

MOLECULAR BIOLOGY

Nuclear 2'-O-methylation regulates RNA splicing through its binding protein FUBP1

Boyang Gao^{1,2}, Bochen Jiang^{2,3}, Zhongyu Zou^{2,3}, Bei Liu^{2,3}, Weijin Liu⁴, Li Chen³, Lisheng Zhang^{4,5}, Chuan He^{2,3*}

2'-O-methylation (N_m) is an abundant RNA modification exists on different mammalian RNA species. However, potential N_m recognition by proteins has not been extensively explored. Here, we used RNA affinity purification, followed by mass spectrometry to identify N_m -binding proteins. The N_m -binding protein candidates exhibit enriched binding at known N_m sites. Some candidates display nuclear localization and functions. We focused on the splicing factor FUBP1. Electrophoretic mobility shift assay validated preference of FUBP1 to N_m -modified RNA. As FUBP1 predominantly binds intronic regions, we profiled N_m sites in chromatin-associated RNA (caRNA) and found N_m enrichment within introns. Depletion of N_m led to skipped exons, suggesting N_m -dependent splicing regulation. The caRNA N_m sites overlap with FUBP1-binding sites, and N_m depletion reduced FUBP1 occupancy on modified regions. Furthermore, *FUBP1* depletion induced exon skipping in N_m -modified genes, supporting its role in mediating N_m -dependent splicing regulation. Overall, our findings identify FUBP1 as an N_m -binding protein and uncover previously unrecognized nuclear functions for RNA N_m modification.

INTRODUCTION

Throughout the life cycles of RNA, chemical modifications play critical roles in regulating RNA processing, metabolism, and function (1, 2). These modifications can alter the physical properties of the RNA molecule (3, 4) or recruit specific binding proteins (readers) to modulate RNA function and subsequent cellular pathways (5–9). While changes in physical properties primarily influence RNA structure-dependent regulation, the recruitment of reader proteins can affect diverse downstream processes, including splicing (10), degradation (5), translation (6, 9), and RNA transport to specialized cellular locations (11–13). In addition, recruitment of binding proteins may alter the surrounding state of the modified RNA; examples include chromatin state regulation by N^6 -methyladenosine (m^6A) methylation of chromatin-associated RNA (caRNA) (7, 8, 14).

2'-O-methylation (N_m) is one of the most abundant modifications (15). It can be found in almost all RNA species, including ribosomal RNA (rRNA), tRNA, small nuclear RNA, small nucleolar RNA (snoRNA), microRNA, and mRNA. It has been shown to regulate ribosome biogenesis (16, 17), gene expression (18, 19), innate immune sensing (20, 21), and cell fate decisions (22). N_m entails methylation at the 2'-OH position of the ribose on the RNA backbone. Consequently, it can occur on any of the four ribonucleotide residues, namely, 2'-O-methyladenosine (A_m), 2'-O-methylguanosine (G_m), 2'-O-methylcytidine (C_m), and 2'-O-methyluridine (U_m). The methylation further stabilizes the ribose 3'-endo conformation, favoring the A-type RNA helix and restricting strand flexibility (23–25). Consequently, N_m installation alters physical properties of the modified RNA. This effect orchestrates N_m -dependent cellular functions of different RNA species, such as stabilizing ribosome structure

through modified rRNA (26) and controlling translational efficiency through internal mRNA N_m modification (18). However, whether N_m could be recognized by binding proteins (readers) remains largely unexplored, particularly within the internal regions of mRNA or premature mRNA. While previous studies have identified over a thousand internal N_m sites (27), the corresponding function remains elusive. The processing, metabolism, or function of N_m -modified RNA could be unveiled by the identification of N_m -binding proteins. A previous study reported a functional relationship between cleavage and polyadenylation specificity factor subunit 7 (CPSF7) binding and N_m -mediated alternative polyadenylation, suggesting CPSF7 as a potential N_m -binding protein (28). However, more comprehensive identification of N_m -binding proteins is still lacking.

In this work, we conducted RNA affinity purification, followed by mass spectrometry (MS) to identify N_m -binding protein candidates, among which we focused on the splicing factor, far upstream element binding protein 1 (FUBP1). Up-regulation of *FUBP1* has been suggested to promote proliferation of multiple types of cancers (29, 30). Initially characterized as a transcription factor modulating MYC expression, FUBP1 is now recognized as an RNA binding protein (RBP) involved in pre-mRNA splicing (31, 32). Previous studies have indicated that FUBP1 stabilizes splicing machineries at 3' splice sites, including U2AF2 and SF1. In addition, it interacts with components of U1 small nuclear ribonucleoproteins, potentially facilitating splice sites pairing in long introns. Understanding the binding specificity of FUBP1 could elucidate the mechanism underlying alternative splicing (AS) events that could be affected by N_m modification through FUBP1.

Using electrophoretic mobility shift assay (EMSA), we confirmed the binding preference of FUBP1 to N_m -modified oligos, supporting its role as an N_m -binding protein. N_m -mut-seq (27) analysis of caRNA in HepG2 cells identified 5575 N_m sites, with more than half of the intragenic N_m sites localized in introns. Disruption of N_m installation led to altered splicing patterns, especially increased exon skipping. These N_m sites were bound by FUBP1 in a manner responsive to N_m depletion. Last, FUBP1 preferentially regulates exon skipping at N_m -modified regions. Together, our study identifies FUBP1 as

¹Department of Molecular Genetics and Cell Biology, University of Chicago, Chicago, IL 60637, USA. ²Howard Hughes Medical Institute, The University of Chicago, Chicago, IL 60637, USA. ³Department of Chemistry, The University of Chicago, Chicago, IL 60637, USA. ⁴Division of Life Science, The Hong Kong University of Science and Technology (HKUST), Kowloon, Hong Kong SAR, China. ⁵Department of Chemistry, The Hong Kong University of Science and Technology (HKUST), Kowloon, Hong Kong SAR, China.

*Corresponding author. Email: chuanhe@uchicago.edu

one of the first examples of N_m -binding proteins and highlights a previously unrecognized role of N_m in splicing regulation.

RESULTS

RNA affinity purification followed by MS identified N_m -binding protein candidates

To identify proteins that may preferentially recognize internal RNA N_m modification, we designed biotinylated RNA probes based on published internal N_m sites in HepG2 (27). Given that N_m can occur on four distinct types of ribonucleotides (A_m , G_m , C_m , and U_m), probes with a single type of N_m may not fully capture the binding proteins recognizing four different N_m modifications. Considering the predominant presence of G_m and A_m in mRNA N_m modifications (27), we designed two probes: one with G_m modification (G_m -1) and the other with A_m modification (A_m -2) (Fig. 1A). Each probe was accompanied by a corresponding control oligo lacking the N_m modification, denoted as Ctrl-1 and Ctrl-2. These probes were selected on the basis of two reported N_m sites, with relatively high mutation ratio and responsiveness to the knockdown (KD) of their

methyltransferase fibrillarin (FBL; Fig. 1) (27). The G_m -1 sequence is in the 3' untranslated region (3'UTR) of *RPL13*, while the A_m -2 sequence situates in the coding sequence (CDS) of *UQCRC2*.

We performed affinity purification with G_m -1 and A_m -2 using HepG2 cell lysate, respectively, and used MS for protein identification of the pull-down fraction. Excess probes were used in these experiments (fig. S1A). Peptide abundance was normalized to the total peptide abundance, and fold changes of N_m -modified versus control groups were computed. With a cutoff of a $\log_2$ fold change of >1.58 , an adjusted $P < 0.05$, and a peptide number ≥ 4 , we identified 37 enriched proteins in the G_m -1 group and 33 enriched proteins in A_m -2 group (Fig. 1B and fig. S1B). These enriched proteins are tentative binding proteins for G_m and A_m modifications. The identified protein candidates include mRNA binding proteins and proteins involved in RNA-related metabolic pathways, based on gene ontology (GO) analysis (Fig. 1C), validating the reliability of our assay.

In the candidate protein list, we observed proteins present in both G_m - and A_m -enriched groups, indicating them as potential N_m binders without base specificity. One example is FUBP1; another homolog within the same protein family, FUBP3, is also weakly enriched by

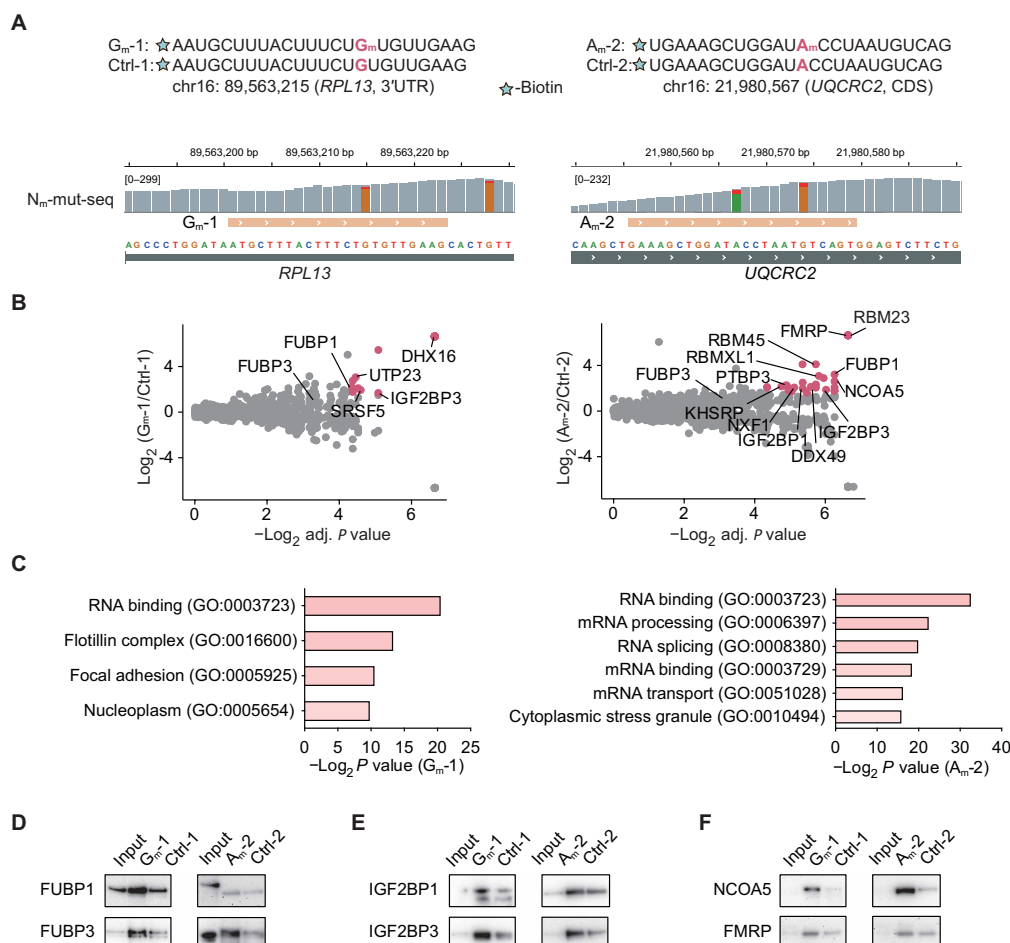


Fig. 1. RNA affinity purification followed by proteomics identified N_m -binding protein candidates. (A) Design of G_m -1 and A_m -2 probes based on known N_m sites in HepG2 mRNA. Integrative genomics viewer (IGV) tracks of published N_m -mut-seq showed read depth of mutated (red) and unmutated reads (G in orange and A in green). (B) Proteins that may preferentially bind to RNA G_m -1 (left) and A_m -2 (right) probes over control unmodified probes. Tentative N_m -binding proteins (red) have a $\log_2$ fold change of >1.58 , an adjusted $P < 0.05$, and a peptide number of ≥ 4 . (C) Gene ontology (GO) analysis of tentative N_m -binding proteins. (D to F) Immunoblotting of N_m -binding protein candidate enrichment by oligo pull-down. The ratio of the corresponding lysate amount of input versus pull-down samples are 1:100.

both oligo pairs. The structural similarity within the protein family suggests that both two proteins may preferentially bind to N_m . To validate this speculation, we detected the protein enrichment by immunoblotting after RNA affinity purification using the two pairs of probes (Fig. 1A). As expected, both FUBP1 and FUBP3 were enriched by the two N_m probes compared to their respective controls (Fig. 1D). This observation indicates that FUBP1 and FUBP3 are tentative candidates recognizing both G_m and A_m modifications on RNA.

