Published in final edited form as:
Science. 2023 February 17; 379(6633): 677–682. doi:10.1126/science.abj9090.

Exon architecture controls mRNA m^{6}A suppression and gene expression

P. Cody He[1,2,3,†], **Jiangbo Wei**[1,3,†], **Xiaoyang Dou**[1,3,†], **Bryan T. Harada**[1,3,†,‡], **Zijie Zhang**[1,3,4], **Ruiqi Ge**[1,3], **Chang Liu**[1,3,§], **Li-Sheng Zhang**[1,3], **Xianbin Yu**[1,3], **Shuai Wang**[5], **Ruitu Lyu**[1,3], **Zhongyu Zou**[1,3], **Mengjie Chen**[6,7], **Chuan He**[1,2,3,*]

[1]Department of Chemistry, Department of Biochemistry and Molecular Biology, Institute for Biophysical Dynamics, The University of Chicago, Chicago, IL 60637, USA.

[2]Committee on Immunology, The University of Chicago, Chicago, IL 60637, USA.

[3]Howard Hughes Medical Institute, The University of Chicago, Chicago, IL 60637, USA.

[4]State Key Laboratory for Conservation and Utilization of Bio-Resources, School of Life Sciences, Yunnan University, Kunming, Yunnan 650091, China.

[5]Department of Neurobiology, The University of Chicago, Chicago, IL 60637, USA.

[6]Department of Human Genetics, The University of Chicago, Chicago, IL 60637, USA.

[7]Section of Genetic Medicine, Department of Medicine, The University of Chicago, Chicago, IL 60637, USA.

Abstract

N^6–methyladenosine (m^{6}A) is the most abundant mRNA modification and plays crucial roles in diverse physiological processes. Utilizing a Massively Parallel Assay for m^{6}A (MPm^{6}A), we discover that m^{6}A specificity is globally regulated by “suppressors” that prevent m^{6}A deposition in unmethylated transcriptome regions. We identify Exon Junction Complexes (EJCs) as m^{6}A suppressors that protect exon junction-proximal RNA within coding sequences from methylation

*Corresponding author. chuanhe@uchicago.edu.
‡Present address: Cell Press, Cambridge, MA 02139, USA.
§Present address: Department of Cellular and Molecular Medicine, University of California, San Diego, La Jolla, CA 92093, USA.
†These authors contributed equally to this work.

Author contributions: Conceptualization: PCH; Methodology: PCH, JW, XD, BTH, ZZ, CH; Formal analysis: PCH, XD, XY, RL; Investigation: PCH, JW, XD, BTH, SW, RG, CL, LZ, ZZ; Visualization: PCH, XD, JW; Funding acquisition: PCH, BTH, CH; Project administration: CH; Supervision: MC, CH; Writing – original draft: PCH, CH; Writing – review & editing: PCH, JW, XD, BTH, ZZ, SW, RL, MC, CH.

Competing interests: C.H. is a scientific founder and a scientific advisory board member of Accent Therapeutics, Inc., Aferna Bio, Inc., and AccuraDX Inc. The other authors declare no competing interests.

SUPPLEMENTARY MATERIALS
Materials and Methods
Supplementary Text
Figs. S1 to S30
Tables S1 to S6
References (35–79)

and regulate mRNA stability through m^{6}A suppression. EJC suppression of m^{6}A underlies multiple global characteristics of mRNA m^{6}A specificity, with the local range of EJC protection sufficient to suppress m^{6}A deposition in average-length internal exons, but not in long internal and terminal exons. EJC-suppressed methylation sites co-localize with EJC-suppressed splice sites, suggesting that exon architecture broadly determines local mRNA accessibility to regulatory complexes.

N^6–methyladenosine (m^{6}A), the most prevalent mRNA modification in mammals, influences wide-ranging aspects of gene expression in diverse physiological and pathophysiological processes (1–3). The METTL3-METTL14 methyltransferase complex installs m^{6}A methylation on mRNA in a common DRACH (D = A, G, or U; R= A or G; H= A, C, or U) sequence motif, but only a fraction of DRACH sequences (~5%) in a subset of cellular transcripts are selected for methylation (4). Additionally, m^{6}A exhibits a marked regional bias in its transcriptomic distribution, being strongly enriched in unusually long internal exons and near stop codons (5, 6). Despite the central importance of specific m^{6}A deposition in m^{6}A-mediated gene regulation, the mechanistic basis for m^{6}A specificity has remained poorly understood.

In this study, we discover the existence of prevalent regulatory mechanisms that restrict m^{6}A methylation to specific transcript regions through targeted suppression of m^{6}A in unmethylated regions. We find that pre-mRNA splicing selectively suppresses m^{6}A deposition in average-length internal exons, but not in longer exons. We identify Exon Junction Complexes (EJC) as major m^{6}A suppressors that mediate this effect and control several key characteristics of global m^{6}A specificity. EJC depletion results in pervasive aberrant methylation of mRNAs and m^{6}A-mediated transcript destabilization. EJCs, together with interacting proteins, package and protect long stretches of proximal RNA from cellular methylation deposition, which may represent a general mechanism by which exon architecture and EJC positioning determine local mRNA accessibility to regulatory machineries.

Massively Parallel Assay for m^{6}A

The extent to which global m^{6}A specificity is controlled has important implications for m^{6}A regulation but is poorly understood (4). We approached this problem by asking: is the local sequence surrounding an m^{6}A methylated site, when uncoupled from its endogenous context, sufficient to specify methylation at that site? Conversely, is the local sequence surrounding an unmethylated DRACH site, when uncoupled from its endogenous context, sufficient to prevent methylation at that site?

To assess this systematically on an epitranscriptome-wide scale, we developed a Massively Parallel Reporter Assay (MPRA) that enables high-throughput assessment of the m^{6}A methylation status of thousands of designed sequences, which we termed Massively Parallel assay for m^{6}A (MPm^{6}A) (Fig. 1A and fig. S1A). In the MPm^{6}A workflow, thousands of endogenously methylated m^{6}A sites and endogenously unmethylated DRACH sites, with 102 nucleotides of sequence surrounding each site, were synthesized and cloned into the 3′UTR of a plasmid-based intronless GFP transgene. For each sequence, we also

designed a corresponding negative control sequence in which all DRACH motifs were mutated to prevent methylation. The sequences were expressed and then m^{6}A methylated through transfection into cells, or, when specified, through in vitro transcription and in vitro m^{6}A methylation. The methylation status of each individual sequence was assessed by its enrichment following m^{6}A-immunoprecipitation (IP) of mRNA. We selected 6,897 HeLa m^{6}A sites and 3,058 unmethylated DRACH sites to assay in HeLa cells and validated the assay's precision and accuracy (fig. S1, B to D).

Widespread mRNA m^{6}A suppression controls m^{6}A epitranscriptome specificity

When we compared the methylation levels of the endogenously methylated sequences to their negative control sequences, we found that 92.8% of the sequences exhibited significant methylation in this reporter assay (Fig. 1B and fig. S1E), indicating that most endogenously methylated sites do not strictly require their larger surrounding native context for methylation. Unexpectedly, 90.2% of endogenously unmethylated sequences also exhibited significant methylation (Fig. 1B and fig. S1E). The MPm^{6}A enrichment scores of the endogenously unmethylated group were similar to the endogenously methylated group, despite their diverging endogenous methylation states (Fig. 1B). We observed similar results when the sequences were in vitro transcribed and methylated with recombinant METTL3-METTL14 (fig. S2). Thus, thousands of endogenously unmethylated DRACH sites became methylated when they were uncoupled from their endogenous contexts and expressed in an artificial reporter context. We term these identified sites "suppressed m^{6}A sites". We validated these results for three selected sequences (fig. S1F), and confirmed that methylation was not notably influenced by the CMV promoter of the MPm^{6}A plasmid (fig. S3). We observed similar results when sequences were expressed within CDS or 5′UTR, though m^{6}A enrichment was significantly lower for many sequences when placed in the CDS or 5′UTR versus in the 3′UTR (fig. S4, A to D). This suggests that 5′ regions are generally less conducive for m^{6}A methylation than 3′ regions (fig. S4E). Collectively, these results reveal the existence of thousands of suppressed m^{6}A sites that are silenced by unknown mechanisms.

We noted that suppressed m^{6}A sites were enriched in the CDS and 3′UTR and were depleted near the stop codon, which is the reverse of endogenous m^{6}A site enrichment (Fig. 1C and fig. S5A). Further, suppressed m^{6}A sites in internal exons reside within much shorter exons (median = 167 nt) than endogenous m^{6}A sites (median = 915 nt) (Fig. 1D). These observations suggest that endogenous m^{6}A enrichment in long internal exons may be a consequence of suppression of m^{6}A deposition in shorter internal exons, which comprise most exons (90% of internal exons are < 246 nt) (Fig. 1E). 942 genes containing suppressed m^{6}A sites did not contain any endogenous m^{6}A peaks on their transcripts (fig. S5B). Suppression of these sites appears to involve suppression of m^{6}A deposition rather than active demethylation, as binding sites for RBM15, a METTL3-METTL14 methyltransferase complex accessory subunit, were highly enriched near endogenous m^{6}A sites compared to suppressed m^{6}A sites (fig. S6A), while FTO and ALKBH5 binding sites were not significantly enriched near suppressed sites and exhibited little binding near suppressed sites

overall (fig. S6B). These results were unexpected as previous reports on m^{6}A specificity had mainly focused on activating mechanisms (7–12). Our MPm^{6}A assay suggests the existence of unknown m^{6}A "suppressors" that govern global m^{6}A specificity by suppressing m^{6}A deposition.

