functions of HUR and PTBP1 in mRNA pre-processing (**Figure S4G**).

Figure 4. HUR and PTBP1 control RNA processing cooperatively with an optimal distance between their binding sites

- (A) Venn diagram shows the overlap between HUR and PTBP1 peaks (top). Density plot (bottom) demonstrates the distance between peak summits of overlapping peaks. The x-axis represents the distance in nucleotides (nt) (log₁₀ scale, bandwidth 0.15). The bimodal distribution is separated at 30 nt (dashed line).
- (B) Scatter plot compares HUR- and PTBP1-bound competitive peaks (dots, read count normalized, log₂ transformed). Peaks either differentially bound by PTBP1 (light green) and HUR (dark green) or not differentially bound (grey) are highlighted (fold change ≥2, FDR-adjusted $p \leq 0.05$). Size and transparency of the dots are proportional to the absolute fold change.
- (C) Schematic representation of the co-RAPseq assay for HUR (dark green) and PTBP1 (light green)
- (D) Illustration of the expected effect on peak widths of cooperative binding events

- (E) Venn diagrams (top) show peak overlaps between the co-RAPseq with HUR (left) and PTBP1 (right) RAPseq assays, respectively. The intersections represent the binding sites used to assess the differences in peak widths. Boxplots (bottom) demonstrate differences in peak widths between the single- and co-RAPseq assays. The y-axis shows peak deviation from the median library size in nt per assay. Color codes are as in Figure 4C. Asterisks represent the significance of two-sided Wilcoxon rank sum tests between single- and co-RAPseq assays (***, $p < 0.001$).
- (F) Model summarizing the cooperative binding of PTBP1 and HUR to their respective motifs (top). LOESS fitted regression line shows the trend in binding site enrichments (y-axis) when plotted as a function of the distance between the HUR and PTBP1 motifs (x-axis). The fold change in RNA binding is highest at 27 nt (dashed line).

T7-RAPseq reveals m⁶A-dependent YTHDF1 binding to mediate mRNA regulation

To assess the transcriptome-wide relevance of RNA modifications in RBP-RNA recognition, we developed T7-RAPseq. We generated a modification-devoid RNA substrate by *in vitro* transcribing the native RNA substrate using T7 RNA polymerase (Figure 5A). When comparing RAPseq and T7-RAPseq, our results showed identical transcript distributions, high correlations (Spearman $\rho = 0.98$ and Pearson $r = 0.97$) (Figure 5B) and uniform read coverage over spliced and unspliced long and short transcripts (Figure S5A) confirming that the *in vitro* transcription step maintains substrate composition.

We applied T7-RAPseq to determine modification-dependent targets of the m⁶A binder YTHDF1 (X. Wang et al., 2015). According to previous reports (X. Wang et al., 2015), YTHDF1 binds to a preferred GRAC (R is an A or G) motif of the m⁶A DRACH consensus sequence (D is A, U or G and H is A, U or C). About 75% (1,414/1,895) of the identified YTHDF1 binding sites were strictly dependent on the presence of m⁶A (Figure 5C, Table S4) and the binding affinity drastically decreased in the absence of RNA modifications (one-tailed Wilcoxon rank sum test, $p < 0.001$) (Figure 5D) corroborating m⁶A dependency in RNA binding. We observed that GGACU was the most abundant m⁶A-DRACH motif (Figure 5E) either as a 5- or its two 4-mers GGAC and GACU (Figure S5B). Although enrichment scores for the pentamer were higher than for the individual tetramers (one-tailed Wilcoxon rank sum test, $p < 0.001$) (Figure S5C), the enrichment scores are comparable when the two tetramers can act cooperatively (one-tailed Wilcoxon rank sum test, $p < 0.001$) (Figure S5C). This demonstrated that GGACU is the optimal binding motif and the two suboptimal tetramers (GGAC and/or GACU) can increase the frequency of YTHDF1 interactions with RNA and/or induce the formation of YTHDF1 dimers. In accordance with the m⁶A dependency of YTHDF1 (Xu et al., 2015), GGACU pentamer containing binding sites with a m⁶A had higher enrichment scores when compared to m⁶A depleted motifs (one-tailed Wilcoxon rank sum test, $p < 0.001$) (Figure 5F).

We next investigated pathways that could be affected by the presence or absence of YTHDF1 m⁶A binding. We found 48 significantly enriched (FDR-adjusted $p < 0.01$) parent GO terms with the highest m⁶A dependency in genes involved in stem cell maintenance and chromatin binding (Figure 5G and S5D). We used public m⁶A-sequencing data to validate the top 4 binding sites in MLLT6, ZFP36L2, POLR2D and MBD3 mRNAs (Figure 5G-H) (Dominissini et al., 2012). In conclusion, the deterministic difference in RAPseq and T7-RAPseq data allowed to discern a clear m⁶A-dependent on-off switch for YTHDF1 binding (Figure 5I) and that T7-RAPseq can be used to distinguish genome-wide contributions of RNA modifications to RBP binding activities.

Figure 5. m⁶A is necessary for YTHDF1 to bind RNA at steady-state protein levels

- (A) Schematic illustration of the T7-RAPseq method
- (B) Scatter plot compares the T7 *in vitro*-transcribed (T7-RAPseq) and the native RNA (RAPseq) substrates (log₂ transformed transcripts per million, TPM). Numbers within the plot show Spearman (ρ), Pearson (r) correlation coefficient, linear regression (adjusted R^2) and the number of observations. Density plots (right and top, bw=0.2) show the overall distributions of TPMs and their medians (4.1 log₂ TPM, dashed line).
- (C) Venn diagram overlaps YTHDF1 binding sites between the native RAP and the T7-RAP substrates.

- (D) Violin and overlaid box plots show overall YTHDF1 enrichments ($\log_2$ transformed peak binding scores + 1) in both RAPseq (green points) with T7-RAPseq (grey points) replicates. Each point represents one binding site ($n=1,895$). Asterisks represent significance (Wilcoxon one-tailed rank sum test ***, $p<0.001$).
- (E) Barplot shows the frequency of DRACH 5-mers motif within YTHDF1 peaks normalized by their abundance at control sites. Top motif is highlighted (green).
- (F) Boxplots demonstrates the peak fold change of YTHDF1 enrichments in RNA substrates ($n=1,059$) with a modified (RAPseq, green) or modification depleted (T7-RAPseq, grey) GGACU motif. Dashed lines mark fold change values of 1 and 2. Asterisks represent the significance (Wilcoxon one-tailed rank sum test ***, $p<0.001$).
- (G) Dotplot illustrates ranking m^6A -dependent YTHDF1 binding sites within mRNAs of genes belonging to two GO terms (Figure S5D). RAPseq binding site signal (library size normalized read counts $\log_2$ transformed, y-axis) is ranked (x-axis). Top 4 targets are highlighted.
- (H) Genome tracks show YTHDF1 binding to m^6A within the 3'UTRs of the top 4 binding sites (MLLT6, MBD3, POLR2D and ZFP36L2) (Figure 5G). Genomic locations with scale bar indicate the length of the genomic region in bases and gene features (blue rectangle: exon and 3'UTR, arrow: direction of transcription) (x-axis) and normalized reads per million (RPM) (y-axis) are shown. Vertical yellow lines highlight bound site. Tracks for two replicates using native (dark green) or T7 *in vitro* transcribed (grey) input RNA as well as m^6A -specific RNA immunoprecipitation (IP, light green) and the respective input control are presented.
- (I) Model illustrates m^6A -dependent roles of YTHDF in post-transcriptional regulation. The presence of m^6A (yellow circle) enables YTHDF1 (green reniform cartoon) binding to its RNA targets (grey box) and thereby enhances their stability and translation under steady state conditions. In contrast, the absence of m^6A drastically abolishes RNA binding of YTHDF1 resulting in RNA decay and decreased translation.