In addition, insulin-like growth factor 2 mRNA binding protein 1 (IGF2BP1) and IGF2BP3 emerged as candidate binders of G_m and A_m . IGF2BPs were initially identified as m^6A -binding proteins that stabilize target mRNAs and facilitates their storage under stress conditions (33). Our previous work also revealed their role in recognizing 7-methylguanosine (m^7G) (34). While IGF2BP1 primarily binds m^6A and stabilizes associated mRNAs, IGF2BP3 preferentially recognizes m^7G and promotes mRNA degradation. Immunoblotting following affinity purification of both probe pairs confirmed the preferential binding of IGF2BP1 and IGF2BP3 to N_m -modified RNA (Fig. 1E). These findings may suggest a complex mechanism of RNA modification recognition by the IGF2BP proteins.

In addition to the FUBP and IGF2BP families, we also validated enrichment of fragile X messenger ribonucleoprotein 1 (FMRP) and nuclear receptor coactivator 5 (NCOA5). FMRP and NCOA5 were enriched by both G_m -1/Ctrl-1 and A_m -2/Ctrl-2 probe pairs through affinity purification, followed by immunoblotting (Fig. 1F). The RBP

FMRP has previously been shown to colocalize with internal N_m sites (28). Our validation by RNA affinity purification further supports its tentative role as an N_m -binding protein.

Notably, NCOA5 is a transcriptional coactivator that interacts with estrogen receptors (35). The enrichment of NCOA5 by RNA affinity purification suggests it as another example of a transcription factor that binds to and is potentially regulated by RNA. We analyzed a published RNA binding region identification dataset (36) and observed enriched NCOA5 peptides, indicating its involvement in RNA binding as an RBP (fig. S1C). Its preferential binding to N_m may suggest a possible role for N_m in transcriptional regulation. Overall, we validated six proteins with enriched binding to N_m -modified probes: FUBP1, FUBP3, IGF2BP1, IGF2BP3, FMRP, and NCOA5.

N_m -binding protein candidates are enriched at internal mRNA N_m sites

To provide further cellular evidence of the association between binding of these proteins and N_m modifications, we investigated their RNA binding profiles using publicly available eCLIP datasets from the ENCODE project (37, 38). Among the six identified tentative N_m binders, FUBP3, IGF2BP1, and IGF2BP3 have eCLIP data generated in HepG2 cells. Metagene plots generated from these datasets revealed enriched binding of all four proteins around the reported confident N_m sites (Fig. 2, A to C) (27), supporting the association of binding by these proteins with N_m modification. This

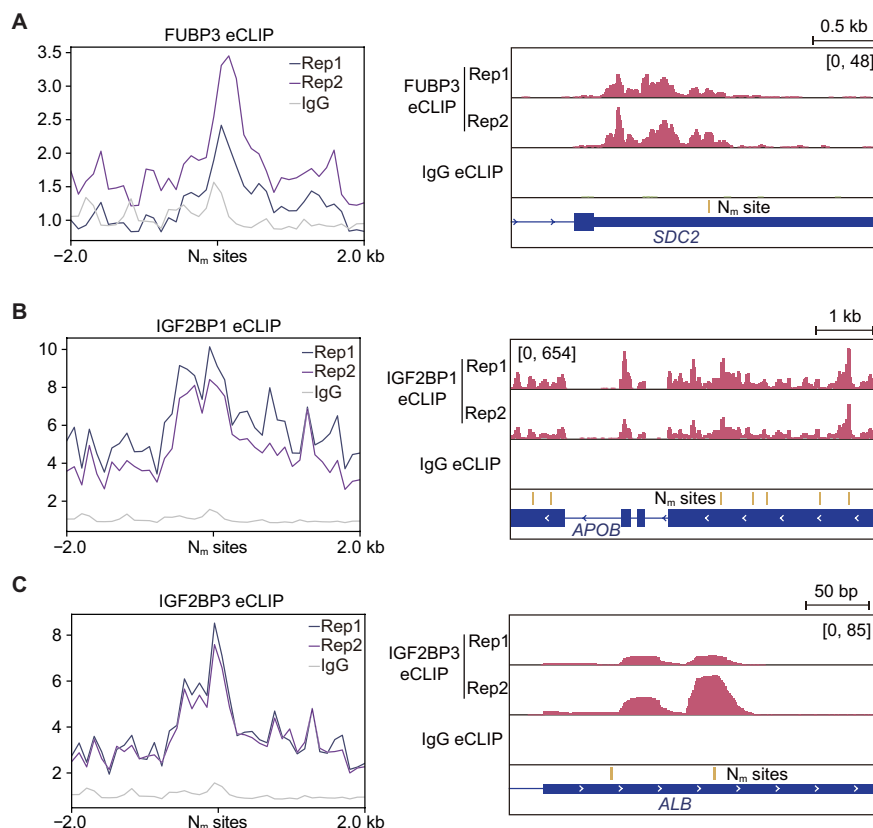


Fig. 2. RNA binding sites of N_m -binding protein candidates were enriched at N_m sites. (A to C) Metagene plots of eCLIP signals at HepG2 mRNA N_m sites (left) and representative IGV tracks (right) from binding sites of N_m binding protein candidates FUBP3 (A), IGF2BP1 (B), and IGF2BP3 (C). Ranges of the signal intensity are annotated in the brackets.

is consistent with previous reports, where IGF2BP1 and IGF2BP3 have also been suggested for their colocalization with internal N_m sites in other cell lines (28). In vitro validation of the direct binding preference is still lacking for these three proteins, so their roles as N_m-binding proteins require further validations.

We examined the sequence contexts of the three proteins at their N_m-associated binding sites. The enriched N_m motifs from the overlapped binding sites closely resembled their own RNA binding motifs (fig. S2, A to C), suggesting that the binding selectivity of these N_m-binding protein candidates are determined by their canonical sequence contexts but may be further enhanced by the N_m modification. To further investigate how RNA binding preferences affect potential N_m site selectivity, we analyzed the distribution of their binding targets and associated N_m sites across mRNA regions. While N_m sites were generally enriched in the CDS compared to the 3'UTR and 5'UTR, N_m-binding protein with 3'UTR preferences, FUBP3, predominantly bound to N_m sites within the 3'UTR, highlighting its preferences to its canonical binding sites (fig. S2D). Conversely, IGF2BP1 and IGF2BP3 exhibited a stronger enrichment of bound N_m sites in the CDS than their overall binding sites (fig. S2D), suggesting a potentially important contribution of selective N_m installation to RNA binding by these RBPs.

The reported function of IGF2BP1 in stabilizing mRNA aligns with the general effect of N_m on mRNA levels as previously described (19, 27, 28). Accordingly, we analyzed RNA level changes using published KD datasets for *IGF2BP1* in HepG2 cells (33). N_m-modified mRNA transcripts showed a more notable decrease in

expression following *IGF2BP1* KD compared to unmodified transcripts (fig. S2E), supporting the association between IGF2BP1 binding and N_m modifications. Further investigation into IGF2BP1 binding to methylated RNA is needed to elucidate the interplay among m⁶A, m⁷G, and N_m modifications.

FUBP1 preferentially binds internal mRNA N_m sites

The enrichment of nuclear proteins in the N_m binder candidate list suggests intriguing functions of N_m in the cell nucleus. We chose FUBP1 as an example for more detailed investigations, as FUBP1 is a well-characterized splicing factor (29, 32). Its preference for N_m has been validated using the two probe pairs in both proteomics and immunoblotting analyses. Nevertheless, the RNA affinity purification approach leaves a possibility of indirect FUBP1 binding mediated by an N_m-recognizing partner protein. To confirm the direct interaction between FUBP1 and the N_m modification, we recombinantly expressed and purified FUBP1 with a C-terminal strep-tag from the Expi293 expression system (fig. S3A). We then assessed the binding affinity of the purified FUBP1 toward two probe pairs by EMSA. We observed preferential binding of FUBP1 to both N_m-modified RNA probes compared to their respective controls (Fig. 3A). As a sugar 2'-OH modification, N_m recognition by FUBP1 could be more subtle, when compared with other well-recognized RNA modifications such as m⁶A by the reader YTH N6-methyladenosine RNA binding proteins (YTHDF proteins) (5). Our EMSA results support that the preferential binding of FUBP1 to N_m is mediated by a direct interaction. Future structural characterization may reveal details of this recognition.