Pre-mRNA splicing suppresses m^{6}A methylation proximal to splice sites

We next examined the relative enrichment of binding sites for 120 RBPs near endogenously methylated versus suppressed m^{6}A sites to identify candidate suppressors (13). Several spliceosome components (*BUD13*, *SF3B4*, *EFTUD2*) were significantly enriched near suppressed sites, suggesting that splicing may suppress m^{6}A (fig. S6A). Because suppressed m^{6}A sites in CDS primarily reside within average-length internal exons, we hypothesized that the splicing of average-length internal exons may suppress m^{6}A methylation. To test this, we cloned a suppressed m^{6}A site from an average-length internal exon in the *CRY1* gene (fig. S7A) into a rabbit beta-globin minigene reporter (BG), as well as a version with the introns removed (BG Δi1,i2). First, we cloned the suppressed *CRY1* site and 50/51 nt of flanking sequence into the internal exon, or in the last exon of these constructs. Notably, the spliced construct strongly suppressed methylation of the sequence placed within the internal exon, but not within the last exon (fig. S7B). Removal of either intron (BG Δi1 CRY1 102, BG Δi2 CRY1 102) resulted in partial loss of suppression, indicating that splicing of both introns contributed to methylation suppression (Fig. 2A). Consistent with this notion, deletion of all splice sites also resulted in a decrease in m^{6}A suppression (fig. S7C). Cloning in 912 nt of the *CRY1* exonic sequence surrounding the suppressed site into the internal exon (BG CRY1 912), forming a long internal exon, resulted in a loss of suppression (Fig. 2A). We hypothesized that the suppression is dependent on the proximity of the m^{6}A site, located within the center of the exon, to splice sites. Expanding the length of the BG CRY1 102 internal exon by cloning in larger amounts of *CRY1* flanking exonic sequence resulted in a progressive loss of suppression, with a > 476 nt internal exon unable to suppress m^{6}A (Fig. 2B and fig. S7D). These results reveal a causal role for pre-mRNA splicing in m^{6}A regulation.

Exon junction complexes control m^{6}A epitranscriptome specificity

We next sought to understand the mechanism by which splicing suppresses m^{6}A deposition. Exon junction complexes (EJCs) are deposited by spliceosomes onto mRNA ~24 nt upstream of exon-exon junctions and plays multifaceted roles in gene expression regulation (14, 15). Notably, two recent studies reported that EJCs efficiently block splicing at proximal aberrant splice sites (16, 17). Additionally, EJCs, together with interacting serine and arginine-rich (SR) proteins, package and compact mRNA and can protect long stretches of proximal RNA from nuclease accessibility in vitro, and also block 5′ to 3′ exonuclease degradation in vivo (18, 19). We reasoned that suppressed m^{6}A sites within average-length internal exons are within relatively close proximity to both an upstream and downstream EJC. Conversely, m^{6}A sites within long internal exons and near stop codons (which generally reside in long last exons) can be hundreds of nucleotides away from the nearest EJC. We therefore hypothesized that EJCs could mediate the splice site-proximal suppression of m^{6}A we observed.

We knocked down (KD) the core EJC factor EIF4A3 in HeLa cells and assessed the effect on m^{6}A deposition transcriptome-wide using m^{6}A-MeRIP-seq. Notably, 24,350 regions were significantly hypermethylated upon *EIF4A3* KD, while 3,140 regions were hypomethylated (Fig. 3A). 39% of these hypermethylated regions exhibited a greater than 8-fold increase in m^{6}A enrichment compared to the non-targeting siRNA control. We knocked down RBM8A, another core EJC factor (20), and observed similar, though relatively milder, transcriptome-wide m^{6}A changes, with 14,034 significantly hypermethylated regions observed, of which 57% overlapped with hypermethylated regions observed in *EIF4A3* KD (Fig. 3A and fig. S8, A and B). The relatively milder m^{6}A changes upon *RBM8A* KD may result from relatively lower KD efficiency (table S1) or may indicate a stronger requirement of EIF4A3 for suppression. Concordant with these transcriptome-wide m^{6}A changes, using UHPLC-QQQ-MS/MS, we found that *EIF4A3* KD increased global levels of m^{6}A in polyadenylated RNA by two-fold, while *RBM8A* KD resulted in a ~25% increase (fig. S8C).

94% of hypermethylated regions from *EIF4A3* KD and 82% of hypermethylated regions from *RBM8A* KD did not overlap with m^{6}A peaks identified under the non-targeting siRNA control conditions, suggesting that these regions contain newly methylated suppressed m^{6}A sites (Fig. 3, A and B, and fig. S8D). Indeed, out of 1,024 CDS sequences identified by MPm^{6}A to contain suppressed m^{6}A sites, 46% become methylated upon *EIF4A3* and/or *RBM8A* KD, including the *CRY1* suppressed site (fig. S8E), with three selected suppressed sites validated (fig. S8, F and G) (21). Furthermore, *EIF4A3* KD substantially alleviated the previously observed m^{6}A suppression within the internal exon of BG CRY1 102 (fig. S8H).

Consistent with our model, newly methylated and hypermethylated regions were highly enriched in average-length internal exons within CDSs (Fig. 3, C to E, and fig. S8, I and J), with transcriptome-wide increases in m^{6}A enrichment in exon junction-proximal regions observed (fig.S9, A and B) upon *EIF4A3* or *RBM8A* KD. *EIF4A3* KD disrupted m^{6}A epitranscriptome specificity globally, resulting in substantial loss of enrichment of m^{6}A peaks in long internal exons and increased density of m^{6}A in the CDS relative to the stop codon (Fig. 3, D and E). It was previously reported that the peak of m^{6}A density near stop codons on metagene plots can be more precisely visualized as an increased enrichment 150 nt past the start of last exons (6). *EIF4A3* KD resulted in a global increase in m^{6}A enrichment < 150 nt past the start of last exons (fig. S9, A to C), indicating that EJC suppression of methylation proximal to last exon-exon junctions is responsible for the characteristic m^{6}A peak density near stop codons. While most transcripts exhibited hypermethylation and contained one or more endogenous m^{6}A peaks upon *EIF4A3* KD, over a thousand transcripts that ordinarily lack endogenous m^{6}A peaks also gained aberrant m^{6}A methylation upon *EIF4A3* KD, revealing a major role for EJCs in suppressing m^{6}A deposition on the subset of transcripts that ordinarily are not subject to m^{6}A regulation (fig. S9, D to F).

The widespread suppression of m^{6}A by the EJCs also implies that many m^{6}A are deposited following splicing, which we confirmed using pulse-chase metabolic labeling experiments and UHPLC-QQQ-MS/MS (supplementary text and fig. S10). Two genes used in gene therapies for mucopolysaccharidosis type II and spinal muscular atrophy, *IDS* and *SMN*,

contain EJC-suppressed m^{6}A sites in their mRNAs, respectively. As expected, when these mRNAs were expressed from cDNA constructs, and thus not bound by EJCs, they were significantly hypermethylated relative to the corresponding endogenous mRNAs (fig. S11). Further, lncRNAs that contain three or more exons globally exhibit EJC suppression of m^{6}A in internal exons, while those with two or less do not (supplementary text and fig. S12). We depleted *EIF4A3* with a different siRNA in HeLa cells, and knocked down *EIF4A3* in HEK293T cells as well as in a glioblastoma cancer cell line (U87) that is sensitive to EIF4A3 perturbation (22), and observed similar transcriptome-wide m^{6}A changes in each case (figs. S13 to S15). Altogether, our results indicate that spliceosomes widely suppress m^{6}A methylation via deposition of EJCs that protect proximal RNA from methylation.

EJCs regulate mRNA stability by suppressing m^{6}A methylation

m^{6}A is known to mainly accelerate mRNA degradation via the reader protein YTHDF2 (23, 24). Accordingly, we observed globally reduced mRNA half-life of hypermethylated transcripts (~90%) upon *EIF4A3* KD (Fig. 4, A and B). Consistently, we observed generally increased YTHDF2 binding on hypermethylated mRNAs, accompanied with the decreased mRNA half-life (Fig. 4C). *YTHDF2* KD could rescue accelerated degradation of YTHDF2 target transcripts upon *EIF4A3* KD (fig. S16). Further, the density of EJC-loading on transcripts (estimated by the number of exons within CDS regions per 1 kb) correlated with transcriptome-wide mRNA stability (fig. S17). Higher EJC density on transcripts tended to correlate with reduced m^{6}A methylation and higher mRNA stability, and the strength of this correlation was diminished by *Mettl3* KO (fig. S17, A and B).

We also found that *METTL3* depletion could generally reduce the expression level changes of hypermethylated genes upon EJC depletion in HeLa cells (supplementary text and fig. S18), indicating that these EJC-dependent gene expression changes are at least in part mediated by m^{6}A methylation.

While the vast majority of hypermethylated transcripts were destabilized by *EIF4A3* KD, a small subset of hypermethylated transcripts were stabilized (Fig. 4A). One example is p53, which mediates neurodevelopmental defects in mouse models of EJC haploinsufficiency (25). The *TP53* transcript was hypermethylated but also up-regulated upon *EIF4A3* KD. Mechanistically, we observed increased binding to *TP53* mRNA by IGF2BP proteins, which are known to stabilize methylated transcripts (supplementary text and fig. S19). In summary, while the predominant effect of EJC-mediated m^{6}A suppression is to stabilize mRNAs by preventing the YTHDF2-mediated decay, in a minority of instances hypermethylated transcripts can be stabilized by other mechanisms, such as binding by IGF2BPs (26).