Non-canonical RBPs have specialized RNA interactomes

Despite the increased identification of ncRBPs through latest technological advances, knowledge about the modes of ncRBP-RNA binding is limited. In order to test whether RAPseq can discover RNA targets of ncRBPs, we intersected RBPomes identified in four independent studies (Castello et al., 2012; Queiroz et al., 2019; Trendel et al., 2019; Urdaneta et al., 2019). Across all assays, one third (243/718) of all common RBPs were ncRBPs (**Figure 6A**). We selected 26 ncRBPs containing different domain composition and canonical functions (**Table S1**). While cRBPs bound on average to 9,574 sites, we observed two groups of ncRBPs binding to either more or fewer than 1000 sites using RAPseq. Ten ncRBPs had a mean of 1,757 (range: 1,102-3,134) and the remaining 16 on average 113 (range: 7-534) binding sites (**Figure S6A**). PEBP1 had the highest (3,134) and CCDC124 (7) the lowest number of binding sites (**Figure S6B**). While cRBPs had a median of 2.5 (range: 2.1-3.2) binding sites, ncRBPs had one binding site across the gene body (**Figure S6C**).

By determining interactions between ncRBPs and RNA classes, we found that most binding sites (median: 97.2%, range: 60-100) resided in mRNAs. Compared to the other ncRBPs, RAN exhibited preferences for exonic regions and PRDX6 for 5'UTRs. CCT2 and TKT had the highest proportional enrichment for long and HSPA5 for small noncoding RNAs (ncRNAs) (**Figure S6D**). In comparison to cRBPs, ncRBPs exhibited weaker binding strengths (**Figure S6E**). HSPA8 and PEBP1 displayed both the highest number of binding sites and the strongest enrichments (**Figure S6C and S6F**). These results implied more specialized functions as post-transcriptional regulators as well as a higher dependency on regulatory signals and interactions due to less *bona fide* functions in competitive RNA binding.

To assess the specialization for each ncRBP-RNA interactome, we performed a functional gene enrichment analysis of the bound mRNAs. For 14 ncRBPs, we identified 48 GO terms that were significantly and specifically enriched (FDR-adjusted $p \leq 0.05$) (**Figure 6B**). GO terms with the highest percentage of binding sites in 5' and 3'UTRs were obtained for genes enabling apolipoprotein binding (for FSCN1) and negative regulation of inflammatory response (for STMN1). In addition, we found mRNA enrichments for genes in cortical cytoskeleton (for MAPRE1) and epithelial cell migration (for STMN1). The canonical function of these two ncRBPs is to either promote (MAPRE1) or destabilize (STMN1) microtubule elongation (**Figure 6B-D**) (Askham et al., 2002; Steinmetz et al., 2000; van Haren et al., 2018; Wittmann et al., 2004). Since we profiled the MAPRE1- and STMN1-RNA interactomes in liver cancer cells (HepG2), we inspected the association with hepatocellular carcinoma (LIHC). MAPRE1 and STMN1 have been linked to promote liver cancer progression through mechanisms exerted by the canonical functions in cytoskeletal remodeling (Aiyama et al., 2020; Lu et al., 2018; Orimo et al., 2008; Wong et al., 2008). We found that upregulation of MAPRE1 and STMN1 gene expression correlated with the severity of the pathological stage and significantly decreased patient survival in LIHC patients (**Figure 6E-F**). A subset of MAPRE1 and STMN1 RNA targets (**Figure 6C-D**) was significantly associated with poorer patient survival in LIHC (**Figure 6G**). Thus, as post-transcriptional regulators, MAPRE1 and STMN1 could increase LIHC progression by either enhancing stability and/or localized translation of their mRNA targets.

Figure 6. ncRBPs bind to cancer-associated transcriptomes

(A) Four-way Venn Diagram overlaps reported RBPs identified by RNA interactome capture (RIC), orthogonal organic phase separation (OOPS), protein and phenol-toluol extraction (PTex) and cross-linked RNA extraction (XRNAX). The number of RBPs discovered by

- each method is written in parenthesis. Common RBPs (white) are divided into canonical or noncanonical.
- (B) Circle plots (left) display ncRBP-RNA interactomes functionally enriched in gene ontology (GO) terms. Circle diameter is proportional to the number of enriched genes per category; circle color corresponds to the type of GO term (BP: biological process, CC: cellular compartment, MF: molecular function) and circle transparency to FDR-adjusted *p*-values. Stacked barplot (right) shows the proportional binding to mRNA features (color-coded) for each GO term.
- (C, D) Interaction network shows functionally enriched (C) MAPRE1- and (D) STMN1-RNA interactomes. Nodes represent gene names of MAPRE1 and STMN1-bound mRNAs and color displays peak binding scores intensities for all binding sites (sum log₂ transformed; light blue: low, dark blue: high) mapping to the gene representing the node. Highlighted genes (yellow) are further studied in G.
- (E, F) Liver Hepatocellular Carcinoma (LIHC) association of (E) MAPRE1 and (F) STMN1. Boxplots (left) show MAPRE1 and STMN1 expression levels (log₂ TPM+1) in LIHC patients (red) and healthy human liver tissues (green) from TCGA and GTex datasets. Asterisks indicate *p*-values (one-way ANOVA F test, ***p*<0.01). Violin plots (middle) show increasing MAPRE1 and STMN1 expression levels with LIHC stage severity. Linear regression and *p*-values (one-way ANOVA F test) are noted. Kaplan-Meier plots (right) show association of MAPRE1 and STMN1 low (green) and high (yellow) expression levels and ten-year survival within the LIHC cohort. Low and high groups are separated by median gene expression level. Patient number (n) per group, hazard ratios (HR) and *p*-values (Mantel-Cox test) are reported.
- (G) Kaplan-Meier plots display an association between expression levels of genes highlighted in panels C and D (yellow) and overall survival (as in E and F, right). Corresponding hazard ratios (HR) and *p*-values (Mantel-Cox test) are shown (n=180-183 patients for each high and low group).

Cancer-associated variants of the IGF2BP family show altered RNA binding

The family of insulin-like growth factor-2 mRNA binding proteins (IGF2BP) comprises three paralogs (IGF2BP1, IGF2BP2 and IGF2BP3). In human and mouse, IGF2BP genes show oncofetal expression patterns with high expression during embryonic development and in liver cancer (**Figures S7A-C**) (Cardoso-Moreira et al., 2019; Rudolph et al., 2016). We inspected frequently occurring IGF2BP variants in cancer types and found one common amino acid change in IGF2BP1 (R167C/H) and two in IGF2BP3 (I474M and R525C) (**Figure 7A**). Frequent missense mutations were absent in IGF2BP2 but a cancer-associated splicing variant (isoform B) (**Figure 7A**) was previously identified (Tybl et al., 2011; J.-Y. Zhang et al., 1999).