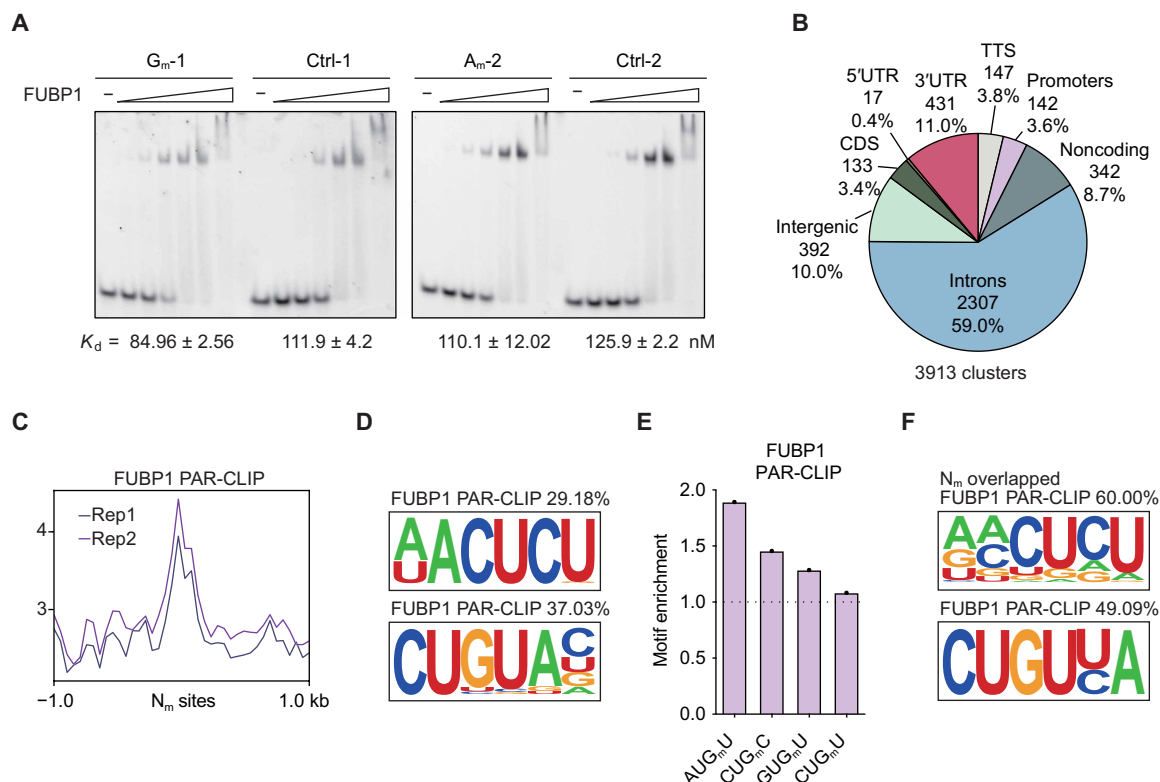


Fig. 3. Splicing regulator FUBP1 prefers N_m-modified RNA. (A) EMSA of FUBP1 binding toward G_m-1 and A_m-2 probes versus their controls. Probe concentration, 20 nM; protein concentration, starting from 400 nM with twofold dilution. (B) Distribution of FUBP1 PAR-CLIP binding clusters across transcript elements. TTS, transcription termination sites. (C) Metagene plot of FUBP1 PAR-CLIP signals at HepG2 mRNA N_m sites. (D) FUBP1-binding motifs identified in PAR-CLIP. (E) Enrichment of N_m motifs at the FUBP1-binding sites. (F) FUBP1-binding motifs identified in PAR-CLIP clusters overlapping with mRNA N_m sites.

In addition to biochemical evidence, we examined whether FUBP1 binding occurs at endogenous N_m sites within various RNA sequence contexts. We conducted photoactivatable ribonucleoside-enhanced cross-linking and immunoprecipitation (PAR-CLIP) of FUBP1 in HepG2 cells. Genome-wide mutation analysis revealed a T-to-C mutation ratio of >55% (fig. S3B). We identified 3913 FUBP1-binding sites (fig. S3C), with ~59% of these located within intronic regions (Fig. 3B), consistent with the reported intron preference of FUBP1 binding (32). With the published mRNA N_m sites at base resolution, we first investigated FUBP1 binding at reported mature mRNA N_m sites. A metagene plot revealed a clear enrichment of PAR-CLIP reads at the confident N_m sites (Fig. 3C). FUBP1-binding motifs were found to be U-rich (Fig. 3D), consistent with the U-rich motifs associated with mRNA N_m . Notably, N_m motifs enriched at FUBP1-binding sites closely resembled the RNA binding motifs shared across all its binding sites (Fig. 3, E and F). These results indicate an intrinsic preference

of FUBP1 for N_m -modified regions, further supporting its role as an N_m -binding protein.

Abundant N_m modifications on caRNA regulate splicing events

The preference of splice factor FUBP1 for N_m modifications indicates an unrecognized intron-dependent function of N_m . To explore whether N_m modifications are also present in intronic regions, we measured the overall levels of N_m in rRNA-depleted HepG2 caRNA by ultrahigh-performance liquid chromatography coupled with triple quadrupole MS (UHPLC-QQQ-MS/MS), with quantitative polymerase chain reaction (qPCR) validation of the rRNA contamination comparable to or lower than mRNA (fig. S4A). The intensity ratios of A_m/A and C_m/C are notably higher in caRNA compared to mRNA (Fig. 4A). The G_m/G ratio in caRNA is about one-third of that in mRNA; however, the considerably longer length of introns

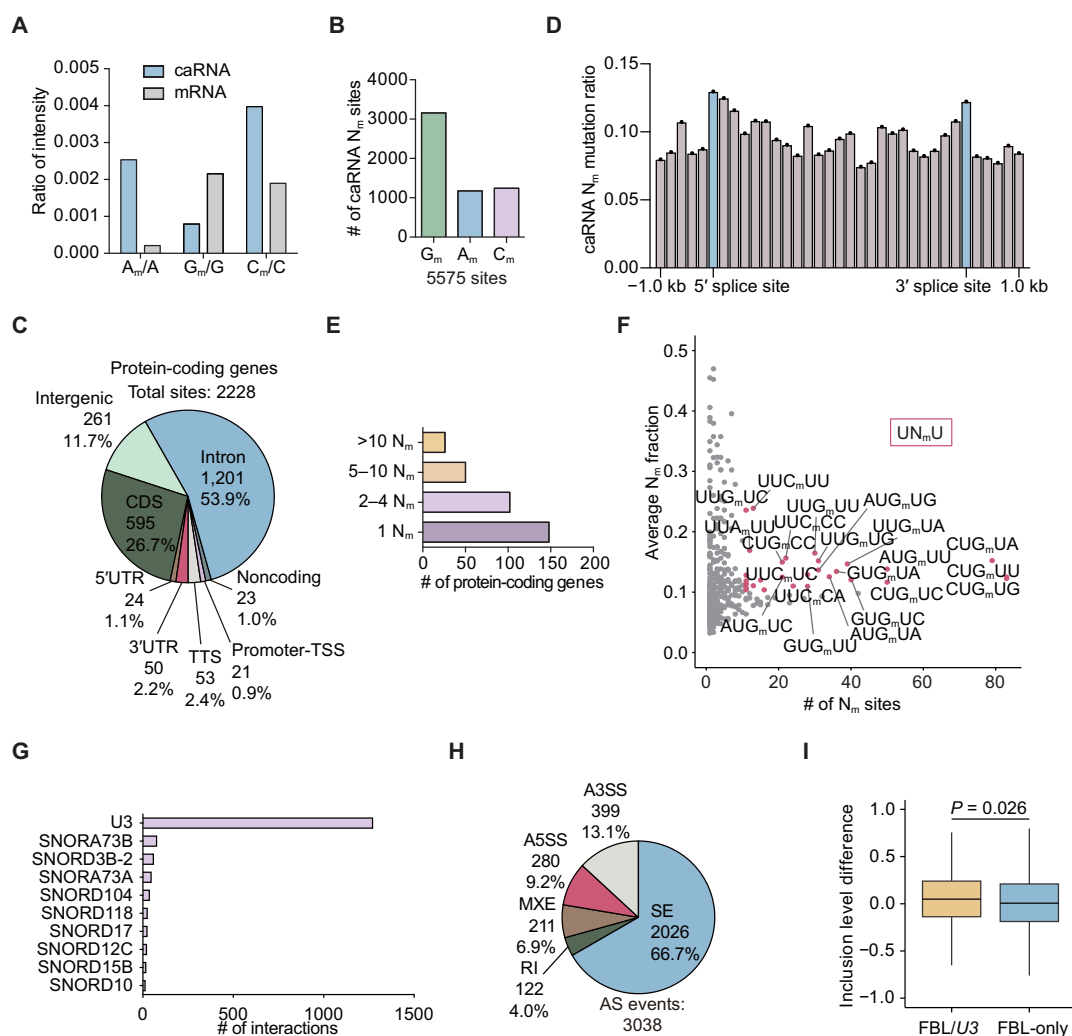


Fig. 4. N_m modifications on HepG2 caRNA affect splicing events. (A) UHPLC-QQQ-MS/MS of N_m abundances in HepG2 mRNA and caRNA. (B) Composition of G_m , A_m , and C_m among the 5575 caRNA N_m sites. (C) Transcript element distribution of 2228 caRNA N_m sites in protein-coding genes. (D) Averaged caRNA N_m mutation ratio across intronic regions in a 200-bp binning. Intronic sequences at the boundary of splice sites were marked in blue. (E) Profile of modified protein-coding gene transcripts with the corresponding numbers of N_m sites. (F) Number and average mutation ratio of all five-base N_m motifs. UN_mU was enriched in more frequent and highly modified motifs. (G) Number of interactions with various snoRNA detected by snoKARR-seq at all caRNA N_m sites. (H) Profile of FBL-dependent AS events, with a significance threshold false discovery rate (FDR) of <0.1. (I) Inclusion level differences of FBL/U3 or FBL-only SE events.

likely compensates, resulting in an overall more G_m sites in introns than those in exons. Our findings suggest enrichment of N_m modifications in caRNA.

To examine caRNA N_m distribution profiles, we performed N_m -mut-seq on HepG2 caRNA. N_m -mut-seq uses an imbalanced deoxy-nucleotide triphosphate (dNTP) supply in reverse transcription (RT) reactions to induce mutations of G_m , A_m , and C_m into T (27). During analysis, the treated samples could be controlled by RT reactions using balanced dNTP supply (input), as well as a set of spike-in that does not harbor N_m modifications but undergoes RT with imbalanced dNTP (background). We analyzed the mutations of caRNA N_m -mut-seq by JACUSA (39) and designated sites as N_m -modified when all three replicates showed a sequencing depth of ≥ 10 , a mutated read depth of ≥ 3 , mutation ratios of >3 -fold of input mutation ratios, and >1.5 -fold of background mutation ratios. On the basis of these criteria, we identified 5575 N_m sites (fig. S4B), far more than the reported 1051 N_m sites or 494 confident N_m sites found in HepG2 mRNA (27).

Similar to mRNA, G_m sites are the most abundantly modified N_m sites in caRNA (Fig. 4B), despite the relatively higher abundances of A_m and C_m detected by UHPLC-QQQ-MS/MS (Fig. 4A). The numbers of confident A_m and C_m sites in caRNA are comparable. Most of caRNA N_m sites are located within protein-coding genes (fig. S4C), with $\sim 53.9\%$ of them residing in introns (Fig. 4C). This is consistent with our hypothesis of intron-dependent function of N_m . These intronic N_m sites show modest enrichment near 5' and 3' splice sites (Fig. 4D), suggesting a potential role of N_m in splicing regulation. Our previous work has shown that internal N_m sites on mRNA often appear densely clustered along a stretch of RNA (27). We hypothesized that this clustering may amplify the relatively weak binding preferences of N_m -binding proteins toward modified RNA. We observed that 2228 caRNA N_m sites on protein-coding genes were distributed across only 323 genes, with more than half of these genes harboring multiple N_m sites (Fig. 4E). In addition to protein-coding genes, several long noncoding RNAs (lncRNAs) also accumulate hundreds of N_m sites (fig. S4D), suggesting a previously unrecognized layer of N_m -dependent regulation in their respective functions.