Consistent with a general role for m^{6}A in promoting translation (12, 27), *EIF4A3* KD led to slightly increased translation efficiency of hypermethylated transcripts, with more highly hypermethylated transcripts exhibiting greater increases in translation efficiency (fig. S20), although the impact was modest relative to the effects observed on mRNA stability.

Differential m^{6}A methylation across tissues and species through EJC suppression

Our model suggests that the cellular EJC levels may impact global mRNA m^{6}A deposition in different tissues. Indeed, we observed a negative correlation between *EIF4A3* expression level and global mRNA m^{6}A modification level in 25 different human tissues with available transcriptome-wide m^{6}A profiles (fig. S21A) (28). We examined the top 10% of genes with the strongest correlations and found that the majority (> 70%) exhibited a negative correlation between m^{6}A and *EIF4A3* levels in different tissues. Further, m^{6}A levels of these genes also negatively correlated with their transcript abundances (fig. S21B). Similar trends were also observed in mouse tissues (fig. S21C). These results further support m^{6}A suppression by EJCs and subsequently mRNA stability regulation in mammalian tissues.

Notably, we observed the lowest *EIF4A3* expressions in brain tissues, which exhibited the highest overall mRNA m^{6}A levels (fig. S21A). We further compared the methylome of the human cerebellum (lowest *EIF4A3* level and highest overall mRNA m^{6}A) with that of the heart (higher *EIF4A3* level and lower overall mRNA m^{6}A). Regions that are hypermethylated in the cerebellum (compared to heart) reside within short internal exons (fig. S21D), suggesting reduced m^{6}A suppression due to low *EIF4A3* expression in cerebellum. This association between high m^{6}A level and low *EIF4A3* expression in cerebellum was attenuated upon depletion of *METTL3* (fig. S21C). These observations further indicate the widespread suppression by EJCs contributes to tissue-specific m^{6}A deposition. We also found that a subset of EJC-suppressed m^{6}A sites physiologically escape suppression in certain tissues via methylation of alternative transcript isoforms. These isoforms contain longer exons and thus altered EJC positioning; methylation of these isoforms generates tissue-specific m^{6}A patterns (supplementary text and fig. S22).

Lastly, the effect of exon-intron architecture on mRNA stability may have co-evolved with YTHDF2 in vertebrates. The strong correlation between EJC loading, represented by the number of exons, and mRNA level across tissues is maintained across humans, mice, and zebrafish, but not fly and worm, which lack YTHDF2 orthologs (supplementary text and fig. S23).

EJCs and peripheral EJC factor RNPS1 protect exon junction-proximal RNA regions from aberrant mRNA processing

We did not observe interactions between the methyltransferase complex and EJC complexes (fig. S24), suggesting that steric hindrance from EJCs, rather than a specific inhibitory interaction, accounts for methylation suppression. Nuclear EJCs bound with the peripheral EJC factor RNPS1 multimerize and associate with wide variety of SR and SR-like proteins to package and compact mRNA into higher-order, megadalton-scale mRNPs that ensheathe proximal RNA well beyond the canonical EJC deposition sites (18, 29, 30). Tens to hundreds of nucleotides of proximal RNA could be protected by this mega-complex from nuclease digestion due to this packaging (18, 31). To examine whether the mRNA packaging function of the EJC-mediates suppression of proximal methylation, we isolated

EJCs/EJC-bound RNA from cellular extracts, digested away physically accessible RNA with in vitro nuclease treatment, and then measured m^{6}A levels on the EJC-protected RNA footprints (fig. S25, A and B). EJC-protected footprints were strongly depleted of m^{6}A, indicating that these inaccessible RNA regions are largely protected from m^{6}A deposition within cells (fig. S25C). EJCs also protected these footprints from in vitro methylation by recombinant METTL3-METTL14 (fig. S25D). This was not due to general inhibition of methyltransferase activity or lack of methylatable sites on the EJC footprints, as free, unmethylated RNA spiked into the methylation reaction as well as deproteinized footprints were both robustly methylated (fig. S25, D and E). Therefore, EJCs suppress local m^{6}A deposition by packaging proximal RNA.

We next asked whether the peripheral EJC factor RNPS1, which associates with high molecular weight EJCs in these highly packaged mRNP structures (29), plays a role. *RNPS1* knockdown led to substantial transcript m^{6}A hypermethylation within average-length internal exons and CDS regions (Fig. 5, A to C, and fig. S26, A to C). We detected fewer hypermethylated regions overall compared to depletion of the core EJC factors; however, we did observe high overlap (45%) between si*RNPS1* hypermethylated regions and si*EIF4A3*/si*RBM8A* hypermethylated regions (Fig. 5C and fig. S26C). In contrast, depletion of UPF1, a central NMD factor that interacts with the EJC in the cytoplasm, did not result in m^{6}A changes similar to those of the core EJC (fig. S27).

The ability of EJCs to protect proximal RNA regions from methylation resembles the recently characterized EJC- and RNPS1-mediated suppression of proximal aberrant splice sites and recursive splicing (16, 17). Transcriptome-wide, EJC-suppressed splice sites significantly colocalize with EJC-suppressed m^{6}A sites (supplementary text; Fig. 4, C to E; fig. S26, D to F; and table S2). Altogether, these results suggest that RNPS1-associated EJCs suppress both local cellular m^{6}A methylation and splicing through packaging of proximal RNA and point to exon architecture as an important determinant of local RNA accessibility to regulatory machineries. Additionally, beyond components of the m^{6}A methyltransferase complex, a number of other RBPs also exhibit preferential binding at long internal exons, suggesting that EJCs may regulate mRNA accessibility to a broader range of mRNA regulators in addition to the splicing and m^{6}A methylation machineries through their mRNA packaging function (supplementary text and fig. S28).

Discussion

Previously identified m^{6}A effector proteins fall broadly into three categories according to their activities: “writers”, which catalyze m^{6}A methylation, “readers”, which preferentially bind m^{6}A, and “erasers”, which reverse m^{6}A methylation. Here we establish the EJCs as a member of a new class of m^{6}A regulators: “suppressors”, which broadly suppress the deposition of m^{6}A (fig. S28). EJCs appear to be a major regulator of m^{6}A deposition that mediate multiple key aspects of global m^{6}A epitranscriptome specificity, including enrichment of m^{6}A in long internal exons, depletion of m^{6}A in CDSs and enrichment of m^{6}A in last exons near stop codons, and methylation selectivity for transcripts possessing long internal exons. This mechanism may also explain the high abundance of m^{6}A on certain non-coding RNAs, such as LINE-1 elements that are generally unspliced and thus

not bound by the EJCs (32, 33). Further, our systematic analysis of m^{6}A determinants using MPm^{6}A may suggest the existence of additional m^{6}A suppressing pathways, including m^{6}A suppression within the CDS, as *EIF4A3* KD does not appear to completely restore methylation to unspliced levels (supplementary text, fig. S8H, and fig. S29).

Our results point to exon length within transcripts as a functionally relevant element for post-transcriptional gene expression regulation. Mammalian EJCs stably bind the vast majority of pre-translational mRNAs in the transcriptome at closely spaced intervals. Long internal exons and terminal exons, which usually encode UTRs, are notably free of EJCs. This widespread binding, in conjunction with their mRNA packaging function, appears to uniquely position EJCs to broadly determine mRNA accessibility to regulatory machineries, such as the m^{6}A methylation and splicing machineries (fig. S30). Our work has relevance for the use of cDNA expression constructs in research studies and gene therapies, as loss of endogenous mRNA exon architecture and EJC protection results in m^{6}A hypermethylation (fig. S11), which could modulate gene expression outcome. Finally, our study also suggests that exon length and architecture co-evolved with mRNA processing steps as an additional regulatory layer of gene expression.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

ACKNOWLEDGMENTS

We thank Tao Pan, Erin Adams, Marcus Clark, Marcelo Nobrega, Jonathan Staley, and Amelia Joslin for comments and suggestions. We thank the Genomics Facility of the University of Chicago and the University of Chicago Comprehensive Cancer Center DNA Sequencing and Genotyping Facility for assistance with sequencing.

Funding:

National Institutes of Health HG008935 (C.H.); National Institutes of Health grant T32 HD007009 (P.C.H); National Institutes of Health grant F32 CA221007 (B.T.H); C.H. is an investigator of the Howard Hughes Medical Institute.

Data and materials availability:

Raw and processed data can be found at NCBI GEO accession GSE162199. Custom scripts available on Zenodo (34). All other data are available in the manuscript or the supplementary materials.