We generated IGF2BP paralogs (wt) and cancer-associated variants to compare their RNA binding capacities by RAPseq (**Figure 7B**). The paralogs showed highly similar binding preferences towards CA-rich motifs, as previously reported (Conway et al., 2016; Schneider et al., 2019). We found that these motif preferences were maintained in the cancer-associated variants (Spearman rank correlations, median=0.96, range: 0.92-0.99, n=28) (**Figure S7D**). In contrast, k-mer enrichment profiles differed when comparing IGF2BP paralogs and variants to either RBFOX2 or YTHDF1 (**Figure S7D**). Compared to the IGF2BP1 paralog, the cancer-associated variant IGF2BP1-R167C showed similar but IGF2BP1-R167H had overall lower signals (**Figure 7C**). We found no global differences between the two IGF2BP2 isoforms. IGF2BP3 variants exhibited weaker binding strengths when compared to the wt (**Figure 7C**). By testing differential RNA binding, we found no differences between the wt and the IGF2BP1-R167C variant but the wt bound to significantly more sites (n=962) than the IGF2BP1-R167H variant (n=34). The two IGF2BP2 variants had a small and balanced number of differently

bound sites (56 for isoform B and 70 for isoform A). IGF2BP3 bound to more sites than the two mutants (268 *versus* 15 for IGF2BP3-I474M and 102 *versus* 3 for IGF2BP3-R525C) (**Figure 7D**). We noticed that the differentially bound sites of IGF2BP2 isoform B occurred more frequent in ncRNAs than those of isoform A (Chi-squared test, $p < 0.003$) (**Figure 7E**). Thirteen ncRNAs were more bound by isoform B compared to only four by isoform A and the highest differences were observed for snoRNAs and miRNAs (**Figures S7E-F**).

At the pathway level, we observed significant differences (FDR-adjusted $p \leq 0.05$) between variants and respective wt paralogs despite the small size effects (**Figure 7F**). The IGF2BP1-R167H variant showed a slight increase in binding towards pathways regulating RNA metabolism and a decrease in binding to factors involved in DNA damage response. This could suggest a gain of function (GoF) shift towards promotion of growth and loss of function (LoF) in regulation of factors safeguarding genotoxic events. For the pathological IGF2BP2 isoform B, we observed a GoF in DNA damage response and a LoF in growth promoting pathways. For IGF2BP3-I474M, we noticed LoFs in chromosome and centrosome organization during mitosis, which as for IGF2BP1-R167H could result in genotoxic events plausibly through missegregation during anaphase and a GoFs in calcium-dependent overproduction of ATP and growth promotion. In the IGF2BP3-R525C mutant, a GoF controlled binding to growth and to mobility promoting mRNAs as well as a LoF affecting binding of the chaperonin complex CCT/TriC. Overall, our data suggested that the profiled mutations alter binding strength but not binding specificities, which can result in post-transcriptional deregulation of their RNA targets and related pathways.

Figure 7 Pathological variants of the IGF2BP family cause functional gains and losses in RNA binding

- (A) Bars illustrate RRM and KH domain locations in each paralog protein (length in amino acids, aa, x-axis). Lollipop plots show aa mutation frequencies (y-axis) in cancer patients. Mutants selected for RAPseq profiling are color-coded. The pathogenic isoform B of IGF2BP2 with the skipped exon 10 is drawn.
- (B) Schematic representation of RAPseq assays performed to profile IGF2BP paralogs and pathological variants (color-coded).

- (C) Density plots represent the distribution of RAPseq signals (read counts normalized to HaloTag in $\log_{10}$) for each IGF2BP variant with respect to their wild types. Dashed lines indicate medians. The number of observations (n) is shown in each panel.
- (D) MA plots demonstrate differentially bound sites (DBS) between wild type and variants (y-axis) as a function of the average signal per binding site (x-axis, in $\log_2$ counts per million). Data points and number represent significantly (color-coded by variant) and not significantly DBS (grey) sites (fold change ≥ 2 , FDR-adjusted $p \leq 0.05$).
- (E) Stacked barplots show the percentage of DBS (y-axis) binned per major RNA feature for each wild type and variants (color-coded) (x-axis). Bars of the same color sum up to 100%. Significant difference of DBS in 3'UTR and ncRNA of the two IGF2BP2 isoforms (Chi-squared test with continuity correction p -value) are highlighted.
- (F) Horizontal barplots display the top 10 differentially bound Reactome pathways between variants and wild types (color-coded, fold change ≥ 2 , FDR-adjusted $p \leq 0.05$).

DISCUSSION

We have developed RAPseq, a method that provides *in vitro* derived RBP-RNA interactomes by using recombinant RBPs and native cellular RNA substrates with endogenous epitranscriptomic modifications maintained. RAPseq can be applied in versatile and large-scale transcriptome-wide studies to address fundamental questions in post-transcriptional gene regulation.

First, RAPseq allowed us to design combinatorial cross-species RBP-RNA interaction assays to dissect vertebrate conservation of HUR's binding specificities and the divergence of its intronic and 3'UTR binding sites between lower and higher vertebrates. Second, we used RAPseq to identify parameters for optimal cooperativity between HUR and PTBP1 and pathways that benefit from this co-binding activity. Third, we produced a modification depleted version of the native RAPseq substrate (T7-RAPseq) and discovered which YTHDF1 binding sites depend on the presence of m⁶A. Forth, we demonstrated that RAPseq can overcome one of biggest limiting factors in current RNA binding assays by multiplexing the profiling of 26 ncRBP-RNA interactomes. We concluded that ncRBPs have specialized rather than global post-transcriptional regulatory roles and found that some moonlighting activities coincided within the same biological processes described for the canonical functions. Fifth, we profiled cancer-associated variants of the IGF2BP family. Except for IGF2BP2, the other variants bound to a larger extent to RNA substrates involved in pathways promoting cell growth. This suggested that increased RNA binding activities of frequently mutated IGF2BPs could provide proliferative advantages in cells of cancer patients.

We showed that RAPseq can be used by individual laboratories to enable consortia-like studies by describing applications of the highest interest. Additional method developments, including co- and T7-RAPseq, extended the spectrum of practical implications. Increasing the proportion of currently underrepresented intronic sequences (**Figure S2G**) in the RAPseq substrate, i.e., by purifying RNA from nuclear isolates will be of relevance in the future to assess alternative splicing events that have co-RBP regulatory dependencies. The increasing number of ncRBPs without known RNA binding functions but with well-established roles in cellular metabolism, opens questions regarding contextuality of their RNA binding activity. For example, binding to RNA and executing known functions could occur simultaneously or may be mutually exclusive. RNA-enzyme-metabolite (REM) networks (Hentze & Preiss, 2010) proven so far, have demonstrated a few examples of linking interdependent functions. RAPseq could be used to perform RNA binding assays in the presence or absence of such metabolites across cell type-specific transcriptomes varying in composition to assess increased or disrupted interdependency of REM networks.

Altogether, RAPseq is a versatile and above all enabling tool that will ease and pave the way towards the description of the complete RBPome and help envision new types of studies that would be unfeasible otherwise.

Acknowledgments

This work was supported by KI KID KI-KID funding (2016–00189, CK), Knut & Alice Wallenberg foundation (KAW 2016.0174, CK); Ruth & Richard Julin foundation (2017–00358, 2018–00328, 2020-00294 CK); SFO-SciLifeLab fellowship (SFO_004, CK); Swedish Research Council (2019-05165, CK); Lillian Sagen & Curt Ericsson research foundation (2021-00427, CK); Gösta Milton’s research foundation (2021-00527, CK); Onassis Foundation (F ZO 057-1/2018-2019, SP), Karolinska Institute (2019-01080 and 2019-01102, IA), Erasmus+ (20170801, IA; 20200115, NP); Swedish National Infrastructure for Computing (SNIC) at UPPMAX (SNIC 2017/7-260, uppstore2018110, SNIC 2020/15-293, SNIC 2020/16-228, CK-IA). We would like to thank the National Genomics Infrastructure in Stockholm funded by Science for Life Laboratory, the members of the Marc Friedländer, Vicent Pelechano, and Claudia Kutter laboratories as well as Carsten Daub for helpful feedback.