N_m installation in HepG2 caRNA is enriched in the UN_mU sequence context (Fig. 4F), similar to the motifs observed in HepG2 mRNA (27). The consistency should be expected, as caRNA modifications in intronic regions and noncoding RNA (ncRNA) are generally installed by the same methyltransferases responsible for mRNA modification, despite their different fates after RNA processing and nuclear export (7). The observed motifs also overlap with U-rich elements recognized by key splicing factors, such as the 5' splice sites GU, the branch sites, the UGCAUG hexanucleotides (40), the polypyrimidine tract (41), etc. This further indicates a role of N_m in splicing regulation. Notably, the U-rich motif is consistent with the known RNA binding preference of FUBP1 (32), further suggesting its role as an N_m -binding protein that recognizes intronic N_m to potentially modulate splicing.

To understand the regulation of N_m installation on caRNA, we analyzed snoRNA interactions with N_m -modified regions using the published snoRNA enriched and kethoxal assisted RNA-RNA interaction sequencing (snoKARR-seq) dataset (42). Most of snoRNA interaction with caRNA N_m sites were mediated through *U3* (Fig. 4G and fig. S4E), highlighting the critical role of *U3* in directing N_m installation on caRNA. Our finding suggests that *U3* KD may serve as an effective strategy to perturb caRNA N_m installation in addition to depletion of the N_m methyltransferase *FBL*.

The preference of N_m binding by the splicing factor FUBP1 may suggest the potential role of N_m in splicing regulation. It has been reported that ectopic overexpression of snoRNA that installs N_m in pre-mRNA branch sites could manipulate splicing of certain RNA transcripts (43, 44). To investigate the role of caRNA N_m in splicing regulation, we depleted *FBL* or *U3* in HepG2 cells (fig. S4F). We examined five types of AS events: alternative 3' splice sites (A3SSs), alternative 5' splice sites (A5SSs), mutually exclusive exons (MXEs), retained introns (RIs), and skipped exons (SEs). AS events dependent on *FBL* or *U3* exhibited similar distributions, with SE events being the most prevalent (Fig. 4H and fig. S4G). In both *FBL* depletion and *U3* KD, more SE events were up-regulated than down-regulated, while other AS types showed nonsignificant or inconsistent changes (fig. S4H). This indicates that N_m installation may play a role in preventing exon skipping. To further examine this, we compared SE events that consistently occurred following both *FBL* and *U3* KD (*FBL/U3*) with those that occurred only or inconsistently after *FBL* KD (*FBL-only*) or *U3* KD (*U3-only*). The *FBL/U3* SE events exhibited a higher degree of up-regulation compared to *FBL-only* and *U3-only* groups (Fig. 4I and fig. S4I), further supporting a role for N_m in suppressing SE events. Overall, our findings reveal abundant N_m modifications in caRNA, with intronic installation affecting splicing regulation.

FUBP1 mediates N_m -dependent splicing regulation

Our N_m -mut-seq analysis of HepG2 caRNA revealed a substantial presence of N_m modifications within intronic regions, potentially involved in splicing regulation. Given that FUBP1 is a known splicing factor implicated in splice site recognition (32), we hypothesized that it may mediate N_m -dependent splicing regulation. To investigate this, we first validated the association of FUBP1 with caRNA N_m sites. Of the 5575 caRNA N_m sites, 2030 were bound by FUBP1, as determined by PAR-CLIP analysis (Fig. 5A). Correspondingly, FUBP1 PAR-CLIP signals were enriched around caRNA N_m sites (Fig. 5B). Although FUBP1 primarily targets protein-coding genes (fig. S3C), it also binds to the heavily modified lncRNAs *MALAT1* and *NEAT1* in multiple clusters (fig. S5A). Together, these findings demonstrate that FUBP1 preferentially binds to N_m sites in both caRNA and mRNA, consolidating its role as an N_m -binding protein.

If FUBP1 acts as a binding protein for N_m , then its binding should be affected by the depletion of N_m . To test this, we conducted *FBL* KD. As PAR-CLIP provides mutation-based identification with less reliable measurement of differential peaks, we performed cross-linking immunoprecipitation sequencing (CLIP-seq) of FUBP1 following *FBL* depletion. Consistent with our PAR-CLIP results, FUBP1 CLIP-seq peaks were predominantly located within intronic regions and protein-coding genes (fig. S5, B and C). Differential peak analysis revealed a decrease in FUBP1 peak signals overlapping with N_m sites, while the overall peak intensity remained largely unchanged (Fig. 5C). Similarly, gene-based differential binding measured by FUBP1 RNA immunoprecipitation sequencing (RIP-seq) also demonstrated a significant decrease in N_m -modified genes compared to unmodified ones after *FBL* KD (Fig. 5D). These findings collectively confirmed that FUBP1 binding to N_m -modified RNA is affected by the presence of N_m modification.

With FUBP1 CLIP-seq identifying valid N_m -binding sites, we further designed N_m -modified oligos based on FUBP1 target sequences responsive to N_m depletion. EMSA revealed enhanced binding of FUBP1 to N_m -modified oligos compared to their controls (Fig. 5E and fig. S5D), consolidating FUBP1 as an N_m -binding protein.

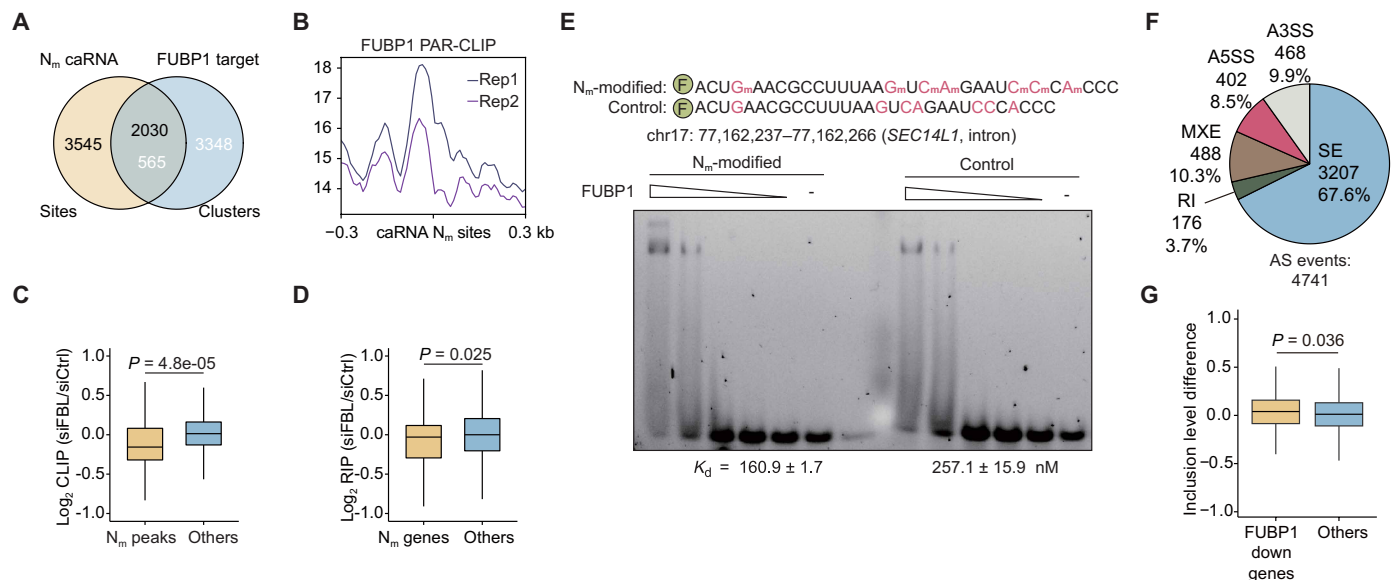


Fig. 5. FUBP1 binds to N_m -modified caRNA to affect splicing. (A) Overlap between caRNA N_m sites and FUBP1 PAR-CLIP clusters extending 200 bp on each side. (B) Metagen plot of FUBP1 PAR-CLIP signals at ± 0.3 kb around HepG2 caRNA N_m sites. (C) FUBP1 CLIP signal changes at peaks overlapping with N_m sites (N_m peaks) versus others after *FBL* KD. (D) FUBP1 RIP signal changes at gene transcripts with N_m modifications (N_m genes) versus others after *FBL* KD. (E) EMSA of FUBP1 binding toward N_m -modified or control probes originated from *SEC14L1* intron. Probe concentration, 10 nM; protein concentration, starting from 400 nM with twofold dilution. (F) Profile of the FUBP1-dependent AS events. (G) Inclusion level differences of SE events after FUBP1 depletion in genes with down-regulated FUBP1 CLIP-seq peaks after siFBL (FUBP1 down genes) versus other genes (others).

Notably, both oligos harboring multiple N_m sites showed more pronounced differences in binding affinity compared to single-modified probes (Fig. 3A), suggesting that densely clustered N_m may facilitate the recruitment of N_m -binding proteins.

After confirming FUBP1 as a caRNA N_m -binding protein, we aimed to explain the N_m -dependent splicing regulation by FUBP1. FUBP1 facilitates bridging of 5' and 3' splice sites (32), and its depletion primarily induces exon skipping (29). Consistent with this, we knocked down *FUBP1* in HepG2 cells (fig. S5E) and observed that SE was the main AS event (Fig. 5F). This aligns with the N_m -dependent splicing regulation observed after *FBL* and *U3* KD (Fig. 4H and fig. S4G). Among the SE events, up-regulation occurred more frequently than down-regulation (fig. S5F), consistent with the biased up-regulation induced by *FBL* or *U3* depletion (fig. S4H). To further elucidate the relationship between N_m modification, FUBP1 binding, and splicing changes, we examined SE events in genes with caRNA N_m modifications. We found that N_m -modified genes showed greater up-regulation of SE events after FUBP1 depletion than nonmodified genes, supporting that FUBP1 mediates N_m -dependent inhibition of exon skipping. (fig. S5, G and H). Similarly, genes with reduced FUBP1 binding following *FBL* depletion exhibited higher SE up-regulation after *FUBP1* KD (Fig. 5G), further reinforcing the causal connection between N_m modification, FUBP1 binding, and splicing changes.

DISCUSSION

N_m modifications affect the physical properties of mRNA, exerting crucial regulation on the translation efficiency of modified genes (18). However, the “reader” proteins of N_m modifications on protein-coding genes remain elusive. Characterization of N_m -binding proteins could reveal N_m -dependent function. In this study, we used

RNA affinity purification, followed by LC-MS/MS to identify N_m -binding proteins. Immunoblotting validated the enrichment of FUBP1, FUBP3, IGF2BP1, IGF2BP3, FMRP, and NCOA5. This is consistent with previous reports that IGF2BP1, IGF2BP3, and FMRP showed colocalized RNA binding with N_m modifications (28). Published eCLIP datasets of FUBP3, IGF2BP1, and IGF2BP3 further supported their enrichment around endogenous mRNA N_m sites. While functions of IGF2BP proteins align with the effect of mRNA N_m on RNA expression levels, the identified FUBP family proteins and NCOA5 suggest unrecognized realms of nuclear N_m regulation. However, without quantitative biochemical validation, the enrichment of these proteins could be indirect, and their roles as N_m -binding proteins need further validations.