REFERENCES AND NOTES

1. Frye M, Harada BT, Behm M, He C, RNA modifications modulate gene expression during development. Science 361, 1346–1349 (2018).doi:10.1126/science.aau1646 [PubMed: 30262497]
2. Gilbert WV, Bell TA, Schaening C, Messenger RNA modifications: Form, distribution, and function. Science 352, 1408–1412 (2016).doi:10.1126/science.aad8711 [PubMed: 27313037]
3. Roundtree IA, Evans ME, Pan T, He C, Dynamic RNA modifications in gene expression regulation. Cell 169, 1187–1200 (2017).doi:10.1016/j.cell.2017.05.045 [PubMed: 28622506]
4. He PC, He C, m^{6}A RNA methylation: From mechanisms to therapeutic potential. EMBO J. 40, e105977 (2021).doi:10.15252/embj.2020105977 [PubMed: 33470439]

5. 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, Topology of the human and mouse m^6A RNA methylomes revealed by m^6A-seq. Nature 485, 201–206 (2012).doi:10.1038/nature11112 [PubMed: 22575960]
6. Ke S, Alemu EA, Mertens C, Gantman EC, Fak JJ, Mele A, Haripal B, Zucker-Scharff I, Moore MJ, Park CY, Vågbø CB, Kusśnierczyk A, Klungland A, Darnell JE Jr., R. B. Darnell, A majority of m^6A residues are in the last exons, allowing the potential for 3′ UTR regulation. Genes Dev. 29, 2037–2053 (2015).doi:10.1101/gad.269415.115 [PubMed: 26404942]
7. Barbieri I, Tzelepis K, Pandolfini L, Shi J, Millán-Zambrano G, Robson SC, Aspris D, Migliori V, Bannister AJ, Han N, De Braekeleer E, Ponstingl H, Hendrick A, Vakoc CR, Vassiliou GS, Kouzarides T, Promoter-bound METTL3 maintains myeloid leukaemia by m^6A-dependent translation control. Nature 552, 126–131 (2017).doi:10.1038/nature24678 [PubMed: 29186125]
8. Bertero A, Brown S, Madrigal P, Osnato A, Ortmann D, Yiangou L, Kadiwala J, Hubner NC, de Los Mozos IR, Sadée C, Lenaerts A-S, Nakanoh S, Grandy R, Farnell E, Ule J, Stunnenberg HG, Mendjan S, Vallier L, The SMAD2/3 interactome reveals that TGF_β controls m^6A mRNA methylation in pluripotency. Nature 555, 256–259 (2018).doi:10.1038/nature25784 [PubMed: 29489750]
9. Fish L, Navickas A, Culbertson B, Xu Y, Nguyen HCB, Zhang S, Hochman M, Okimoto R, Dill BD, Molina H, Najafabadi HS, Alarcón C, Ruggero D, Goodarzi H, Nuclear TARBP2 drives oncogenic dysregulation of RNA splicing and decay. Mol. Cell 75, 967–981.e9 (2019).doi:10.1016/j.molcel.2019.06.001 [PubMed: 31300274]
10. Slobodin B, Han R, Calderone V, Vrielink JAFO, Loayza-Puch F, Elkon R, Agami R, Transcription impacts the efficiency of mRNA translation via co-transcriptional N^6-adenosine methylation. Cell 169, 326–337.e12 (2017).doi:10.1016/j.cell.2017.03.031 [PubMed: 28388414]
11. Huang H, Weng H, Zhou K, Wu T, Zhao BS, Sun M, Chen Z, Deng X, Xiao G, Auer F, Klemm L, Wu H, Zuo Z, Qin X, Dong Y, Zhou Y, Qin H, Tao S, Du J, Liu J, Lu Z, Yin H, Mesquita A, Yuan CL, Hu Y-C, Sun W, Su R, Dong L, Shen C, Li C, Qing Y, Jiang X, Wu X, Sun M, Guan J-L, Qu L, Wei M, Müschen M, Huang G, He C, Yang J, Chen J, Histone H3 trimethylation at lysine 36 guides m^6A RNA modification co-transcriptionally. Nature 567, 414–419 (2019).doi:10.1038/s41586-019-1016-7 [PubMed: 30867593]
12. Zhang Z, Luo K, Zou Z, Qiu M, Tian J, Sieh L, Shi H, Zou Y, Wang G, Morrison J, Zhu AC, Qiao M, Li Z, Stephens M, He X, He C, Genetic analyses support the contribution of mRNA N^6-methyladenosine (m^6A) modification to human disease heritability. Nat. Genet. 52, 939–949 (2020).doi:10.1038/s41588-020-0644-z [PubMed: 32601472]
13. Van Nostrand EL, Freese P, Pratt GA, Wang X, Wei X, Xiao R, Blue SM, Chen J-Y, Cody NAL, Dominguez D, Olson S, Sundararaman B, Zhan L, Bazile C, Bouvrette LPB, Bergalet J, Duff MO, Garcia KE, Gelboin-Burkhart C, Hochman M, Lambert NJ, Li H, McGurk MP, Nguyen TB, Palden T, Rabano I, Sathe S, Stanton R, Su A, Wang R, Yee BA, Zhou B, Louie AL, Aigner S, Fu X-D, Lécuyer E, Burge CB, Graveley BR, Yeo GW, A large-scale binding and functional map of human RNA-binding proteins. Nature 583, 711–719 (2020).doi:10.1038/s41586-020-2077-3 [PubMed: 32728246]
14. Le Hir H, Saulière J, Wang Z, The exon junction complex as a node of post-transcriptional networks. Nat. Rev. Mol. Cell Biol. 17, 41–54 (2016).doi:10.1038/nrm.2015.7 [PubMed: 26670016]
15. Boehm V, Gehring NH, Exon junction complexes: Supervising the gene expression assembly line. Trends Genet. 32, 724–735 (2016).doi:10.1016/j.tig.2016.09.003 [PubMed: 27667727]
16. Boehm V, Britto-Borges T, Steckelberg A-L, Singh KK, Gerbracht JV, Gueney E, Blazquez L, Altmüller J, Dieterich C, Gehring NH, Exon junction complexes suppress spurious splice sites to safeguard transcriptome integrity. Mol. Cell 72, 482–495.e7 (2018).doi:10.1016/j.molcel.2018.08.030 [PubMed: 30388410]
17. Blazquez L, Emmett W, Faraway R, Pineda JMB, Bajew S, Gohr A, Haberman N, Sibley CR, Bradley RK, Irimia M, Ule J, Exon junction complex shapes the transcriptome by repressing recursive splicing. Mol. Cell 72, 496–509.e9 (2018).doi:10.1016/j.molcel.2018.09.033 [PubMed: 30388411]

18. Singh G, Kucukural A, Cenik C, Leszyk JD, Shaffer SA, Weng Z, Moore MJ, The cellular EJC interactome reveals higher-order mRNP structure and an EJC-SR protein nexus. Cell 151, 750–764 (2012).doi:10.1016/j.cell.2012.10.007 [PubMed: 23084401]
19. Lee W-C, Hou B-H, Hou C-Y, Tsao S-M, Kao P, Chen H-M, Widespread exon junction complex footprints in the RNA degradome mark mRNA degradation before steady state translation. Plant Cell 32, 904–922 (2020).doi:10.1105/tpc.19.00666 [PubMed: 31988264]
20. Tange TØ, Shibuya T, Jurica MS, Moore MJ, Biochemical analysis of the EJC reveals two new factors and a stable tetrameric protein core. RNA 11, 1869–1883 (2005).doi:10.1261/rna.2155905 [PubMed: 16314458]
21. Xiao Y, Wang Y, Tang Q, Wei L, Zhang X, Jia G, An elongation- and ligation-based qPCR amplification method for the radiolabeling-free detection of locus-specific N^6-methyladenosine modification. Angew. Chem. Int. Ed. 57, 15995–16000 (2018).doi:10.1002/anie.201807942
22. Tang W, Wang D, Shao L, Liu X, Zheng J, Xue Y, Ruan X, Yang C, Liu L, Ma J, Li Z, Liu Y, LINC00680 and TTN-AS1 stabilized by EIF4A3 promoted malignant biological behaviors of glioblastoma cells. Mol. Ther. Nucleic Acids 19, 905–921 (2020).doi:10.1016/j.omtn.2019.10.043 [PubMed: 32000032]
23. Wang X, Lu Z, Gomez A, Hon GC, Yue Y, Han D, Fu Y, Parisien M, Dai Q, Jia G, Ren B, Pan T, He C, N^6-methyladenosine-dependent regulation of messenger RNA stability. Nature 505, 117–120 (2014).doi:10.1038/nature12730 [PubMed: 24284625]
24. Du H, Zhao Y, He J, Zhang Y, Xi H, Liu M, Ma J, Wu L, YTHDF2 destabilizes m^6A-containing RNA through direct recruitment of the CCR4-NOT deadenylase complex. Nat. Commun. 7, 12626 (2016).doi:10.1038/ncomms12626 [PubMed: 27558897]
25. Mao H, McMahon JJ, Tsai Y-H, Wang Z, Silver DL, Haploinsufficiency for Core Exon Junction Complex Components Disrupts Embryonic Neurogenesis and Causes p53-Mediated Microcephaly. PLOS Genet. 12, e1006282 (2016).doi:10.1371/journal.pgen.1006282 [PubMed: 27618312]
26. Huang H, Weng H, Sun W, Qin X, Shi H, Wu H, Zhao BS, Mesquita A, Liu C, Yuan CL, Hu Y-C, Hüttelmaier S, Skibbe JR, Su R, Deng X, Dong L, Sun M, Li C, Nachtergaele S, Wang Y, Hu C, Ferchen K, Greis KD, Jiang X, Wei M, Qu L, Guan J-L, He C, Yang J, Chen J, Recognition of RNA N^6-methyladenosine by IGF2BP proteins enhances mRNA stability and translation. Nat. Cell Biol. 20, 285–295 (2018).doi:10.1038/s41556-018-0045-z [PubMed: 29476152]
27. Wang X, Zhao BS, Roundtree IA, Lu Z, Han D, Ma H, Weng X, Chen K, Shi H, He C, N^6-methyladenosine modulates messenger RNA translation efficiency. Cell 161, 1388–1399 (2015).doi:10.1016/j.cell.2015.05.014 [PubMed: 26046440]
28. Liu J, Li K, Cai J, Zhang M, Zhang X, Xiong X, Meng H, Xu X, Huang Z, Peng J, Fan J, Yi C, Landscape and regulation of m^6A and m^6A_m methylome across human and mouse tissues. Mol. Cell 77, 426–440.e6 (2020).doi:10.1016/j.molcel.2019.09.032 [PubMed: 31676230]
29. Mabin JW, Woodward LA, Patton RD, Yi Z, Jia M, Wysocki VH, Bundschuh R, Singh G, The exon junction complex undergoes a compositional switch that alters mRNP structure and nonsense-mediated mRNA decay activity. Cell Rep. 25, 2431–2446.e7 (2018).doi:10.1016/j.celrep.2018.11.046 [PubMed: 30466796]
30. Metkar M, Ozadam H, Lajoie BR, Imakaev M, Mirny LA, Dekker J, Moore MJ, Higher-order organization principles of pre-translational mRNPs. Mol. Cell 72, 715–726.e3 (2018).doi:10.1016/j.molcel.2018.09.012 [PubMed: 30415953]
31. Le Hir H, Izaurralde E, Maquat LE, Moore MJ, The spliceosome deposits multiple proteins 20–24 nucleotides upstream of mRNA exon-exon junctions. EMBO J. 19, 6860–6869 (2000).doi:10.1093/emboj/19.24.6860 [PubMed: 11118221]
32. Liu J, Dou X, Chen C, Chen C, Liu C, Xu MM, Zhao S, Shen B, Gao Y, Han D, He C, N^6-methyladenosine of chromosome-associated regulatory RNA regulates chromatin state and transcription. Science 367, 580–586 (2020).doi:10.1126/science.aay6018 [PubMed: 31949099]
33. Wei J, Yu X, Yang L, Liu X, Gao B, Huang B, Dou X, Liu J, Zou Z, Cui X-L, Zhang L-S, Zhao X, Liu Q, He PC, Sepich-Poore C, Zhong N, Liu W, Li Y, Kou X, Zhao Y, Wu Y, Cheng X, Chen C, An Y, Dong X, Wang H, Shu Q, Hao Z, Duan T, He Y-Y, Li X, Gao S, Gao Y, He C, FTO mediates LINE1 m^6A demethylation and chromatin regulation in mESCs and mouse development. Science 376, 968–973 (2022).doi:10.1126/science.abe9582 [PubMed: 35511947]