Author contribution

CK conceptualized the study. IA and CK designed experiments and computational analysis. IA, SP and NP performed experiments. IA and SP performed computational analysis. CK acquired funding. IA and CK wrote the manuscript. All authors read and approved the final manuscript.

Declaration of Interests

The authors declare no competing interests.

Supplementary Figures

Figure S1. RBPs are produced as C-terminal Halo fusions and the RAPseq substrate is produced by random fragmentation. Related to Figure 1

(A) Agarose gel (left) and violin plot (right) show size distribution of the sequenced input library (34.5 million reads). (B) Agarose (left) and SDS polyacrylamide gel (right) represent *in vitro* transcribed RNA (left) and *in vitro* translated protein products of N- and C-terminal RBP-HaloTag® fusion products resulting in truncated or full-length protein products, respectively. Protein products are visualized by HaloTag® bound to Halo Alexa488-Ligand before gel electrophoresis. N-terminal tagged CCT3 displays a mix of full length and truncated translation products. (C) SDS-PAGE gels of all described RBP-Halo fusion proteins. The percentage of loaded protein refers to the amount of *in vitro* translation (ivTLN) reaction used for the RAP binding assay. Theoretical RBP molecular weight (MW) in kDa is reported. A recombinant Halo-GST fusion protein was used as a titration curve. (D) Illustration and bar plots confirm the amount of RNA recovered before (top) or after (bottom) the RAP binding assay using *in vitro* translated HUR-Halo fusion proteins in either the absence (left) or presence (right) of RNase A (barplots, error bars show $\pm$ SEM of three replicates). The eluted RBP bound RNA generally ranges from 1 to 5 ng depending on the RBP, higher amounts are indicative of contaminants from the *in vitro* transcription and translation.

Figure S2. RAPseq confirms and identifies new secondary motifs. Related to Figure 2

(A) Table lists the percentage and corresponding DREME algorithm E-values of motifs in the RBP-bound transcriptome. (B) Boxplots show peak binding scores and number (n) of peaks containing one of the 14 RBFOX2 'GNWYG' motifs. Asterisks represent *p*-values (one-tailed Wilcoxon rank sum test, **p*<0.05, ****p*<0.001). (C) Metaplots compare the frequency of the 14 RBFOX2 'GNWYG' motifs (polygons) with respect to the RBFOX2 (light blue) and three control RBPs (HNRNPA1, HNRNPC and PTBP1) (yellows) peak summits (± 100 nt, bin size 5 nt). Asterisks represent *p*-values (binomial test for differential central enrichment in RBFOX2 to the three control RBPs, **p*<0.05, ****p*<0.001). (D) Boxplots display peak binding scores as a function of the presence of an additional 5' uracil at each GNWYG 5-mer (top) and as a function of number of 5-mers per peak (bottom). (E) Boxplots show peak fold changes as a function of number of RBFOX2 'GNWYG' motif occurrences in RAPseq (blue) and eCLIPseq (grey) assays. Asterisks represent *p*-values (one-tailed Wilcoxon rank sum test, **p*<0.05, ****p*<0.001). (F) Boxplots display peak binding scores as a function of the presence of one or two additional 5' adenosines at the three 4-mers of the YBX3 'CAHC' core motif. (G) Stacked barplots represent the fraction of peaks for six RBPs residing in a mRNA feature in RAPseq (blue) and eCLIPseq (grey) assays.

Figure S3. Combinatorial cross-species RAPseq assays dissects HUR conserved and divergent roles. Related to Figure 3

(A) Crystal structures highlight amino acid differences (black arrows) of HUR orthologs in beta sheets (green) or alpha helices (grey) in the RRM1 to 3. The human HUR is in complex with RNA (uracils in red). (B) Schematic representation of intersections used in Fig. 3G-I (x-axes). (C) Table exemplifies calculations of ortholog average scores. (D) Table lists the percentage of peaks with an uracil triplet (± 50 nt around the peak summit) and corresponding DREME E-value of for the human and fish HUR orthologs in RAPseq assays with the fish substrate. (E) Line plot (smoothing band width 3 nt) shows frequency of tetrapod HUR binding to uracil triplets ± 100 nt of the peak summit (dark blue, human; light blue, zebrafish) and control sites residing 1000 nt distant from the peak summit (dark grey, human; light grey, zebrafish). (F) Boxplots display peak fold changes by number of uracil triplets ± 100 nt of the peak summit (human: dark blue, zebrafish: light blue). (G) Boxplot shows peak fold changes over all binding sites of human (n=3,191) and zebrafish (n=2,975) HUR using zebrafish RNA. Asterisks represent p -values (one-tailed Wilcoxon rank sum test, *** $p < 0.001$). (H) Venn diagram (top) overlaps GO terms of genes bound by human and zebrafish HUR using zebrafish RNA. Pie charts (bottom) display the percent of binding sites within introns and 3'UTRs of genes in the overlapping GO terms (yellow) in human (left) and zebrafish (right). (I) Scatter plot shows the percent of intronic HUR binding sites with respect to 3'UTR for genes in each GO term as a function of the ratio between the percentage of zebrafish and human intronic sites. Each point represents one overlapping GO term (94, yellow) with the top intronic binding site bearing GO terms highlighted (blue). (J) Bar plots (left, human; right, zebrafish) show the percent of binding sites within introns (dark blue, human; light blue, zebrafish) and 3'UTRs (white) of GO terms highlighted (blue) in Figure S3I. GO terms are ranked based on the percentage of intronic sites in zebrafish.

Figure S4. HUR and PTBP1 cooperate to regulate interferon response pathways and secretory processes in HepG2 transcriptomes. Related to Figure 4

(A-C) LOESS fitted regression lines show the trend in binding site enrichments (y-axis) when plotted as a function of the distance between the HUR and PTBP1 motifs (x-axis). Enrichment scores are mapped to the distance of adjacent unbound control sites in (A) HUR and PTBP1 co-RAPseq, (B) single HUR and (C) PTBP1 RAPseq assays. (D) 4-way VENN diagram overlaps genes between co- and single RAPseq assays. (E) Volcano plots demonstrate differentially bound (DB) genes between co-RAPseq *versus* HUR (left) and PTBP1 (right) RAPseq. VENN diagrams (top) illustrate which genes were tested. Genes either differentially bound in the co-RAPseq (dark blue), HUR (blue), PTBP1 (light blue) RAPseq assay or non-DB (grey) are highlighted (fold change ≥ 2 , FDR-adjusted $p \leq 0.05$). Number of DB genes are mentioned. Size and transparency of the dots are proportional to the absolute fold change. (F) Circle plot displays gene ontology (GO) terms enriched for genes bound only and more in the co-RAPseq assay highlighted in blue in the 3-way VENN diagram (top) using all other genes identified as background gene set. Circle diameter is proportional to the number of enriched genes per GO term (CC: cellular compartment, BP: biological process) and color intensity to FDR-adjusted p -values. (G) Pie chart (top) illustrates the distribution of the binding sites to mRNA features in the GO terms listed in Figure S4F. Boxplots (bottom) show fold-change of binding site enrichments for each mRNA feature. Number of binding sites mapping to each mRNA feature (color-coded) are indicated.