We focused on FUBP1 and validated its role as an N_m -binding protein. EMSA confirmed the direct interaction between FUBP1 and N_m -modified RNA, while PAR-CLIP confirmed enrichment of FUBP1 around internal mRNA N_m sites. To further examine nuclear functions of N_m mediated through FUBP1, we profiled N_m modification sites on caRNA by N_m -mut-seq. We identified 5575 caRNA N_m sites, with more than half of caRNA N_m sites in protein-coding genes localized in introns. We found that intronic N_m displayed enrichment around 5' and 3' splice sites, suggesting N_m -mediated splicing regulation. Previous snoKARR-seq data revealed that *U3* is the predominant snoRNA responsible for N_m installation on caRNA. Depletion of *U3*, as well as methyltransferase *FBL*, led to up-regulation of SE events, supporting a role of N_m in splicing regulation. Prior studies have shown that ectopic installation of N_m at splicing branch sites by snoRNA can alter the splicing of certain transcripts (43, 44). Our findings support the N_m -mediated splicing regulation, pointing to the potential of snoRNA mimics as tools for targeted splicing modulation.

With the profiled N_m sites on caRNA, we validated the overlap of FUBP1 binding with caRNA N_m sites. To further explore the causal relationship, we disrupted N_m by *FBL* depletion and detected impaired FUBP1 binding at N_m -modified regions by CLIP-seq and RIP-seq. This further supports FUBP1 as the caRNA N_m -binding protein. *FUBP1* depletion up-regulates SE events preferentially in N_m -modified genes and genes with down-regulated FUBP1 binding after *FBL* depletion, confirming its role in mediated N_m -dependent splicing regulation. Overall, our findings identify FUBP1 as a caRNA N_m -binding protein and uncover a new function of N_m in splicing modulation through FUBP1. Nevertheless, it is likely that FUBP1 also contributes to other regulatory processes upon N_m binding, given that it was initially characterized as a transcriptional regulator.

A total of 2228 identified N_m sites were densely populated on 323 protein-coding genes, suggesting cooperation among N_m sites in recruiting binding proteins. This may compensate for the modest preference of FUBP1 for a single N_m modification observed in EMSA experiments, contributing to notable enrichment of FUBP1 at N_m -modified regions. EMSA using densely modified oligos demonstrated a stronger preference of FUBP1 for N_m modifications. These observations highlight the importance of quantitative analysis in studies of N_m -mediated protein recruitment, as a single N_m may exert weaker effect compared to densely clustered N_m sites. Further investigation is needed to elucidate the underlying mechanism of the weak but additive role of N_m in recruiting its potential binding proteins.

Limitations of the study

The N_m -binding protein candidates we identified, aside from FUBP1, were not fully validated. More quantitative biochemical evidence supporting their interactions with N_m are required in the future. Moreover, the binding affinity of FUBP1 to individual N_m sites appears relatively weak compared to that of canonical “readers” of RNA modifications such as m^6A . Therefore, additional caution and context are needed when interpreting FUBP1 as an N_m -binding protein.

MATERIALS AND METHODS

Cell culture

HepG2 cells (American Type Culture Collection, HB-8065) were cultured with medium containing Dulbecco's modified Eagle's medium (Gibco, 11995040), 10% fetal bovine serum (Gibco, 2614079), and 1% penicillin-streptomycin (Gibco, 15140122) at 37°C with 5% CO₂ in the environment. Cells were passaged when reaching ~90% confluency at 1:4 ratio. Mycoplasma were tested by PCR with primers gggagcaaacaggattagataccct and tgcaccatctgtcactctgttaacctc every half a year.

In the small interfering RNA (siRNA)-mediated KD assay, cells at ~90% confluency were passaged at 1:4 ratio into 15-cm plates. Within 16 hours after passaging, 60 µl of Lipofectamine RNAiMAX Transfection Reagent (Invitrogen, 13778150) and 200 pmol of siRNA were diluted in 1 ml of Opti-MEM I Reduced Serum Medium (Gibco, 31985070), respectively. The solutions were mixed together and incubated at room temperature for 5 min before being added into the cell culture. KD reactions for fewer cells were scaled down on the basis of the bottom area of culture plates. The siRNA used in this paper include siControl (QIAGEN, 1027310), si*FBL* (QIAGEN, SI04164951), siControl2 (Invitrogen, 4390846), and si*FUBP1* (Invitrogen, s16966). Cells were harvested 48 hours posttransfection. KD efficiency was measured by qPCR with primers from Origene (HP205317).

U3 KD was conducted with *U3* allele-specific oligonucleotide (ASO) (mU*mC*mA*mC*mC*T*T*C*A*A*C*C*C*T*C*T*mC*mC*mA*mC*mU) controlled by green fluorescent protein ASO (mC*mU*mG*mC*mC*A*T*C*C*A*A*A*T*C*G*mU*mU*mA*mU*mC). Transfection was done by Lipofectamine 3000 transfection reagents (Invitrogen, L3000015), where 600 pmol of ASO and 8 µl of P3000 reagents were added to 125 µl of Opti-MEM medium. In parallel, 6 µl of Lipofectamine 3000 reagents were diluted in 125 µl of Opti-MEM medium. The two mixtures were mixed together, incubated at room temperature for 10 min, and then added to 2 ml of medium in a six-well plate. HepG2 cells corresponding to one-third of a 10-cm plate at ~90% confluency were then added to the transfected medium. Cells were cultured for 72 hours before harvest. KD efficiency was measured by qPCR primer pair CGTGTAGAGCACCGAAAACC and CACTCAGACCGCGTTCTCTC.

Immunoblotting

Samples were lysed in 2× NuPAGE LDS sample buffer (Invitrogen, NP0007) supplemented with 1:20 (v/v) 2-mercaptoethanol (Sigma-Aldrich, M6250-1L). After incubation at 95°C for 10 min, the denatured samples were loaded into 4 to 12% NuPAGE bis-tris gels (Invitrogen, NP0322BOX) and transferred onto nitrocellulose membranes (Bio-Rad, 1620115). Samples were blocked by 5% bovine serum albumin (BSA) (Fisher Scientific, BP1600-1) in PBS Tween-20 (PBST; Thermo Fisher Scientific, 28352) for 30 min, followed by overnight incubation at 4°C in primary antibodies with designated dilution ratios in 3% BSA diluted by PBST. Membranes were washed three times and incubated in the secondary antibody conjugated to HRP at room temperature for 1 hour. Protein signals were developed by SuperSignal West Dura Extended Duration Substrate (Thermo Fisher Scientific, 34075). Antibodies used in this study and their dilution ratios are as follows: anti-FUBP1 (Abcam, ab192867; 1:1000), anti-FUBP3 (Abcam, ab181025; 1:1000), anti-IGF2BP1 (MBL International, RN007P; 1:1000), anti-IGF2BP3 (MBL International, RN009P; 1:1000), anti-NCOA5 (Proteintech, 20175-1-AP; 1:500), anti-FMRP (Abcam, ab17722; 1:1000), and anti-rabbit immunoglobulin G (IgG) linked with HRP (Cell Signaling Technology, 7074S; 1:2000).

RNA affinity purification

The experiment followed the protocol in the previous publication (45) with adjustment. A total of 30 µl of Dynabeads MyOne Streptavidin C1 bead suspension was washed with RNA binding buffer [50 mM Hepes (pH 7.5), 150 mM NaCl, 0.5% NP-40 substitute, and 10 mM MgCl₂] and incubated in RNA binding buffer supplemented with yeast tRNA (100 µg/ml; Invitrogen, AM7119) for 1 hour at 4°C with rotation. After two rounds of washing, 400 pmol of N_m -modified or control probes were incubated with beads suspended in RNA binding buffer for 30 min at 4°C with rotation. Beads conjugated with oligos were washed with RNA wash buffer [50 mM Hepes (pH 7.5), 250 mM NaCl, 0.5% NP-40 substitute, and 10 mM MgCl₂] and then with protein incubation buffer [10 mM tris-HCl (pH 7.5), 150 mM KCl, 1.5 mM MgCl₂, 0.1% NP-40 substitute, and 0.5 mM dithiothreitol (DTT)] twice. HepG2 cell pellets in the volume of 45 µl were lysed in 400 µl of lysis buffer [50 mM tris (pH 7.5), 100 mM NaCl, 1% NP-40 substitute, 0.5% sodium deoxycholate, 100× proteinase inhibitor cocktail (Sigma-Aldrich, P8340), and 100× SUPERase-In RNase (ribonuclease) Inhibitor (Invitrogen, AM2696)] for 30 min at 4°C with rotation. The supernatant of the lysate was harvested by centrifugation at 12,000g for 15 min and separated to two equal

halves after saving 5% as input. The lysate was incubated with beads conjugated with oligos, supplemented with tRNA (50 µg/ml), 0.5 mM DTT, and 100× SUPERase-In, and incubated at room temperature for 30 min and 4°C for 90 min with rotation. The beads were washed three times with protein incubation buffer before removal of all supernatant.

LC-MS/MS analysis

Samples were prepared with three replicates, harvested on dry beads, and frozen on dry ice when shipped to the MS center. The beads were heated in 3× reducing LDS sample buffer with 15 mM DTT and 2 M biotin for 10 min at 95°C, and the supernatant was loaded on 12% bis-tris propane SDS–polyacrylamide gel electrophoresis gel for removal of detergent. The gel was run shortly and stained with colloidal Coomassie blue for gel cut of the whole lane. Gel pieces were reduced with DTT, alkylated with iodoacetamide, washed properly, and digested with trypsin overnight at 37°C. The extracted peptides were dried down, redissolved in 2.5% acetonitrile-water solution with 0.1% formic acid, and then run by nano-LC-MS/MS using a 2-hour gradient on a 0.075 mm-by-250 mm C18 column feeding into an Orbitrap Eclipse mass spectrometer.

The quantitation was done by Proteome Discoverer (Thermo Fisher Scientific; version 2.4). All MS/MS samples were searched with Mascot (Matrix Science, London, UK; version 2.6.2) using cRAP_20150130.fasta (124 entries) and uniprot-human_20201207 database (75777 entries) assuming trypsin digestion, with a fragment ion mass tolerance of 0.02 Da and a parent ion tolerance of 10.0 parts per million. The specified variable modifications included asparagine and glutamine deamidation, methionine oxidation, lysine and arginine methylation, and cysteine carbamidomethylation. Peptides were validated by Percolator with a 0.01 posterior error probability threshold. The data were searched with a decoy database to set the false discovery rate (FDR) to 1%. The peptides were quantified using the precursor abundance based on intensity, with the peak abundance normalized by total peptide amount. The sum of peptide group abundances for each sample were normalized to the maximum sum of the analyzed samples. The protein ratios were calculated using summed abundance for each replicate separately and the geometric median of the resulting ratios was used as the protein ratios. To compensate for missing values in some of the replicates, the replicate-based resampling imputation mode was used. The significance of differential expression was generated by analysis of variance (ANOVA) test, and the *P* values were adjusted by the Benjamini-Hochberg method.