34. He PC, Dou X, Custom scripts associated with “Exon architecture controls mRNA m6A suppression and gene expression”, Zenodo (2023); 10.5281/zenodo.7541415.
35. Uphoff CC, Drexler HG, in Cancer Cell Culture: Methods and Protocols, Langdon SP, Ed., vol. 88 of Methods in Molecular Medicine (Humana Press, 2004), pp. 319–326. [PubMed: 14634244]
36. Yue Y, Liu J, Cui X, Cao J, Luo G, Zhang Z, Cheng T, Gao M, Shu X, Ma H, Wang F, Wang X, Shen B, Wang Y, Feng X, He C, Liu J, VIRMA mediates preferential m^6A mRNA methylation in 3’UTR and near stop codon and associates with alternative polyadenylation. Cell Discov. 4, 10 (2018).doi:10.1038/s41421-018-0019-0 [PubMed: 29507755]
37. Zhou Y, Zeng P, Li Y-H, Zhang Z, Cui Q, SRAMP: Prediction of mammalian N^6-methyladenosine (m^6A) sites based on sequence-derived features. Nucleic Acids Res. 44, e91 (2016).doi:10.1093/nar/gkw104 [PubMed: 26896799]
38. Ernst J, Kheradpour P, Mikkelsen TS, Shoresh N, Ward LD, Epstein CB, Zhang X, Wang L, Issner R, Coyne M, Ku M, Durham T, Kellis M, Bernstein BE, Mapping and analysis of chromatin state dynamics in nine human cell types. Nature 473, 43–49 (2011).doi:10.1038/nature09906 [PubMed: 21441907]
39. Kim D, Langmead B, Salzberg SL, HISAT: A fast spliced aligner with low memory requirements. Nat. Methods 12, 357–360 (2015).doi:10.1038/nmeth.3317 [PubMed: 25751142]
40. Zhang Z, Zhan Q, Eckert M, Zhu A, Chryplewicz A, De Jesus DF, Ren D, Kulkarni RN, Lengyel E, He C, Chen M, RADAR: Differential analysis of MeRIP-seq data with a random effect model. Genome Biol. 20, 294 (2019).doi:10.1186/s13059-019-1915-9 [PubMed: 31870409]
41. Yu G, Wang L-G, Han Y, He Q-Y, clusterProfiler: An R package for comparing biological themes among gene clusters. OMICS 16, 284–287 (2012).doi:10.1089/omi.2011.0118 [PubMed: 22455463]
42. Kent WJ, Sugnet CW, Furey TS, Roskin KM, Pringle TH, Zahler AM, Haussler D, The human genome browser at UCSC. Genome Res. 12, 996–1006 (2002).doi:10.1101/gr.229102 [PubMed: 12045153]
43. Bolger AM, Lohse M, Usadel B, Trimmomatic: A flexible trimmer for Illumina sequence data. Bioinformatics 30, 2114–2120 (2014).doi:10.1093/bioinformatics/btu170 [PubMed: 24695404]
44. Anders S, Pyl PT, Huber W, HTSeq—A Python framework to work with high-throughput sequencing data. Bioinformatics 31, 166–169 (2015).doi:10.1093/bioinformatics/btu638 [PubMed: 25260700]
45. Langmead B, Trapnell C, Pop M, Salzberg SL, Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biol. 10, R25 (2009).doi:10.1186/gb-2009-10-3-r25 [PubMed: 19261174]
46. Love MI, Huber W, Anders S, Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 15, 550 (2014).doi:10.1186/s13059-014-0550-8 [PubMed: 25516281]
47. Martin M, Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet. J. 17, 10 (2011).doi:10.14806/ej.17.1.200
48. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M, Gingeras TR, STAR: Ultrafast universal RNA-seq aligner. Bioinformatics 29, 15–21 (2013).doi:10.1093/bioinformatics/bts635 [PubMed: 23104886]
49. Liao Y, Smyth GK, Shi W, featureCounts: An efficient general purpose program for assigning sequence reads to genomic features. Bioinformatics 30, 923–930 (2014).doi:10.1093/bioinformatics/btt656 [PubMed: 24227677]
50. Ingolia NT, Ghaemmaghami S, Newman JRS, Weissman JS, Genome-wide analysis in vivo of translation with nucleotide resolution using ribosome profiling. Science 324, 218–223 (2009).doi:10.1126/science.1168978 [PubMed: 19213877]
51. Buschmann T, Bystrykh LV, Levenshtein error-correcting barcodes for multiplexed DNA sequencing. BMC Bioinformatics 14, 272 (2013).doi:10.1186/1471-2105-14-272 [PubMed: 24021088]
52. Agarwal V, Bell GW, Nam J-W, Bartel DP, Predicting effective microRNA target sites in mammalian mRNAs. eLife 4, e05005 (2015).doi:10.7554/eLife.05005 [PubMed: 26267216]