Figure S5. YTHDF1 exhibits extensive m⁶A dependencies as an RNA-binding protein. Related to Figure 5

(A) Genome tracks demonstrate similarities between the native (RAP input, blue) and *in vitro* transcribed (T7-RAP input, grey) substrates used for profiling YTHDF1 binding. Tracks exemplify an even transcriptome and fragment representation between the two substrates over a long spliced (top), long unspliced (middle) and small (bottom) transcripts. Genomic locations with scale bar indicate the length of the genomic region in bases and gene features (black rectangle: exon and UTR, grey line: intron, arrow: direction of transcription) (x-axis) and normalized read density (RPM, y-axis) for RAP- and T7-RAPseq are shown. (B) Dotplot highlights the enrichment of k-mers of various lengths containing the GGACU motif or its two encompassed 4-mers, GGAC and GACU respectively. The fraction of k-mer counts is weighted the ratio between their abundance at bound sites with respect to control sites (y-axis). (C) Line plot of the empirical cumulative distribution functions (ECDF) of YTHDF1 dependency on m⁶A presence in each motif at each binding site. The m⁶A dependency is computed as the ratio between peak binding scores observed at YTHDF1 binding sites when using native RNA or *in vitro* transcribed RNA as substrates. The values are log₂ transformed. Asterisks represent *p*-values (one-tailed Wilcoxon rank sum test, ****p*<0.001). '2 or more' refers to binding of a minimum of two tetramers. (D) Bar plots show significantly enriched parent Gene Ontology (GO) terms for YTHDF1-bound genes over the expressed transcriptome (FDR-adjusted *p*≤0.05). GO terms were ranked by m⁶A dependency (left) and the relative number of binding sites that were mapped to each GO term (right). The m⁶A dependency per GO term is represented by the ratio of cumulative peak binding scores per pathway between RAP and T7-RAPseq assays.

Figure S6. ncRBP have reduced number of endogenous RNA targets compared to cRBPs. Related to Figure 6

(A-B) Boxplots show number of (A) all binding sites and (B) binding sites per gene identified in RAPseq for canonical (cRBPs, dark blue) and noncanonical RBPs (ncRBPs). ncRBPs are split into two groups based on more (light blue) or fewer (grey) than 1000 binding sites. The number (n) of RBPs for each group is indicated. (C) Barplots display the number of binding sites identified in RAPseq for all 26 ncRBP. (D) Barplots demonstrate the binding site occurrences across RNA biotypes for each ncRBPs mapping to small noncoding RNAs (top, excluding long noncoding RNA), all noncoding RNAs (middle) or all biotypes (bottom). (E) Boxplots show the fold-change of enrichments of RBP binding sites (n) over HaloTag control for the three groups. Yellow boxes represent downsampling per group. (F) Boxplots display the fold-change of enrichments for ncRBPs with more than 1000 binding sites considering all (top) or downsampled (n=1000) binding sites.

Figure S7. The binding specificity of IGF2BP family disease variants is preserved. Related to Figure 7

(A) Line plots show IGF2BP paralog mRNA expression levels (in reads per kilobase per million mapped reads, RPKM) over human (left) and mouse (right) liver development. Birth is indicated by dashed lines. (B) Barplots display IGF2BP gene family expression levels in liver cancer cell lines and healthy adult liver in human (left, HepG2 and Huh7) and mouse (right, Hepa1.6 and Hepa1c1c7). (C) Boxplots show IGF2BP gene family expression levels in TCGA liver cancer cohorts. Each dot represents one patient. (D) Scatter plots (upper-right triangle) illustrate k-mer enrichment correlations between the binding specificities of the three IGF2BP paralogs and their respective variants. RBFOX2 and YTHDF1 are included as controls. Colored dots highlight the binding motifs of each RBP. Spearman rank correlations (ρ) for each comparison are reported (lower-left triangle). Axes present z-transformed k-mer enrichment values. (E) Barplot shows the number (y-axis) and type (x-axis) of differentially bound ncRNAs by the two IGF2BP2 isoforms. (F) Genome tracks demonstrate IGF2BP2 isoform A (grey) or B (orange) binding over HaloTag control (grey) to two snoRNAs and two miRNAs. Genomic locations with scale bar indicate the length of the genomic region in bases (x-axis) and normalized read density (RPM, y-axis) are shown.

References

- Aiyama, T., Orimo, T., Yokoo, H., Ohata, T., Hatanaka, K. C., Hatanaka, Y., Fukai, M., Kamiyama, T., & Taketomi, A. (2020). Adenomatous polyposis coli-binding protein end-binding 1 promotes hepatocellular carcinoma growth and metastasis. *PLOS ONE*, 15(9), e0239462. <https://doi.org/10.1371/JOURNAL.PONE.0239462>
- Askham, J. M., Vaughan, K. T., Goodson, H. V., & Morrison, E. E. (2002). Evidence That an Interaction between EB1 and p150Glued Is Required for the Formation and Maintenance of a Radial Microtubule Array Anchored at the Centrosome. *https://doi.org/10.1091/Mbc.E02-01-0061*, 13(10), 3627–3645. <https://doi.org/10.1091/MBC.E02-01-0061>
- Baltz, A. G., Munschauer, M., Schwanhäusser, B., Vasile, A., Murakawa, Y., Schueler, M., Youngs, N., Penfold-Brown, D., Drew, K., Milek, M., Wyler, E., Bonneau, R., Selbach, M., Dieterich, C., & Landthaler, M. (2012). The mRNA-Bound Proteome and Its Global Occupancy Profile on Protein-Coding Transcripts. *Molecular Cell*, 46(5), 674–690. <https://doi.org/10.1016/J.MOLCEL.2012.05.021>
- Begg, B. E., Jens, M., Wang, P. Y., Minor, C. M., & Burge, C. B. (2020). Concentration-dependent splicing is enabled by Rbfox motifs of intermediate affinity. *Nature Structural & Molecular Biology* 2020 27:10, 27(10), 901–912. <https://doi.org/10.1038/s41594-020-0475-8>
- Cardoso-Moreira, M., Halbert, J., Valloton, D., Velten, B., Chen, C., Shao, Y., Liechti, A., Ascensão, K., Rummel, C., Ovchinnikova, S., Mazin, P. V., Xenarios, I., Harshman, K., Mort, M., Cooper, D. N., Sandi, C., Soares, M. J., Ferreira, P. G., Afonso, S., ... Kaessmann, H. (2019). Gene expression across mammalian organ development. *Nature* 2019 571:7766, 571(7766), 505–509. <https://doi.org/10.1038/s41586-019-1338-5>
- Castello, A., Fischer, B., Eichelbaum, K., Horos, R., Beckmann, B. M., Strein, C., Davey, N. E., Humphreys, D. T., Preiss, T., Steinmetz, L. M., Krijgsvelde, J., & Hentze, M. W. (2012). Insights into RNA Biology from an Atlas of Mammalian mRNA-Binding Proteins. *Cell*, 149(6), 1393–1406. <https://doi.org/10.1016/J.CELL.2012.04.031>
- Castello, A., Fischer, B., Frese, C. K., Horos, R., Alleaume, A. M., Foehr, S., Curk, T., Krijgsvelde, J., & Hentze, M. W. (2016). Comprehensive Identification of RNA-Binding Domains in Human Cells. *Molecular Cell*, 63(4), 696–710. <https://doi.org/10.1016/J.MOLCEL.2016.06.029>
- Christodoulou-Vafeiadou, E., Ioakeimidis, F., Andreadou, M., Giagkas, G., Stamatakis, G., Reczko, M., Samiotaki, M., Papanastasiou, A. D., Karakasiliotis, I., & Kontoyiannis, D. L. (2018). Divergent Innate and Epithelial Functions of the RNA-Binding Protein HuR in Intestinal Inflammation. *Frontiers in Immunology*, 0(NOV), 2732. <https://doi.org/10.3389/FIMMU.2018.02732>
- Conway, A. E., Van Nostrand, E. L., Pratt, G. A., Aigner, S., Wilbert, M. L., Sundararaman, B., Freese, P., Lambert, N. J., Sathe, S., Liang, T. Y., Essex, A., Landais, S., Burge, C. B., Jones, D. L., & Yeo, G. W. (2016). Enhanced CLIP Uncovers IMP Protein-RNA Targets in Human Pluripotent Stem Cells Important for Cell Adhesion and Survival. *Cell Reports*, 15(3), 666–679. <https://doi.org/10.1016/J.CELREP.2016.03.052>
- Dominissini, D., Moshitch-Moshkovitz, S., Schwartz, S., Salmon-Divon, M., Ungar, L., Osenberg, S., Cesarkas, K., Jacob-Hirsch, J., Amariglio, N., Kupiec, M., Sorek, R., & Rechavi, G. (2012). Topology of the human and mouse m6A RNA methylomes revealed by m6A-seq. *Nature* 2012 485:7397, 485(7397), 201–206. <https://doi.org/10.1038/nature11112>
- Dominissini, D., Nachtergaele, S., Moshitch-Moshkovitz, S., Peer, E., Kol, N., Ben-Haim, M. S., Dai, Q., Di Segni, A., Salmon-Divon, M., Clark, W. C., Zheng, G., Pan, T., Solomon,