Protein purification

FUBP1-strep were expressed in 100 ml of Expi293F cells (Gibco, A14527) for 48 hours, lysed in lysis buffer [20 mM tris (pH 8.0), 150 mM KCl, 2 mM EDTA, 100× phenylmethylsulfonyl fluoride (PMSF), 1% NP-40 substitute, and 0.5 mM DTT] at 4°C for 30 min. Supernatant was harvested by centrifugation and diluted to threefold volume before passing through 0.22-µm filter. Samples were loaded to 200 µl of Strep-Tactin Sepharose resin (IBA, 21201010) and incubated at 4°C for 1 hour. The resin was flowed through by lysate twice and washed with 25 ml of wash buffer [50 mM tris (pH 8.0), 250 mM NaCl, 0.05% NP-40 substitute, 0.5 mM DTT, and 100× PMSF]. Protein was eluted six times with elution buffer [50 mM tris, (pH 8.0), 250 mM NaCl, 0.01% NP-40 substitute, 0.5 mM DTT, and 2.5 mM desthiobiotin] with 100 µl each time, concentrated with 30-kDa

Amicon Ultra Centrifugal Filter (Millipore, UFC503008), and exchanged to storage buffer [50 mM tris, (pH 8.0), 150 mM KCl, 0.1 mM EDTA, and 20% glycerol]. The purified protein was aliquoted, snap frozen in liquid nitrogen, and stored at –80°C.

Electrophoretic mobility shift assay

Probes for EMSA were labeled with 6-carboxyfluorescein (FAM) fluorophore at the 5′ end, with sequence noted in the figures. Probes were refolded by incubation at 75°C for 2 min and natural cooling to room temperature for 10 min. FUBP1 proteins of designated concentrations were incubated with FAM-labeled oligos in binding buffer [10 mM tris (pH 7.5), 140 mM KCl, 10 mM NaCl, 1 mM MgCl₂, 10% glycerol, 1 mM DTT, and SUPERase-In RNase Inhibitor (1 U/µl)] at room temperature for 30 min. The mixtures were loaded to 4 to 20% Novex tris-borate EDTA (TBE) gels (EC62255BOX) that have been rerun at 90 V for 30 min at 4°C in 0.5× TBE. Gels were run for 2 hours before imaging. The fluorescence signal intensity was quantified by ImageJ, and *K_d* value was calculated by GraphPad.

UHPLC-QQQ-MS/MS quantification

HepG2 mRNA were purified by two rounds of polyadenylate [poly(A)] selection following the commercial protocol of Dynabeads mRNA DIRECT Purification Kit (Invitrogen, 61011). HepG2 chromatin were purified following the published protocol reported previously (7), with adjustment of lysis buffer to 10 mM tris-HCl (pH 7.5), 0.15% NP-40 substitute, and 150 mM NaCl. The caRNA were purified from fractionated chromatin, followed by two rounds of RiboMinus reactions according to the commercial instruction of RiboMinus Eukaryote System v2 (Invitrogen, A15026).

The verification of rRNA contamination was performed by qPCR for 18S (with primer pair CGGACAGGATTGACAGATTG and CAAATCGCTCCACCAACTAA) and 45S (with primer pair TCGCT-GCGATCTATTGAAAG and AGGAAGACGAACGGAAGGAC) rRNA. The same amount of RNA was used for RT. The absolute amount of cDNA product was quantitated by the control of PCR product of the two primer pairs with known amount of DNA template.

For UHPLC-QQQ-MS/MS detection, 100 ng of RNA was digested with Nuclease P1 [New England Biolabs (NEB), M0660S] at 37°C overnight, followed by digestion with shrimp alkaline phosphatase (rSAP, NEB, M0371S) in rCutSmart buffer 2 hours at 37°C. The samples were diluted and filtered by 0.22-µm polyvinylidene difluoride filter (Millipore, SLGVR33RB) and then injected into a C18 reversed phase column coupled online to Agilent 6460 LC-MS/MS spectrometer in positive electrospray ionization mode. Nucleosides were quantified using nucleoside-to-base transitions of *A_m* (282 to 136), *A* (268 to 136), *G_m* (198.1 to 152.1), *G* (284 to 152), *C_m* (258.2 to 112), and *C* (244 to 112). The signal intensity of *N_m* nucleotides was normalized to the corresponding unmodified nucleotides to enable comparison of samples with different length.

Photoactivatable ribonucleoside-enhanced cross-linking and immunoprecipitation

The experiment was performed with an adapted protocol from a previous report (46). Two biological replicates of 150 million HepG2 cells were cultured with 200 µM 4-Thiouridine for 14 hours. Cells were cross-linked by 365-nm ultraviolet (UV) at 1500 J/m² twice, harvested, and lysed by iCLIP lysis buffer [50 mM tris (pH 7.5), 100 mM NaCl, 1% NP-40 substitute, 0.1% SDS, 0.5% sodium deoxycholate, 100× protease inhibitor cocktail, and 100× SUPERase-In

RNase Inhibitor] at 4°C for 15 min with rotation. To release FUBP1 proteins associated with the chromatin, cell lysate was supplemented with 1% SDS, sonicated at 30% amplitude with 2-s:4-s cycles for 1 min on ice. The lysate was 10-fold diluted by iCLIP lysis buffer without SDS, centrifuged to save the supernatant, and treated with RNase T1 (0.2 U/μl; Thermo Fisher Scientific, EN0642) at 22°C for 15 min before being quenched on ice for 5 min. Protein G beads (Invitrogen, 10009D) were conjugated with 20 μg of FUBP1 antibody (Abcam, ab192867) or rabbit IgG (Cell Signaling Technology, 2729) by incubation at 4°C for 1 hour, then washed, and mixed with RNase T1-treated lysate to rotate at 4°C for 4 hours. Beads were washed three times with CLIP wash buffer [50 mM tris (pH 7.5), 300 mM KCl, 0.05% NP-40 substitute, 1000× proteinase inhibitor cocktail, and 1000× SUPERase-In RNase Inhibitor], digested with RNase T1 (10 U/μl) at 22°C for 8 min, washed with CLIP high-salt buffer [50 mM tris (pH 7.5), 500 mM KCl, 0.05% NP-40 substitute, 1000× protease inhibitor cocktail, and 1000× SUPERase-In RNase Inhibitor] and T4 polynucleotide kinase (T4 PNK) buffer without DTT three times each, and underwent end repair by T4 PNK (NEB, M0201L). The immunoprecipitation was validated by biotinylation, eluted with proteinase K digestion (Thermo Fisher Scientific, EO0491), with purified RNA used to construct libraries with NEBNext Small RNA Library Prep Set (NEB, E7330S). Sequencing was performed by Illumina NovaSeq 6000 paired-end 50-bp reads.

Bioinformatic analysis of PAR-CLIP

For both FUBP1 PAR-CLIP and its IgG control, adapters were trimmed by Cutadapt (47), and reads were mapped to hg38 human genome by HISAT2 (48), with parameter “--reorder --no-unal --pen-noncansplice 12.” Duplicates were filtered by Picard MarkDuplicates. R1 reads with mapping quality higher than 30 were used for identification of FUBP1-binding clusters by wavCluster (49, 50), with removal of reads containing “^” in the MD tag. The identified clusters were overlapped and then annotated by Homer (51) annotatepeaks. The binding clusters identified by IgG were supposed to be removed from the FUBP1-binding clusters. Nevertheless, we identified no clusters from IgG control. Metaplots centering at N_m sites were generated by deepTools (52). Motifs were identified by Homer findMotifsGenome with parameter “-rna -size 200 -len 6”.

Cross-linking immunoprecipitation sequencing

The experiment followed the procedure in the previous publication (53). Three biological replicates of 150 million HepG2 cells were harvested after 48 hours of FBL KD, with the KD efficiency validated by RT-qPCR with primers from Origene (HP205317). Samples were cross-linked three times by UV (1500 J/m²) at 254 nm, lysed, centrifuged, and digested with the same condition as PAR-CLIP. After saving 2% of the lysate as input, lysates were incubated with beads treated the same as PAR-CLIP. The following steps were exactly the same, with the exception that input samples were treated separately with RNase T1 (10 U/μl) at 22°C for 8 min, followed by proteinase K digestion and end repair in the next. Sequencing by Illumina NovaSeq X was conducted for single-end 100-bp reads.

Bioinformatic analysis of CLIP-seq

Cutadapt (47) was used to trim the adapters, and HISAT2 (48) mapping to the reference genome (hg38) was performed, with parameter “--reorder --no-unal --pen-noncansplice 12.” Peaks were identified by Piranha (54) with input samples as the covariate, and bin size

designated as 50 bp. The identified peaks of siControl and siFBL were merged and annotated by Homer (51) annotatepeaks. Differential binding was analyzed by DiffBind (55), with relative log expression normalization considering the background (normalize = DBA_NORM_RLE, background = TRUE). Down-regulated peaks were defined as a log₂ fold change of <−0.58 and an FDR < 0.1. Peaks within 2 kb of N_m sites were considered as N_m modified peaks for computing binding fold change.

RNA sequencing

HepG2 cell RNA was harvested by TRIzol reagent (Invitrogen, 15596026) after 48 hours of siRNA-mediated KD or 72 hours of ASO-mediated KD with three biological replicates. RNA was purified with the manufacturer's procedures, and the KD was measured by RT-qPCR with primers from Origene (HP207343). After poly(A) RNA selection with Dynabeads mRNA DIRECT Purification Kit and the samples were constructed to libraries with SMARTer Stranded Total RNA-Seq Kit v2–Pico Input Mammalian (Takara, 634412). Next-generation sequencing was performed with Illumina NovaSeq 6000 paired-end 150-bp reads.

RNA immunoprecipitation sequencing

Two biological replicates of HepG2 cells were harvested after 48 hours of FBL KD. The KD efficiency was validated by RT-qPCR. Cells were lysed by iCLIP lysis buffer at 4°C for 15 min with rotation, followed by centrifugation at 13,000g for 15 min. Protein G beads were conjugated with 16 μg of FUBP1 antibody by incubation at 4°C for 1 hour, then washed, and mixed with lysate supernatant to rotate at 4°C for 4 hours. Beads were washed three times with CLIP wash buffer and eluted with proteinase K digestion (Thermo Fisher Scientific, EO0491). Input samples harvested before immunoprecipitation was also digested with proteinase K for RNA recovery. The purified RNA was constructed for library with SMARTer Stranded Total RNA-Seq Kit v2–Pico Input Mammalian. Next-generation sequencing was performed with Illumina NovaSeq 6000 paired-end 150-bp reads.

Bioinformatic analysis for RNA sequencing and RIP-seq

Adapters were trimmed by Cutadapt (47), and reads were mapped to hg38 human genome by HISAT2 (48), with parameter “--reorder --no-unal --pen-noncansplice 12.” Reads were counted for each gene by HTSeq (56), and differential expression analysis by DESeq2 (57) was conducted. For RIP-seq samples, the differential expression analysis compares immunoprecipitation versus input samples, where the fold change was considered as enrichment of each gene. The enrichment was further compared between control and FBL KD samples to measure the differential binding of FUBP1 upon N_m depletion.

The AS events were detected by rMATS (58), where we used the input samples of RIP-seq for analysis upon FBL KD. Only junction reads were used for calculation of AS scores. Events with an FDR of <0.1 were considered as events.