53. Ke S, Pandya-Jones A, Saito Y, Fak JJ, Vågbø CB, Geula S, Hanna JH, Black DL, Darnell JE Jr., R. B. Darnell, m^{6}A mRNA modifications are deposited in nascent pre-mRNA and are not required for splicing but do specify cytoplasmic turnover. Genes Dev. 31, 990–1006 (2017).doi:10.1101/gad.301036.117 [PubMed: 28637692]
54. Yan F, Al-Kali A, Zhang Z, Liu J, Pang J, Zhao N, He C, Litzow MR, Liu S, A dynamic N^6-methyladenosine methylome regulates intrinsic and acquired resistance to tyrosine kinase inhibitors. Cell Res. 28, 1062–1076 (2018).doi:10.1038/s41422-018-0097-4 [PubMed: 30297871]
55. Middleton R, Gao D, Thomas A, Singh B, Au A, Wong JJ-L, Bomane A, Cosson B, Eyras E, Rasko JEJ, Ritchie W, IRFinder: Assessing the impact of intron retention on mammalian gene expression. Genome Biol. 18, 51 (2017).doi:10.1186/s13059-017-1184-4 [PubMed: 28298237]
56. Herrmann CJ, Schmidt R, Kanitz A, Artimo P, Gruber AJ, Zavolan M, PolyASite 2.0: A consolidated atlas of polyadenylation sites from 3′ end sequencing. Nucleic Acids Res. 48, D174–D179 (2020).doi:10.1093/nar/gkz918 [PubMed: 31617559]
57. Baltz AG, 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, The mRNA-bound proteome and its global occupancy profile on protein-coding transcripts. Mol. Cell 46, 674–690 (2012).doi:10.1016/j.molcel.2012.05.021 [PubMed: 22681889]
58. Bartosovic M, Molares HC, Gregorova P, Hrossova D, Kudla G, Vanacova S, N6-methyladenosine demethylase FTO targets pre-mRNAs and regulates alternative splicing and 3′-end processing. Nucleic Acids Res. 45, 11356–11370 (2017).doi:10.1093/nar/gkx778 [PubMed: 28977517]
59. Busch A, Hertel KJ, HEXEvent: A database of Human EXon splicing Events. Nucleic Acids Res. 41, D118–D124 (2013).doi:10.1093/nar/gks969 [PubMed: 23118488]
60. Bray NL, Pimentel H, Melsted P, Pachter L, Near-optimal probabilistic RNA-seq quantification. Nat. Biotechnol. 34, 525–527 (2016).doi:10.1038/nbt.3519 [PubMed: 27043002]
61. Louloupi A, Ntini E, Conrad T, Ørom UAV, Transient N-6-methyladenosine transcriptome sequencing reveals a regulatory role of m6A in splicing efficiency. Cell Rep. 23, 3429–3437 (2018).doi:10.1016/j.celrep.2018.05.077 [PubMed: 29924987]
62. Zhou KI, Shi H, Lyu R, Wylder AC, Matuszek Ż, Pan JN, He C, Parisien M, Pan T, Regulation of co-transcriptional pre-mRNA splicing by m^{6}A through the low-complexity protein hnRNPG. Mol. Cell 76, 70–81.e9 (2019).doi:10.1016/j.molcel.2019.07.005 [PubMed: 31445886]
63. Xiao W, Adhikari S, Dahal U, Chen Y-S, Hao Y-J, Sun B-F, Sun H-Y, Li A, Ping X-L, Lai W-Y, Wang X, Ma H-L, Huang C-M, Yang Y, Huang N, Jiang G-B, Wang H-L, Zhou Q, Wang X-J, Zhao Y-L, Yang Y-G, Nuclear m^{6}A reader YTHDC1 regulates mRNA splicing. Mol. Cell 61, 507–519 (2016).doi:10.1016/j.molcel.2016.01.012 [PubMed: 26876937]
64. Viphakone N, Sudbery I, Griffith L, Heath CG, Sims D, Wilson SA, Co-transcriptional loading of RNA export factors shapes the human transcriptome. Mol. Cell 75, 310–323.e8 (2019).doi:10.1016/j.molcel.2019.04.034 [PubMed: 31104896]
65. Neugebauer KM, Nascent RNA and the Coordination of Splicing with Transcription. Cold Spring Harb. Perspect. Biol. 11, a032227 (2019).doi:10.1101/cshperspect.a032227 [PubMed: 31371351]
66. Akhtar J, Kreim N, Marini F, Mohana G, Brüne D, Binder H, Roignant J-Y, Promoter-proximal pausing mediated by the exon junction complex regulates splicing. Nat. Commun. 10, 521 (2019).doi:10.1038/s41467-019-08381-0 [PubMed: 30705266]
67. Silver DL, Watkins-Chow DE, Schreck KC, Pierfelice TJ, Larson DM, Burnetti AJ, Liaw H-J, Myung K, Walsh CA, Gaiano N, Pavan WJ, The exon junction complex component *Magoh* controls brain size by regulating neural stem cell division. Nat. Neurosci. 13, 551–558 (2010).doi:10.1038/nn.2527 [PubMed: 20364144]
68. Wang Z, Murigneux V, Le Hir H, Transcriptome-wide modulation of splicing by the exon junction complex. Genome Biol. 15, 551 (2014).doi:10.1186/s13059-014-0551-7 [PubMed: 25476502]
69. Lu C-C, Lee C-C, Tseng C-T, Tarn W-Y, Y14 governs p53 expression and modulates DNA damage sensitivity. Sci. Rep. 7, 45558 (2017).doi:10.1038/srep45558 [PubMed: 28361991]
70. Sendinc E, Valle-Garcia D, Jiao A, Shi Y, Analysis of m6A RNA methylation in *Caenorhabditis elegans*. Cell Discov. 6, 47 (2020).doi:10.1038/s41421-020-00186-6 [PubMed: 32695436]

71. Kan L, Ott S, Joseph B, Park ES, Dai W, Kleiner RE, Claridge-Chang A, Lai EC, A neural m^6A/Ythdf pathway is required for learning and memory in *Drosophila*. Nat. Commun. 12, 1458 (2021).doi:10.1038/s41467-021-21537-1 [PubMed: 33674589]

72. Jia G, Fu Y, Zhao X, Dai Q, Zheng G, Yang Y, Yi C, Lindahl T, Pan T, Yang Y-G, He C, *N*6-methyladenosine in nuclear RNA is a major substrate of the obesity-associated FTO. Nat. Chem. Biol. 7, 885–887 (2011).doi:10.1038/nchembio.687 [PubMed: 22002720]

73. Wei J, Liu F, Lu Z, Fei Q, Ai Y, He PC, Shi H, Cui X, Su R, Klungland A, Jia G, Chen J, He C, Differential m^6A, m^6A_m, and m^1A Demethylation Mediated by FTO in the Cell Nucleus and Cytoplasm. Mol. Cell 71, 973–985.e5 (2018).doi:10.1016/j.molcel.2018.08.011 [PubMed: 30197295]

74. Su R, Dong L, Li C, Nachtergaele S, Wunderlich M, Qing Y, Deng X, Wang Y, Weng X, Hu C, Yu M, Skibbe J, Dai Q, Zou D, Wu T, Yu K, Weng H, Huang H, Ferchen K, Qin X, Zhang B, Qi J, Sasaki AT, Plas DR, Bradner JE, Wei M, Marcucci G, Jiang X, Mulloy JC, Jin J, He C, Chen J, R-2HG Exhibits Anti-tumor Activity by Targeting FTO/m^6A/MYC/CEBPA Signaling. Cell 172, 90–105.e23 (2018).doi:10.1016/j.cell.2017.11.031 [PubMed: 29249359]

75. Zheng G, Dahl JA, Niu Y, Fedorcsak P, Huang C-M, Li CJ, Vågbø CB, Shi Y, Wang W-L, Song S-H, Lu Z, Bosmans RPG, Dai Q, Hao Y-J, Yang X, Zhao W-M, Tong W-M, Wang X-J, Bogdan F, Furu K, Fu Y, Jia G, Zhao X, Liu J, Krokan HE, Klungland A, Yang Y-G, He C, ALKBH5 is a mammalian RNA demethylase that impacts RNA metabolism and mouse fertility. Mol. Cell 49, 18–29 (2013).doi:10.1016/j.molcel.2012.10.015 [PubMed: 23177736]

76. Zhang S, Zhao BS, Zhou A, Lin K, Zheng S, Lu Z, Chen Y, Sulman EP, Xie K, Bögler O, Majumder S, He C, Huang S, m^6A Demethylase ALKBH5 Maintains Tumorigenicity of Glioblastoma Stem-like Cells by Sustaining FOXM1 Expression and Cell Proliferation Program. Cancer Cell 31, 591–606.e6 (2017).doi:10.1016/j.ccell.2017.02.013 [PubMed: 28344040]

77. Molinie B, Wang J, Lim KS, Hillebrand R, Lu Z-X, Van Wittenberghe N, Howard BD, Daneshvar K, Mullen AC, Dedon P, Xing Y, Giallourakis CC, m6A-LAIC-seq reveals the census and complexity of the m6A epitranscriptome. Nat. Methods 13, 692–698 (2016).doi:10.1038/nmeth.3898 [PubMed: 27376769]

78. Hu L, Liu S, Peng Y, Ge R, Su R, Senevirathne C, Harada BT, Dai Q, Wei J, Zhang L, Hao Z, Luo L, Wang H, Wang Y, Luo M, Chen M, Chen J, He C, m^6A RNA modifications are measured at single-base resolution across the mammalian transcriptome. Nat. Biotechnol. 40, 1210–1219 (2022).doi:10.1038/s41587-022-01243-z [PubMed: 35288668]

79. Ha KCH, Blencowe BJ, Morris Q, QAPA: A new method for the systematic analysis of alternative polyadenylation from RNA-seq data. Genome Biol. 19, 45 (2018). doi:10.1186/s13059-018-1414-4 [PubMed: 29592814]

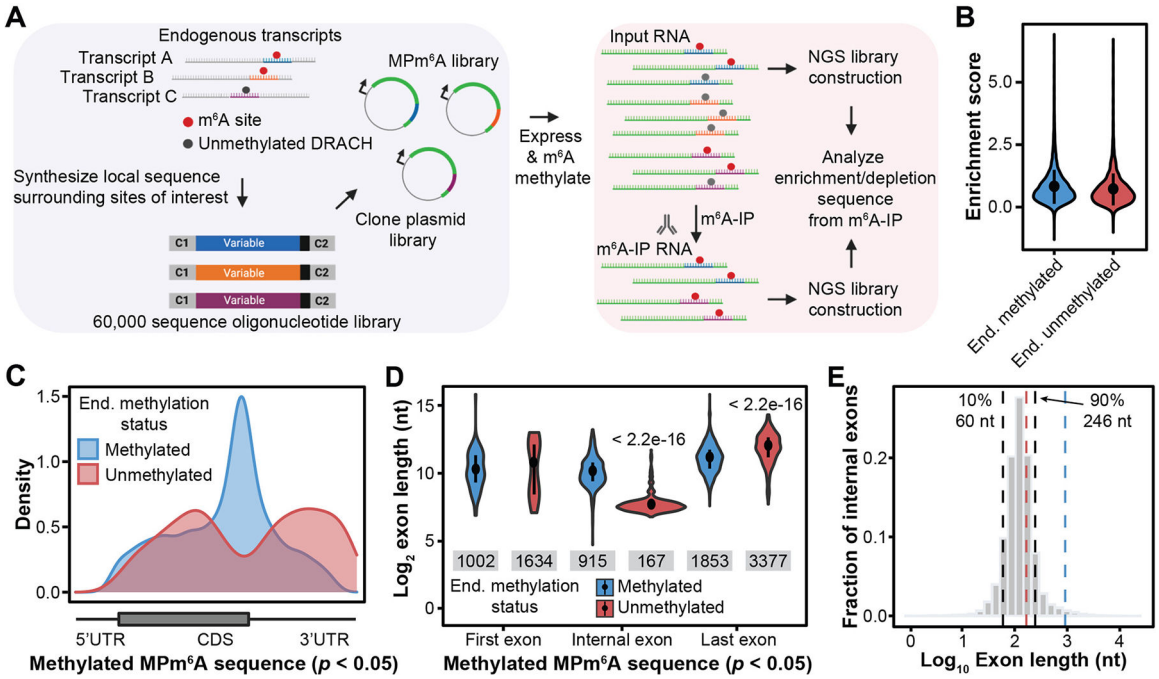

Fig. 1. MPm^{6}A reveals suppression of thousands of m^{6}A sites in unmethylated transcriptome regions.