- O., Eyal, E., Hershkovitz, V., Han, D., Doré, L. C., Amariglio, N., Rechavi, G., & He, C. (2016). The dynamic N1-methyladenosine methylome in eukaryotic messenger RNA. *Nature* 2016 530:7591, 530(7591), 441–446. <https://doi.org/10.1038/nature16998>
- Galbán, S., Kuwano, Y., Pullmann, R., Martindale, J. L., Kim, H. H., Lal, A., Abdelmohsen, K., Yang, X., Dang, Y., Liu, J. O., Lewis, S. M., Holcik, M., & Gorospe, M. (2008). RNA-Binding Proteins HuR and PTB Promote the Translation of Hypoxia-Inducible Factor 1α. *Molecular and Cellular Biology*, 28(1), 93–107. <https://doi.org/10.1128/MCB.00973-07>
- Geng, G., Xu, C., Peng, N., Li, Y., Liu, J., Wu, J., Liang, J., Zhu, Y., & Shi, L. (2021). PTBP1 is necessary for dendritic cells to regulate T-cell homeostasis and antitumour immunity. *Immunology*, 163(1), 74–85. <https://doi.org/10.1111/IMM.13304>
- Good, P. J. (1995). A conserved family of elav-like genes in vertebrates. *Proceedings of the National Academy of Sciences*, 92(10), 4557–4561. <https://doi.org/10.1073/PNAS.92.10.4557>
- Hafner, M., Katsantoni, M., Köster, T., Marks, J., Mukherjee, J., Staiger, D., Ule, J., & Zavolan, M. (2021). CLIP and complementary methods. *Nature Reviews Methods Primers* 2021 1:1, 1(1), 1–23. <https://doi.org/10.1038/s43586-021-00018-1>
- Hafner, M., Landthaler, M., Burger, L., Khorshid, M., Hausser, J., Berninger, P., Rothballer, A., Ascano, M., Jungkamp, A.-C., Munschauer, M., Ulrich, A., Wardle, G. S., Dewell, S., Zavolan, M., & Tuschl, T. (2010). Transcriptome-wide Identification of RNA-Binding Protein and MicroRNA Target Sites by PAR-CLIP. *Cell*, 141(1), 129–141. <https://doi.org/10.1016/J.CELL.2010.03.009>
- Han, D., Liu, J., Chen, C., Dong, L., Liu, Y., Chang, R., Huang, X., Liu, Y., Wang, J., Dougherty, U., Bissonnette, M. B., Shen, B., Weichselbaum, R. R., Xu, M. M., & He, C. (2019). Anti-tumour immunity controlled through mRNA m6A methylation and YTHDF1 in dendritic cells. *Nature* 2019 566:7743, 566(7743), 270–274. <https://doi.org/10.1038/s41586-019-0916-x>
- Hentze, M. W., Caughman, S. W., Rouault, T. A., Barriocanal, J. G., Dancis, A., Harford, J. B., & Klausner, R. D. (1987). Identification of the iron-responsive element for the translational regulation of human ferritin mRNA. *Science*, 238(4833), 1570–1573. <https://doi.org/10.1126/SCIENCE.3685996>
- Hentze, M. W., & Preiss, T. (2010). The REM phase of gene regulation. *Trends in Biochemical Sciences*, 35(8), 423–426. <https://doi.org/10.1016/J.TIBS.2010.05.009>
- Huang, H., Weng, H., Sun, W., Qin, X., Shi, H., Wu, H., Zhao, B. S., Mesquita, A., Liu, C., Yuan, C. L., Hu, Y.-C., Hüttelmaier, S., Skibbe, J. R., Su, R., Deng, X., Dong, L., Sun, M., Li, C., Nachtergaele, S., ... Chen, J. (2018). Recognition of RNA N6-methyladenosine by IGF2BP proteins enhances mRNA stability and translation. *Nature Cell Biology* 20:3, 20(3), 285–295. <https://doi.org/10.1038/s41556-018-0045-z>
- Jolma, A., Zhang, J., Mondragón, E., Morgunova, E., Kivioja, T., Lavery, K. U., Yin, Y., Zhu, F., Bourenkov, G., Morris, Q., Hughes, T. R., Maher, L. J., & Taipale, J. (2020). Binding specificities of human RNA-binding proteins toward structured and linear RNA sequences. *Genome Research*, 30(7), 962–973. <https://doi.org/10.1101/GR.258848.119>
- Kim, D., Paggi, J. M., Park, C., Bennett, C., & Salzberg, S. L. (2019). Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. *Nature Biotechnology*, 37(8). <https://doi.org/10.1038/s41587-019-0201-4>
- König, J., Zarnack, K., Rot, G., Curk, T., Kayikci, M., Zupan, B., Turner, D. J., Luscombe, N. M., & Ule, J. (2010). iCLIP reveals the function of hnRNP particles in splicing at individual nucleotide resolution. *Nature Structural & Molecular Biology* 2010 17:7, 17(7), 909–915. <https://doi.org/10.1038/nsmb.1838>
- Kutter, C., Brown, G. D., Gonçalves, Â., Wilson, M. D., Watt, S., Brazma, A., White, R. J., &