Nm-mut-seq

Three biological replicates of HepG2 caRNA were fractionated and purified in the same procedures as UHPLC-QQQ-MS/MS analysis. rRNA was removed by RiboMinus Eukaryote System v2, and the residual RNA was cleaned up by RNA Clean & Concentrator-5 (Zymo Research, R1014) with removal of a short RNA of <200 nt. The construction of Nm-mut-seq followed the published procedures (27), using 3' linker 5'rApp-NNNNNATCACGAGATCGGAAGAG-CACACGTCT-3SpC3 and 5' SR adapter supplied in the NEBNext

Small RNA Library Prep Set. Sequencing by Illumina NovaSeq X was conducted with single-end 100-bp reads.

Bioinformatic analysis for Nm-mut-seq

Adapters were trimmed by Cutadapt (47), and duplicates marked by unique molecular identifier (UMI) was removed using BBMap (59). Reads were mapped to hg38 human genome by TopHat2 (60), with parameters “-g 4 -N 3 --read-edit-dist 3.” Mutations to T were identified by JACUSA (39), with further selection of mutated sites in all three replicates having a sequencing depth of ≥ 10 , a mutated read depth of ≥ 3 , mutation ratios of >3 -fold of input mutation ratios, and mutation ratios of >1.5 -fold of background mutation ratios. N_m sites were annotated by Homer (51) annotatepeaks.

Statistical analysis

P values annotated in figures were quantified on the basis of two-tailed Student's *t* test. Chi-squared test was used to statistically test the preferential occurrence of different splicing events with increased or decreased inclusion level differences. The K_d value and SD for EMSA were fitted by GraphPad Prism under the mode of “one site-specific binding with Hill slope.”

Supplementary Materials

The PDF file includes:

Figs. S1 to S5

Legends for tables S1 to S10

Other Supplementary Material for this manuscript includes the following:

Tables S1 to S10

REFERENCES AND NOTES

- B. S. Zhao, I. A. Roundtree, C. He, Post-transcriptional gene regulation by mRNA modifications. *Nat. Rev. Mol. Cell Biol.* **18**, 31–42 (2017).
- H. Shi, J. Wei, C. He, Where, when, and how: Context-dependent functions of RNA methylation writers, readers, and erasers. *Mol. Cell* **74**, 640–650 (2019).
- N. Liu, Q. Dai, G. Zheng, C. He, M. Parisien, T. Pan, *N*⁶-methyladenosine-dependent RNA structural switches regulate RNA-protein interactions. *Nature* **518**, 560–564 (2015).
- H. Zhou, I. J. Kimsey, E. N. Nikolova, B. Sathayamoorthy, G. Grazioli, J. McSally, T. Bai, C. H. Wunderlich, C. Kreutz, I. Andricioaei, H. M. Al-Hashimi, m1A and m1G disrupt A-RNA structure through the intrinsic instability of Hoogsteen base pairs. *Nat. Struct. Mol. Biol.* **23**, 803–810 (2016).
- X. Wang, Z. Lu, A. Gomez, G. C. Hon, Y. Yue, D. Han, Y. Fu, M. Parisien, Q. Dai, G. Jia, B. Ren, T. Pan, C. He, *N*⁶-methyladenosine-dependent regulation of messenger RNA stability. *Nature* **505**, 117–120 (2014).
- X. Wang, B. S. Zhao, I. A. Roundtree, Z. Lu, D. Han, H. Ma, X. Weng, K. Chen, H. Shi, C. He, *N*⁶-methyladenosine modulates messenger RNA translation efficiency. *Cell* **161**, 1388–1399 (2015).
- J. Liu, X. Dou, C. Chen, C. Liu, M. M. Xu, S. Zhao, B. Shen, Y. Gao, D. Han, C. He, *N*⁶-methyladenosine of chromosome-associated regulatory RNA regulates chromatin state and transcription. *Science* **367**, 580–586 (2020).
- X. Dou, Y. Xiao, C. Shen, K. Wang, T. Wu, C. Liu, Y. Li, X. Yu, J. Liu, Q. Dai, K. Pajdzik, C. Ye, R. Ge, B. Gao, J. Yu, S. Sun, M. Chen, J. Chen, C. He, RBFOX2 recognizes *N*⁶-methyladenosine to suppress transcription and block myeloid leukaemia differentiation. *Nat. Cell Biol.* **25**, 1359–1368 (2023).
- Z. Zou, C. Sepich-Poore, X. Zhou, J. Wei, C. He, The mechanism underlying redundant functions of the YTHDF proteins. *Genome Biol.* **24**, 17 (2023).
- W. Xiao, S. Adhikari, U. Dahal, Y. S. Chen, Y. J. Hao, B. F. Sun, H. Y. Sun, A. Li, X. L. Ping, W. Y. Lai, X. Wang, H. L. Ma, C. M. Huang, Y. Yang, N. Huang, G. B. Jiang, H. L. Wang, Q. Zhou, X. J. Wang, Y. L. Zhao, Y. G. Yang, Nuclear m⁶A reader YTHDC1 regulates mRNA splicing. *Mol. Cell* **61**, 507–519 (2016).
- I. A. Roundtree, G. Z. Luo, Z. Zhang, X. Wang, T. Zhou, Y. Cui, J. Sha, X. Huang, L. Guerrero, P. Xie, E. He, B. Shen, C. He, YTHDC1 mediates nuclear export of *N*⁶-methyladenosine methylated mRNAs. *eLife* **6**, e31311 (2017).
- Y. Fu, X. Zhuang, m⁶A-binding YTHDF proteins promote stress granule formation. *Nat. Chem. Biol.* **16**, 955–963 (2020).
- Z. Zou, J. Wei, Y. Chen, Y. Kang, H. Shi, F. Yang, Z. Shi, S. Chen, Y. Zhou, C. Sepich-Poore, X. Zhuang, X. Zhou, H. Jiang, Z. Wen, P. Jin, C. Luo, C. He, FMRP phosphorylation modulates neuronal translation through YTHDF1. *Mol. Cell* **83**, 4304–4317.e8 (2023).
- C. Chen, W. Liu, J. Guo, Y. Liu, X. Liu, J. Liu, X. Dou, R. Le, Y. Huang, C. Li, L. Yang, X. Kou, Y. Zhao, Y. Wu, J. Chen, H. Wang, B. Shen, Y. Gao, S. Gao, Nuclear m⁶A reader YTHDC1 regulates the scaffold function of LINE1 RNA in mouse ESCs and early embryos. *Protein Cell* **12**, 455–474 (2021).
- S. Höfler, T. Carlomagno, Structural and functional roles of 2'-O-ribose methylations and their enzymatic machinery across multiple classes of RNAs. *Curr. Opin. Struct. Biol.* **65**, 42–50 (2020).
- J. N. Yelland, J. P. K. Bravo, J. J. Black, D. W. Taylor, A. W. Johnson, A single 2'-O-methylation of ribosomal RNA gates assembly of a functional ribosome. *Nat. Struct. Mol. Biol.* **30**, 91–98 (2023).
- N. Krogh, M. D. Jansson, S. J. Häfner, D. Tehler, U. Birkedal, M. Christensen-Dalsgaard, A. H. Lund, H. Nielsen, Profiling of 2'-O-Me in human rRNA reveals a subset of fractionally modified positions and provides evidence for ribosome heterogeneity. *Nucleic Acids Res.* **44**, 7884–7895 (2016).
- J. Choi, G. Indrisiunaite, H. DeMirici, K.-W. Jeong, J. Wang, A. Petrov, A. Prabhakar, G. Rechavi, D. Dominissini, C. He, M. Ehrenberg, J. D. Puglisi, 2'-O-methylation in mRNA disrupts tRNA decoding during translation elongation. *Nat. Struct. Mol. Biol.* **25**, 208–216 (2018).
- B. A. Elliott, H.-T. Ho, S. V. Ranganathan, S. Vangaveti, O. Ilkayeva, H. Abou Assi, A. K. Choi, P. F. Agris, C. L. Holley, Modification of messenger RNA by 2'-O-methylation regulates gene expression in vivo. *Nat. Commun.* **10**, 3401 (2019).
- R. Züst, L. Cervantes-Barragan, M. Habjan, R. Maier, B. W. Neuman, J. Ziebuhr, K. J. Szretter, S. C. Baker, W. Barchet, M. S. Diamond, S. G. Siddell, B. Ludewig, V. Thiel, Ribose 2'-O-methylation provides a molecular signature for the distinction of self and non-self mRNA dependent on the RNA sensor Mda5. *Nat. Immunol.* **12**, 137–143 (2011).
- M. Ringard, V. Marchand, E. Decroly, Y. Motorin, Y. Bennisser, FTSJ3 is an RNA 2'-O-methyltransferase recruited by HIV to avoid innate immune sensing. *Nature* **565**, 500–504 (2019).
- S. J. Häfner, M. D. Jansson, K. Altinel, K. L. Andersen, Z. Abay-Nørgaard, P. Ménard, M. Fontenas, D. M. Sørensen, D. M. Gay, F. S. Arendrup, D. Tehler, N. Krogh, H. Nielsen, M. L. Kraushar, A. Kirkeby, A. H. Lund, Ribosomal RNA 2'-O-methylation dynamics impact cell fate decisions. *Dev. Cell* **58**, 1593–1609.e9 (2023).
- P. Prusiner, N. Yathindra, M. Sundaralingam, Effect of ribose O(2')-methylation on the conformation of nucleosides and nucleotides. *Biochim. Biophys. Acta* **366**, 115–123 (1974).
- C. C. He, L. A. Hamlow, Z. J. Devereaux, Y. Zhu, Y. w. Nei, L. Fan, C. P. McNary, P. Maitre, V. Steinmetz, B. Schindler, I. Compagnon, P. B. Armentrout, M. T. Rodgers, Structural and energetic effects of O2'-ribose methylation of protonated purine nucleosides. *J. Phys. Chem. B* **122**, 9147–9160 (2018).
- H. Abou Assi, A. K. Rangadurai, H. Shi, B. Liu, M. C. Clay, K. Erharter, C. Kreutz, C. L. Holley, H. M. Al-Hashimi, 2'-O-methylation can increase the abundance and lifetime of alternative RNA conformational states. *Nucleic Acids Res.* **48**, 12365–12379 (2020).
- S. K. Natchiar, A. G. Myasnikov, H. Kratzat, I. Hazemann, B. P. Klaholz, Visualization of chemical modifications in the human 80S ribosome structure. *Nature* **551**, 472–477 (2017).
- L. Chen, L.-S. Zhang, C. Ye, H. Zhou, B. Liu, B. Gao, Z. Deng, C. Zhao, C. He, B. C. Dickinson, Nm-Mut-seq: A base-resolution quantitative method for mapping transcriptome-wide 2'-O-methylation. *Cell Res.* **33**, 727–730 (2023).
- Y. Li, Y. Yi, X. Gao, X. Wang, D. Zhao, R. Wang, L. S. Zhang, B. Gao, Y. Zhang, L. Zhang, Q. Cao, K. Chen, 2'-O-methylation at internal sites on mRNA promotes mRNA stability. *Mol. Cell* **84**, 2320–2336.e2326 (2024).
- J. S. Elman, T. K. Ni, K. E. Mengwasser, D. Jin, A. Wronski, S. J. Elledge, C. Kuperwasser, Identification of FUBP1 as a long tail cancer driver and widespread regulator of tumor suppressor and oncogene alternative splicing. *Cell Rep.* **28**, 3435–3449.e5 (2019).
- P. Jiang, M. Huang, W. Qi, F. Wang, T. Yang, T. Gao, C. Luo, J. Deng, Z. Yang, T. Zhou, Y. Zou, G. Gao, X. Yang, FUBP1 promotes neuroblastoma proliferation via enhancing glycolysis—a new possible marker of malignancy for neuroblastoma. *J. Exp. Clin. Cancer Res.* **38**, 400 (2019).
- A. G. Jacob, R. K. Singh, F. Mohammad, T. W. Bebee, D. S. Chandler, The splicing factor FUBP1 is required for the efficient splicing of oncogene MDM2 pre-mRNA. *J. Biol. Chem.* **289**, 17350–17364 (2014).
- S. Ebersberger, C. Hipp, M. M. Mulorz, A. Buchbender, D. Hubrich, H. S. Kang, S. Martínez-Lumbreras, P. Kristofori, F. X. R. Sutandy, L. Llacsahuanga Allica, J. Schöndelf, C. Bakisoglu, A. Busch, H. Hänel, K. Tretow, M. Welzel, A. Di Liddo, M. M. Möckel, K. Zarnack, I. Ebersberger, S. Legewie, K. Luck, M. Sattler, J. König, FUBP1 is a general splicing factor facilitating 3' splice site recognition and splicing of long introns. *Mol. Cell* **83**, 2653–2672.e15 (2023).