(**A**) Schematic of the MPm^{6}A workflow. (**B**) MPm^{6}A enrichment scores (experimental IP/input – negative control IP/input) for endogenously methylated (n = 6,095) and unmethylated (n = 2,716) sequences, mean ± SD, four biological replicates. (**C**) Metagenes of endogenously methylated and unmethylated sequences that are significantly methylated in MPm^{6}A. (**D**) Exon lengths of endogenously methylated and unmethylated sequences that are significantly methylated in MPm^{6}A. Median, and IQR, Wilcoxon rank sum test. Sample size for each violin plot from left to right is: n = 175, n = 22, n = 696, n = 519, n = 3,539, and n = 1,328. (**E**) Distribution of internal exon lengths in the human genome. Black lines indicate 10th percentile (left, 60 nt) and 90th percentile (right, 246 nt). Blue and red lines indicate median internal exon length for MPm^{6}A endogenously methylated (915 nt) and unmethylated (167 nt) sequences, respectively.

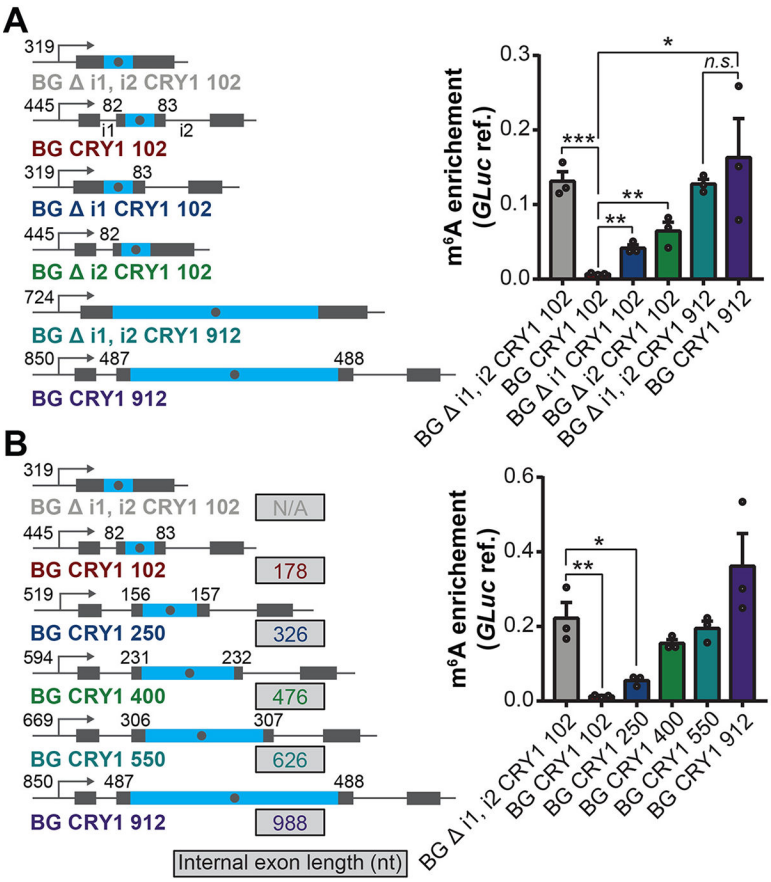

Fig. 2. Pre-mRNA splicing suppresses m^{6}A methylation in average-length exons.

(**A** and **B**) Left: schematic of specified BG CRY1 constructs. Blue regions indicate sequences derived from the *CRY1* endogenous sequence, gray regions indicate sequences derived from rabbit beta-globin (BG). Number following CRY1 refers to the number of nucleotides of exonic sequence surrounding the *CRY1* suppressed m^{6}A site in the *CRY1* endogenous mRNA that was cloned into the BG construct. Grey dot in the blue region denotes the suppressed m^{6}A site; the number at the left and right of the m^{6}A site shows the distance (nt) between the m^{6}A site and the 3′ and 5′ splice site, respectively; the number next to the TSS shows the distance (nt) between the m^{6}A site and the promoter. Δ denotes deletion of the specified intron(s). Details of each construct are described in the supplementary method. Right: m^{6}A enrichment at a *CRY1* suppressed m^{6}A site. Primers amplifying a 62 nt-fragment containing the *CRY1* suppressed m^{6}A site. m^{6}A enrichment was calculated as IP/input normalized to m^{6}A-marked *Gaussia* luciferase RNA spike-in IP/input. Mean ± SEM, two-tailed *t*-test, *$P < 0.05$; **$P < 0.01$, ***$P < 0.001$. Three biological replicates.

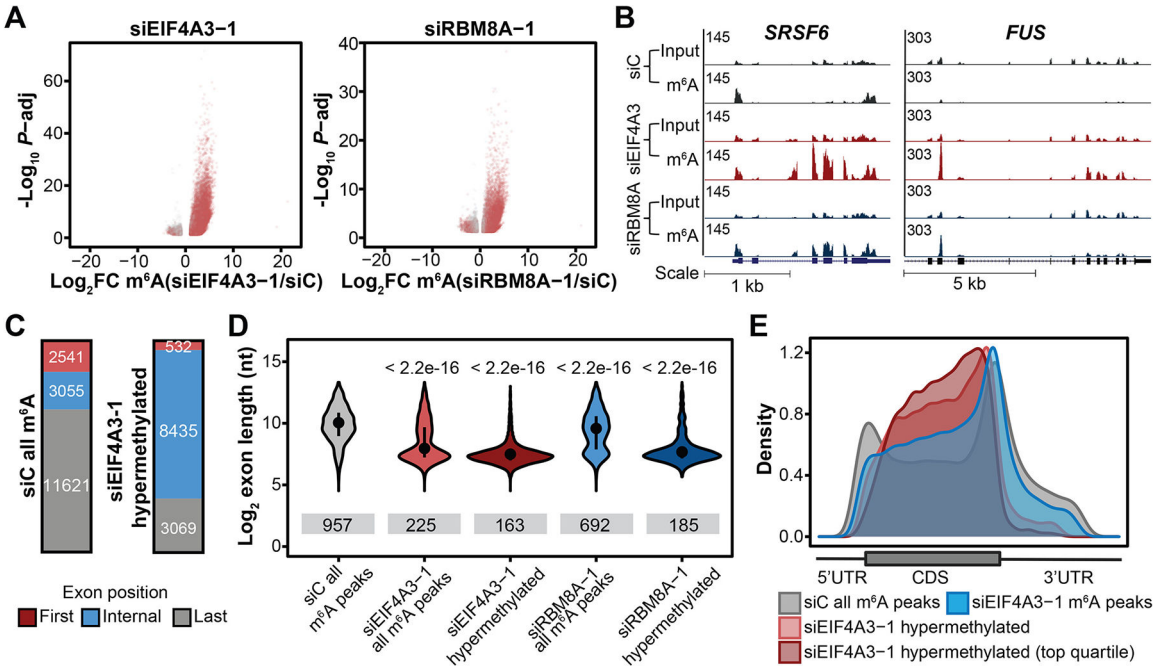

Fig. 3. EJCs protect exon junction-proximal RNA in average-length exons within CDS regions from m^{6}A methylation.

(**A**) Differentially methylated regions upon *EIF4A3* KD (left) and *RBM8A* KD (right) in HeLa cells (FDR<.1, |log$_2$FC|>1). Three biological replicates. Gray and red dots indicate differentially methylated regions that overlap and do not overlap m^{6}A peaks in the control cells, respectively. (**B**) Input and m^{6}A-IP read coverage at *FUS* and *SRSF6* in *EIF4A3* KD, *RBM8A* KD, and control HeLa cells. (**C**) Numbers of *EIF4A3* KD hypermethylated regions (left) and m^{6}A peaks in control cells (right) that reside within first, internal or last exons (**D**) Exon lengths for m^{6}A peaks residing within internal exons in control KD, EIF4A3 KD and RBM8A KD cells, and exon lengths of hypermethylated regions residing within internal exons in EIF4A3 and RBM8A KD cells. Dot and bar represent median and interquartile range, Wilcoxon rank sum test of indicated group vs. siC all m^{6}A peaks. Sample size for each violin plot from left to right is: n = 3166, n = 6659, n = 8438, n = 3817, and n = 3827. (**E**) Metagenes of m^{6}A peaks and significantly hypermethylated m^{6}A regions (and top quartile) in *EIF4A3* KD HeLa cells, and m^{6}A peaks in control cells.