- Odom, D. T. (2011). Pol III binding in six mammals shows conservation among amino acid isotypes despite divergence among tRNA genes. *Nature Genetics*, 43. <https://doi.org/10.1038/ng.906>
- Lambert, N., Robertson, A., Jangi, M., McGeary, S., Sharp, P. A., & Burge, C. B. (2014). RNA Bind-n-Seq: Quantitative Assessment of the Sequence and Structural Binding Specificity of RNA Binding Proteins. *Molecular Cell*, 54(5), 887–900. <https://doi.org/10.1016/J.MOLCEL.2014.04.016>
- Liao, S., Sun, H., & Xu, C. (2018). YTH Domain: A Family of N⁶-methyladenosine (m⁶A) Readers. *Genomics, Proteomics & Bioinformatics*, 16(2), 99–107. <https://doi.org/10.1016/j.gpb.2018.04.002>
- Licatalosi, D. D., Mele, A., Fak, J. J., Ule, J., Kayikci, M., Chi, S. W., Clark, T. A., Schweitzer, A. C., Blume, J. E., Wang, X., Darnell, J. C., & Darnell, R. B. (2008). HITS-CLIP yields genome-wide insights into brain alternative RNA processing. *Nature* 2008 456:7221, 456(7221), 464–469. <https://doi.org/10.1038/nature07488>
- Lin, C., & Miles, W. O. (2019). Beyond CLIP: advances and opportunities to measure RBP–RNA and RNA–RNA interactions. *Nucleic Acids Research*, 47(11), 5490. <https://doi.org/10.1093/NAR/GKZ295>
- Los, G. V., Encell, L. P., McDougall, M. G., Hartzell, D. D., Karassina, N., Zimprich, C., Wood, M. G., Learish, R., Ohana, R. F., Urh, M., Simpson, D., Mendez, J., Zimmerman, K., Otto, P., Vidugiris, G., Zhu, J., Darzins, A., Klaubert, D. H., Bulleit, R. F., & Wood, K. V. (2008). HaloTag: A Novel Protein Labeling Technology for Cell Imaging and Protein Analysis. *ACS Chemical Biology*, 3(6), 373–382. <https://doi.org/10.1021/CB800025K>
- Lu, N. T., Liu, N. M., Patel, D., Vu, J. Q., Liu, L., Kim, C. Y., Cho, P., Khachatoorian, R., Patel, N., Magyar, C. E., Ganapathy, E., Arumugaswami, V., Dasgupta, A., & French, S. W. (2018). Oncoprotein Stathmin Modulates Sensitivity to Apoptosis in Hepatocellular Carcinoma Cells During Hepatitis C Viral Replication: <https://doi.org/10.1177/1179066018785141>, 11. <https://doi.org/10.1177/1179066018785141>
- Meyer, K. D. (2019). DART-seq: an antibody-free method for global m⁶A detection. *Nature Methods* 2019 16:12, 16(12), 1275–1280. <https://doi.org/10.1038/s41592-019-0570-0>
- Orimo, T., Ojima, H., Hiraoka, N., Saito, S., Kosuge, T., Kakisaka, T., Yokoo, H., Nakanishi, K., Kamiyama, T., Todo, S., Hirohashi, S., & Kondo, T. (2008). Proteomic profiling reveals the prognostic value of adenomatous polyposis coli–end-binding protein 1 in hepatocellular carcinoma. *Hepatology*, 48(6), 1851–1863. <https://doi.org/10.1002/HEP.22552>
- 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., Thomas, G. H., Willis, A. E., & Lilley, K. S. (2019). Comprehensive identification of RNA–protein interactions in any organism using orthogonal organic phase separation (OOPS). *Nature Biotechnology* 2019 37:2, 37(2), 169–178. <https://doi.org/10.1038/s41587-018-0001-2>
- Ray, D., Kazan, H., Chan, E. T., Castillo, L. P., Chaudhry, S., Talukder, S., Blencowe, B. J., Morris, Q., & Hughes, T. R. (2009). Rapid and systematic analysis of the RNA recognition specificities of RNA-binding proteins. *Nature Biotechnology* 2009 27:7, 27(7), 667–670. <https://doi.org/10.1038/nbt.1550>
- Ray, D., Kazan, H., Cook, K. B., Weirauch, M. T., Najafabadi, H. S., Li, X., Gueroussov, S., Albu, M., Zheng, H., Yang, A., Na, H., Irimia, M., Matzat, L. H., Dale, R. K., Smith, S. A., Yarosh, C. A., Kelly, S. M., Nabet, B., Mecnas, D., ... Hughes, T. R. (2013). A compendium of RNA-binding motifs for decoding gene regulation. *Nature* 2013 499:7457, 499(7457), 172–177. <https://doi.org/10.1038/nature12311>

- Ripin, N., Boudet, J., Duszczek, M. M., Hinniger, A., Faller, M., Krepl, M., Gadi, A., Schneider, R. J., Šponer, J., Meisner-Kober, N. C., & Allain, F. H.-T. (2019). Molecular basis for AU-rich element recognition and dimerization by the HuR C-terminal RRM. *Proceedings of the National Academy of Sciences*, 116(8), 2935–2944. <https://doi.org/10.1073/PNAS.1808696116>
- Rothamel, K., Arcos, S., Kim, B., Reasoner, C., Lisy, S., Mukherjee, N., & Ascano, M. (2021). ELAVL1 primarily couples mRNA stability with the 3' UTRs of interferon-stimulated genes. *Cell Reports*, 35(8). <https://doi.org/10.1016/J.CELREP.2021.109178>
- Rudolph, K. L. M., Schmitt, B. M., Villar, D., White, R. J., Marioni, J. C., Kutter, C., & Odom, D. T. (2016). Codon-Driven Translational Efficiency Is Stable across Diverse Mammalian Cell States. *PLOS Genetics*, 12(5), e1006024. <https://doi.org/10.1371/JOURNAL.PGEN.1006024>
- Sasanuma, H., Ozawa, M., & Yoshida, N. (2019). RNA-binding protein Ptbp1 is essential for BCR-mediated antibody production. *International Immunology*, 31(3), 157–166. <https://doi.org/10.1093/INTIMM/DXY077>
- Schneider, T., Hung, L.-H., Aziz, M., Wilmen, A., Thaum, S., Wagner, J., Janowski, R., Müller, S., Schreiner, S., Friedhoff, P., Hüttelmaier, S., Niessing, D., Sattler, M., Schlundt, A., & Bindereif, A. (2019). Combinatorial recognition of clustered RNA elements by the multidomain RNA-binding protein IMP3. *Nature Communications* 2019 10:1, 10(1), 1–18. <https://doi.org/10.1038/s41467-019-09769-8>
- Shi, H., Zhang, X., Weng, Y.-L., Lu, Z., Liu, Y., Lu, Z., Li, J., Hao, P., Zhang, Y., Zhang, F., Wu, Y., Delgado, J. Y., Su, Y., Patel, M. J., Cao, X., Shen, B., Huang, X., Ming, G., Zhuang, X., ... Zhou, T. (2018). m6A facilitates hippocampus-dependent learning and memory through YTHDF1. *Nature* 2018 563:7730, 563(7730), 249–253. <https://doi.org/10.1038/s41586-018-0666-1>
- Shwetha, S., Kumar, A., Mullick, R., Vasudevan, D., Mukherjee, N., & Das, S. (2015). HuR Displaces Polypyrimidine Tract Binding Protein To Facilitate La Binding to the 3' Untranslated Region and Enhances Hepatitis C Virus Replication. *Journal of Virology*, 89(22), 11356–11371. <https://doi.org/10.1128/JVI.01714-15>
- Steinmetz, M. O., Kammerer, R. A., Jahnke, W., Goldie, K. N., Lustig, A., & Oostrum, J. van. (2000). Op18/stathmin caps a kinked protofilament-like tubulin tetramer. *The EMBO Journal*, 19(4), 572–580. <https://doi.org/10.1093/EMBOJ/19.4.572>
- Sueyoshi, T., Kawasaki, T., Kitai, Y., Ori, D., Akira, S., & Kawai, T. (2018). Hu Antigen R Regulates Antiviral Innate Immune Responses through the Stabilization of mRNA for Polo-like Kinase 2. *The Journal of Immunology*, 200(11), 3814–3824. <https://doi.org/10.4049/JIMMUNOL.1701282>
- Trendel, J., Schwarzl, T., Horos, R., Prakash, A., Bateman, A., Hentze, M. W., & Krijgsvelde, J. (2019). The Human RNA-Binding Proteome and Its Dynamics during Translational Arrest. *Cell*, 176(1), 391–403.e19. <https://doi.org/10.1016/J.CELL.2018.11.004>
- Tybl, E., Shi, F.-D., Kessler, S. M., Tierling, S., Walter, J., Bohle, R. M., Wieland, S., Zhang, J., Tan, E. M., & Kiemer, A. K. (2011). Overexpression of the IGF2-mRNA binding protein p62 in transgenic mice induces a steatotic phenotype. *Journal of Hepatology*, 54(5), 994–1001. <https://doi.org/10.1016/J.JHEP.2010.08.034>
- Ule, J., Jensen, K. B., Ruggiu, M., Mele, A., Ule, A., & Darnell, R. B. (2003). CLIP Identifies Nova-Regulated RNA Networks in the Brain. *Science*, 302(5648), 1212–1215. <https://doi.org/10.1126/SCIENCE.1090095>
- Urdaneta, E. C., Vieira-Vieira, C. H., Hick, T., Wessels, H.-H., Figini, D., Moschall, R., Medenbach, J., Ohler, U., Granneman, S., Selbach, M., & Beckmann, B. M. (2019). Purification of cross-linked RNA-protein complexes by phenol-toluol extraction. *Nature Communications* 2019 10:1, 10(1), 1–17. <https://doi.org/10.1038/s41467-019-08942-3>