33. H. Huang, H. Weng, W. Sun, X. Qin, H. Shi, H. Wu, B. S. Zhao, A. Mesquita, C. Liu, C. L. Yuan, Y.-C. Hu, S. Hüttelmaier, J. R. Skibbe, R. Su, X. Deng, L. Dong, M. Sun, C. Li, S. Nachtergaele, Y. Wang, C. Hu, K. Ferchen, K. D. Greis, X. Jiang, M. Wei, L. Qu, J.-L. Guan, C. He, J. Yang, J. Chen, Recognition of RNA N⁶-methyladenosine by IGF2BP proteins enhances mRNA stability and translation. *Nat. Cell Biol.* **20**, 285–295 (2018).
34. C. Liu, X. Dou, Y. Zhao, L. Zhang, L. Zhang, Q. Dai, J. Liu, T. Wu, Y. Xiao, C. He, IGF2BP3 promotes mRNA degradation through internal m⁷G modification. *Nat. Commun.* **15**, 7421 (2024).
35. F. Sauv  , L. D. McBroom, J. Gallant, A. N. Moraitis, F. Labrie, V. Gigu  re, CIA, a novel estrogen receptor coactivator with a bifunctional nuclear receptor interacting determinant. *Mol. Cell Biol.* **21**, 343–353 (2001).
36. O. Oksuz, J. E. Henninger, R. Warneford-Thomson, M. M. Zheng, H. Erb, A. Vancura, K. J. Overholt, S. W. Hawken, S. F. Banani, R. Lauman, L. N. Reich, A. L. Robertson, N. M. Hannett, T. I. Lee, L. I. Zon, R. Bonasio, R. A. Young, Transcription factors interact with RNA to regulate genes. *Mol. Cell* **83**, 2449–2463.e13 (2023).
37. The ENCODE Project Consortium, An integrated encyclopedia of DNA elements in the human genome. *Nature* **489**, 57–74 (2012).
38. Y. Luo, B. C. Hitz, I. Gabdank, J. A. Hilton, M. S. Kagda, B. Lam, Z. Myers, P. Sud, J. Jou, K. Lin, U. K. Baymuradov, K. Graham, C. Litton, S. R. Miyasato, J. S. Strattan, O. Jolanki, J. W. Lee, F. Y. Tanaka, P. Adenekan, E. O'Neill, J. M. Cherry, New developments on the encyclopedia of DNA elements (ENCODE) data portal. *Nucleic Acids Res.* **48**, D882–D889 (2020).
39. M. Piechotta, E. Wyler, U. Ohler, M. Landthaler, C. Dieterich, JACUSA: Site-specific identification of RNA editing events from replicate sequencing data. *BMC Bioinformatics* **18**, 7 (2017).
40. Z. Wang, C. B. Burge, Splicing regulation: From a parts list of regulatory elements to an integrated splicing code. *RNA* **14**, 802–813 (2008).
41. S. Sharma, L. A. Kohlstaedt, A. Damianov, D. C. Rio, D. L. Black, Polypyrimidine tract binding protein controls the transition from exon definition to an intron defined spliceosome. *Nat. Struct. Mol. Biol.* **15**, 183–191 (2008).
42. B. Liu, T. Wu, B. A. Miao, F. Ji, S. Liu, P. Wang, Y. Zhao, Y. Zhong, A. Sundaram, T. B. Zeng, M. Majcherska-Agrawal, R. J. Keenan, T. Pan, C. He, snoRNA-facilitated protein secretion revealed by transcriptome-wide snoRNA target identification. *Cell* **188**, 465–483.e22 (2025).
43. X. Zhao, Y.-T. Yu, Targeted pre-mRNA modification for gene silencing and regulation. *Nat. Methods* **5**, 95–100 (2008).
44. J. Ge, H. Liu, Y. T. Yu, Regulation of pre-mRNA splicing in *Xenopus* oocytes by targeted 2'-O-methylation. *RNA* **16**, 1078–1085 (2010).
45. R. R. Edupuganti, S. Geiger, R. G. H. Lindeboom, H. Shi, P. J. Hsu, Z. Lu, S. Y. Wang, M. P. A. Baltissen, P. Jansen, M. Rossa, M. M  ller, H. G. Stunnenberg, C. He, T. Carell, M. Vermeulen, N⁶-methyladenosine (m⁶A) recruits and repels proteins to regulate mRNA homeostasis. *Nat. Struct. Mol. Biol.* **24**, 870–878 (2017).
46. H. Shi, X. Wang, Z. Lu, B. S. Zhao, H. Ma, P. J. Hsu, C. Liu, C. He, YTHDF3 facilitates translation and decay of N⁶-methyladenosine-modified RNA. *Cell Res.* **27**, 315–328 (2017).
47. M. Martin, Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet J* **17**, 3 (2011).
48. D. Kim, J. M. Paggi, C. Park, C. Bennett, S. L. Salzberg, Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. *Nat. Biotechnol.* **37**, 907–915 (2019).
49. C. Sievers, T. Schlumpf, R. Sawarkar, F. Comoglio, R. Paro, Mixture models and wavelet transforms reveal high confidence RNA-protein interaction sites in MOV10 PAR-CLIP data. *Nucleic Acids Res.* **40**, e160 (2012).
50. F. Comoglio, C. Sievers, R. Paro, Sensitive and highly resolved identification of RNA-protein interaction sites in PAR-CLIP data. *BMC Bioinformatics* **16**, 32 (2015).
51. S. Heinz, C. Benner, N. Spann, E. Bertolino, Y. C. Lin, P. Laslo, J. X. Cheng, C. Murre, H. Singh, C. K. Glass, Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. *Mol. Cell* **38**, 576–589 (2010).
52. F. Ram  rez, D. P. Ryan, B. Gr  ning, V. Bhardwaj, F. Kilpert, A. S. Richter, S. Heyne, F. D  ndar, T. Manke, deepTools2: A next generation web server for deep-sequencing data analysis. *Nucleic Acids Res.* **44**, W160–W165 (2016).
53. P. C. He, J. Wei, X. Dou, B. T. Harada, Z. Zhang, R. Ge, C. Liu, L.-S. Zhang, X. Yu, S. Wang, R. Lyu, Z. Zou, M. Chen, C. He, Exon architecture controls mRNA m⁶A suppression and gene expression. *Science* **379**, 677–682 (2023).
54. P. J. Uren, E. Bahrami-Samani, S. C. Burns, M. Qiao, F. V. Karginov, E. Hodges, G. J. Hannon, J. R. Sanford, L. O. Penalva, A. D. Smith, Site identification in high-throughput RNA-protein interaction data. *Bioinformatics* **28**, 3013–3020 (2012).
55. R. Stark, G. Brown, DiffBind: Differential binding analysis of ChIP-seq peak data. *In R package version* **100**, (2011).
56. S. Anders, P. T. Pyl, W. Huber, HTSeq—A Python framework to work with high-throughput sequencing data. *Bioinformatics* **31**, 166–169 (2015).
57. M. I. Love, W. Huber, S. Anders, Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 550 (2014).
58. S. Shen, J. W. Park, Z. X. Lu, L. Lin, M. D. Henry, Y. N. Wu, Q. Zhou, Y. Xing, rMATS: Robust and flexible detection of differential alternative splicing from replicate RNA-seq data. *Proc. Natl. Acad. Sci. U.S.A.* **111**, E5593–E5601 (2014).
59. B. Bushnell. (United States, 2014).
60. D. Kim, G. Pertea, C. Trapnell, H. Pimentel, R. Kelley, S. L. Salzberg, TopHat2: Accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biol.* **14**, R36 (2013).

Acknowledgments: We are grateful to the Genomics Facility at University of Chicago for high-throughput sequencing, the Mass Spectrometry Facility at University of Chicago for UHPLC-QQQ-MS/MS, and the Proteomics & Metabolomics Facility at University of Nebraska–Lincoln for LC-MS/MS analysis. **Funding:** This work is supported by the National Institute of Health RM1 HG008935 and R01 HL155909 (to C.H.). C.H. is an investigator of the Howard Hughes Medical Institute. **Author contributions:** Conceptualization: C.H. and L.Z. Methodology: B.G., L.C., L.Z., W.L., and B.L. Investigation: B.G., L.Z., and B.L. Visualization: B.G., Z.Z., L.Z., and W.L. Resources: C.H., L.Z., and L.C. Funding acquisition: C.H. and L.Z. Formal analysis: B.G., L.Z., and W.L. Software: B.G. and L.Z. Validation: L.Z., B.J., L.C., and B.L. Data curation: L.Z. Project administration: C.H. and L.Z. Supervision: C.H. and L.Z. Writing—original draft: C.H., B.G., and L.Z. Writing—review and editing: C.H., B.G., Z.Z., and L.Z. **Competing interests:** C.H. is a scientific founder; a member of the scientific advisory board and equity holder of Aferna Bio, AllyRNA, and Ellis Bio; a scientific cofounder and equity holder of Accent Therapeutics; and a member of the scientific advisory board of Rona Therapeutics and Element Biosciences. All other authors declare that they have no competing interests. **Data and materials availability:** All data needed to evaluate the conclusions in the paper are present in the paper and/or the Supplementary Materials. The sequencing data are accessible in the Gene Expression Omnibus through accession numbers GSE294792, GSE294793, and GSE294794. Published N_m-mut-seq dataset of internal mRNA (GSE174518); ENCODE datasets of eCLIP experiments targeting FUBP3 (ENCSR486YGP), IGF2BP1 (ENCSR744GEU), and IGF2BP3 (ENCSR993OLA); and their IgG control (GSE77629) were used for analysis.

Submitted 20 April 2025
Accepted 12 September 2025
Published 17 October 2025
10.1126/sciadv.ady3894