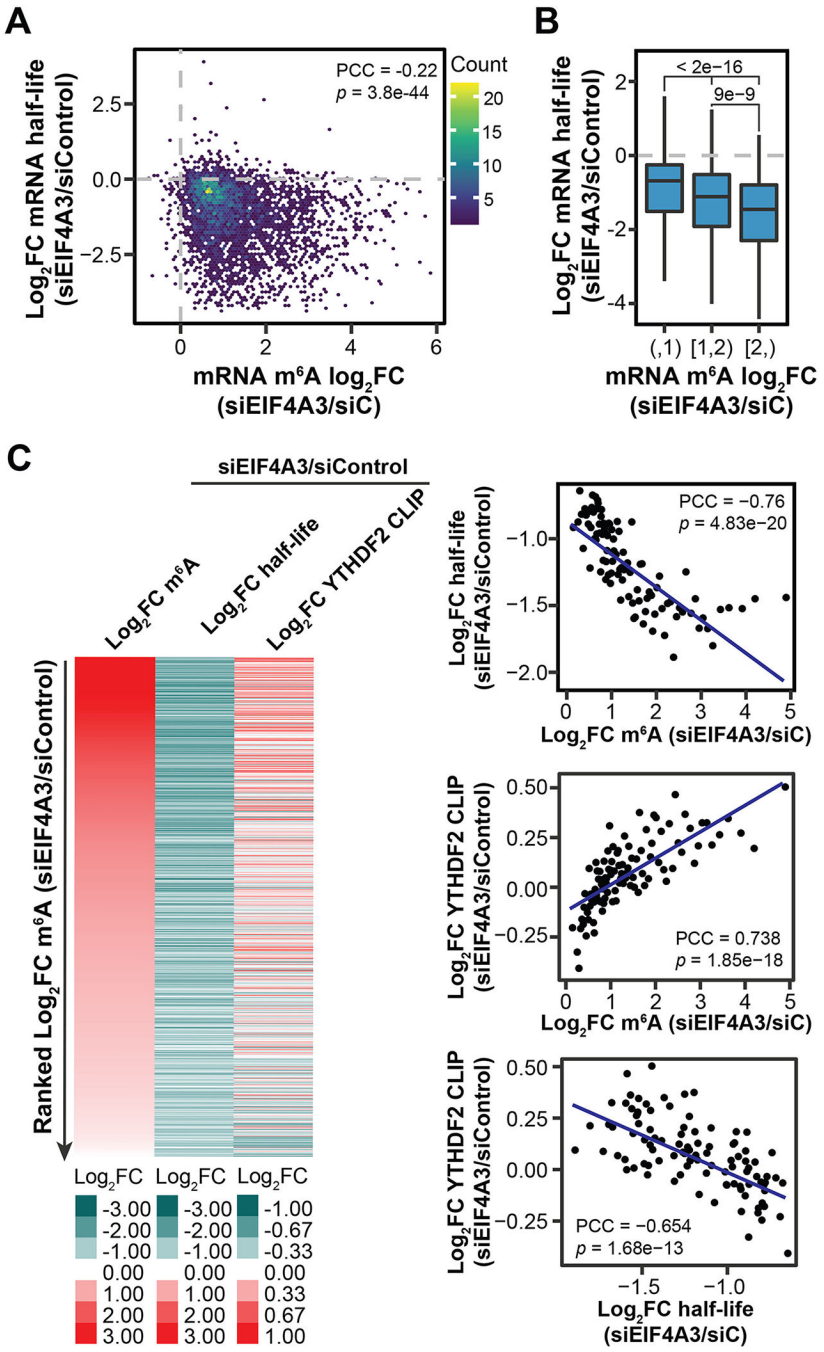

Fig. 4. mRNA m^{6}A hypermethylation upon EJC depletion destabilizes mRNAs.

(**A**) Correlation between fold changes in mRNA half-life and m^{6}A level upon *EIF4A3* KD in HeLa cells (n = 3840). (**B**) Boxplots showing half-life fold changes of hypermethylated mRNAs upon *EIF4A3* KD in HeLa cells. mRNAs were categorized into three groups according to their methylation changes upon *EIF4A3* KD in HeLa cells. *P* values from Wilcoxon rank sum test. Sample size for each boxplot plot from left to right is: n = 1887, n = 1201, and n = 752. (**C**) Left: heatmap showing fold changes in m^{6}A level, mRNA half-life, and YTHDF2 binding upon *EIF4A3* KD in HeLa cells. Right: scatter plots showing the

correlation among fold changes in m^6A level, mRNA half-life, and YTHDF2 binding upon *EIF4A3* KD in HeLa cells. The hypermethylated mRNAs (m^6A $log_2FC > 0$; n = 3424) were categorized into 100 bins based on ranked fold change of m^6A level upon *EIF4A3* KD. For (A) and (C), PCC and *P* values are shown.

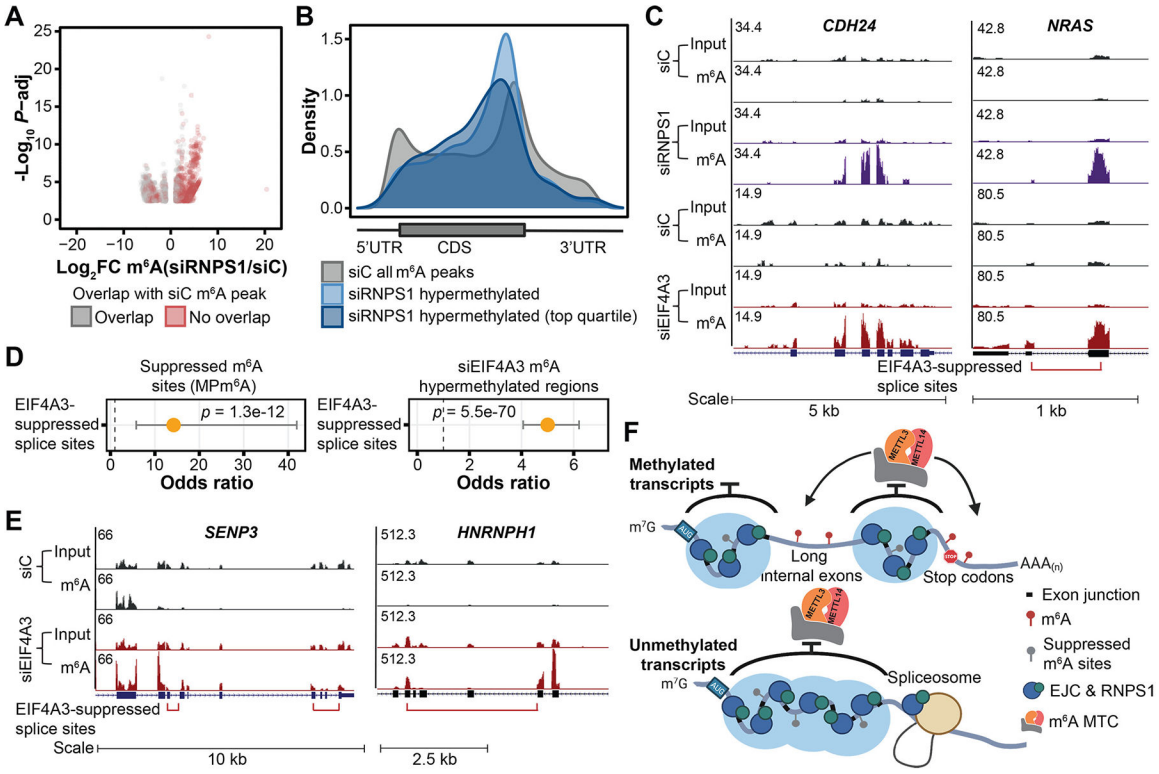

Fig. 5. EJCs and RNPS1 protect proximal RNA regions from aberrant mRNA processing.

(**A**) Differentially methylated regions upon *RNPS1* KD in HeLa cells (FDR < 0.1, $|\log_2 FC| > 1$), three biological replicates. Gray and red dots indicate differentially methylated regions that overlap and do not overlap with m^{6}A peaks in the control KD cells, respectively. (**B**) Metagenes of significantly m^{6}A hypermethylated regions (and top quartile) upon *RNPS1* KD in HeLa cells in comparison with that of all m^{6}A peaks in control cells. (**C**) Input and m^{6}A-IP read coverage at *CDH24* and *NRAS* upon *RNPS1 KD* and *EIF4A3* KD, respectively, as well as corresponding controls in HeLa cells. (**D**) Enrichment of suppressed m^{6}A sites (identified from MPm^{6}A) at EIF4A3-suppressed splice sites (left) and enrichment of EIF4A3 KD hypermethylated regions at EIF4A3-suppressed splice sites (right). Fisher's exact test, dot and bar represent odds ratio and 95% confidence interval. (**E**) Input and m^{6}A-IP read coverage at *SENP3* and *HNRNPH1* in *EIF4A3* KD and control HeLa cells. (**F**) Schematic model depicting that EJCs and RNPS1 (and potentially other EJC-associated proteins) protect exon junction-proximal RNA from m^{6}A deposition through local mRNA packaging. For (C) and (E), red bracket indicates EIF4A3-suppressed splice variant, with ends of bracket indicating the suppressed splice junctions.