- van Haren, J., Charafeddine, R. A., Ettinger, A., Wang, H., Hahn, K. M., & Wittmann, T. (2018). Local control of intracellular microtubule dynamics by EB1 photodissociation. *Nature Cell Biology* 20:3, 20(3), 252–261. <https://doi.org/10.1038/s41556-017-0028-5>
- Van Nostrand, E. L., Freese, P., Pratt, G. A., Wang, X., Wei, X., Xiao, R., Blue, S. M., Chen, J.-Y., Cody, N. A. L., Dominguez, D., Olson, S., Sundararaman, B., Zhan, L., Bazile, C., Bouvrette, L. P. B., Bergalet, J., Duff, M. O., Garcia, K. E., Gelboin-Burkhart, C., ... Yeo, G. W. (2020). A large-scale binding and functional map of human RNA-binding proteins. *Nature* 2020 583:7818, 583(7818), 711–719. <https://doi.org/10.1038/s41586-020-2077-3>
- Van Nostrand, E. L., Pratt, G. A., Shishkin, A. A., Gelboin-Burkhart, C., Fang, M. Y., Sundararaman, B., Blue, S. M., Nguyen, T. B., Surka, C., Elkins, K., Stanton, R., Rigo, F., Guttman, M., & Yeo, G. W. (2016). Robust transcriptome-wide discovery of RNA-binding protein binding sites with enhanced CLIP (eCLIP). *Nature Methods* 2016 13:6, 13(6), 508–514. <https://doi.org/10.1038/nmeth.3810>
- Wang, H., Zeng, F., Liu, Q., Liu, H., Liu, Z., Niu, L., Teng, M., Li, X., & IUCr. (2013). The structure of the ARE-binding domains of Hu antigen R (HuR) undergoes conformational changes during RNA binding. *Urn:Issn:0907-4449*, 69(3), 373–380. <https://doi.org/10.1107/S0907444912047828>
- Wang, X., Zhao, B. S., Roundtree, I. A., Lu, Z., Han, D., Ma, H., Weng, X., Chen, K., Shi, H., & He, C. (2015). N6-methyladenosine Modulates Messenger RNA Translation Efficiency. *Cell*, 161(6), 1388–1399. <https://doi.org/10.1016/J.CELL.2015.05.014>
- Wheeler, E. C., Nostrand, E. L. Van, & Yeo, G. W. (2018). Advances and challenges in the detection of transcriptome-wide protein–RNA interactions. *Wiley Interdisciplinary Reviews: RNA*, 9(1), e1436. <https://doi.org/10.1002/WRNA.1436>
- Wittmann, T., Bokoch, G. M., & Waterman-Storer, C. M. (2004). Regulation of Microtubule Destabilizing Activity of Op18/Stathmin Downstream of Rac1 *. *Journal of Biological Chemistry*, 279(7), 6196–6203. <https://doi.org/10.1074/JBC.M307261200>
- Wong, Q. W. –L., Lung, R. W. –M., Law, P. T. –Y., Lai, P. B. –S., Chan, K. Y. –Y., To, K., & Wong, N. (2008). MicroRNA-223 Is Commonly Repressed in Hepatocellular Carcinoma and Potentiates Expression of Stathmin1. *Gastroenterology*, 135(1), 257–269. <https://doi.org/10.1053/J.GASTRO.2008.04.003>
- Wu, R., Li, A., Sun, B., Sun, J.-G., Zhang, J., Zhang, T., Chen, Y., Xiao, Y., Gao, Y., Zhang, Q., Ma, J., Yang, X., Liao, Y., Lai, W.-Y., Qi, X., Wang, S., Shu, Y., Wang, H.-L., Wang, F., ... Yuan, Z. (2018). A novel m6A reader Prrc2a controls oligodendroglial specification and myelination. *Cell Research* 2018 29:1, 29(1), 23–41. <https://doi.org/10.1038/s41422-018-0113-8>
- Xu, C., Liu, K., Ahmed, H., Loppnau, P., Schapira, M., & Min, J. (2015). Structural Basis for the Discriminative Recognition of N6-Methyladenosine RNA by the Human YT521-B Homology Domain Family of Proteins *. *Journal of Biological Chemistry*, 290(41), 24902–24913. <https://doi.org/10.1074/JBC.M115.680389>
- Xu, C., Wang, X., Liu, K., Roundtree, I. A., Tempel, W., Li, Y., Lu, Z., He, C., & Min, J. (2014). Structural basis for selective binding of m6A RNA by the YTHDC1 YTH domain. *Nature Chemical Biology* 2014 10:11, 10(11), 927–929. <https://doi.org/10.1038/nchembio.1654>
- Zarnegar, B. J., Flynn, R. A., Shen, Y., Do, B. T., Chang, H. Y., & Khavari, P. A. (2016). irCLIP platform for efficient characterization of protein–RNA interactions. *Nature Methods* 2016 13:6, 13(6), 489–492. <https://doi.org/10.1038/nmeth.3840>
- Zhang, J.-Y., Chan, E. K. L., Peng, X.-X., & Tan, E. M. (1999). A Novel Cytoplasmic Protein with RNA-binding Motifs Is an Autoantigen in Human Hepatocellular Carcinoma.

Journal of Experimental Medicine, 189(7), 1101–1110.

<https://doi.org/10.1084/JEM.189.7.1101>

Zhang, L.-S., Liu, C., Ma, H., Dai, Q., Sun, H.-L., Luo, G., Zhang, Z., Zhang, L., Hu, L., Dong, X., & He, C. (2019). Transcriptome-wide Mapping of Internal N7-Methylguanosine Methylome in Mammalian mRNA. *Molecular Cell*, 74(6), 1304-1316.e8. <https://doi.org/10.1016/J.MOLCEL.2019.03.036>

Zhang, Y., Liu, T., Meyer, C. A., Eeckhoutte, J., Johnson, D. S., Bernstein, B. E., Nussbaum, C., Myers, R. M., Brown, M., Li, W., & Liu, X. S. (2008). Model-based Analysis of ChIP-Seq (MACS). *Genome Biology*, 9(9). <https://doi.org/10.1186/gb-2008-9-9-r137>