

Exon junction complex shapes the m⁶A epitranscriptome

Received: 16 August 2022

Accepted: 14 December 2022

Published online: 23 December 2022

Xin Yang^{1,2}, Robinson Triboulet^{1,2,6}, Qi Liu^{1,2,7,8}, Erdem Sendinc^{1,2} & Richard I. Gregory^{1,2,3,4,5}✉

N6-methyladenosine (m⁶A), the most abundant modification of mRNA, is essential for normal development and dysregulation promotes cancer. m⁶A is highly enriched in the 3' untranslated region (UTR) of a large subset of mRNAs to influence mRNA stability and/or translation. However, the mechanism responsible for the observed m⁶A distribution remains enigmatic. Here we find the exon junction complex shapes the m⁶A landscape by blocking METTL3-mediated m⁶A modification close to exon junctions within coding sequence (CDS). Depletion of EIF4A3, a core component of the EJC, causes increased METTL3 binding and m⁶A modification of short internal exons, and sites close to exon-exon junctions within mRNA. Reporter gene experiments further support the role of splicing and EIF4A3 deposition in controlling m⁶A modification via the local steric blockade of METTL3. Our results explain how characteristic patterns of m⁶A mRNA modification are established and uncover a role of the EJC in shaping the m⁶A epitranscriptome.

Chemical modifications of RNA, the recently termed 'epitranscriptome' play key roles in many biological processes¹. N6-methyladenosine (m⁶A) is the most abundant internal modification found in messenger RNA (mRNA)¹. Numerous recent studies linked m⁶A to various aspects of mRNA metabolism, including splicing, localization, stability, and translation^{1–11}. m⁶A is catalyzed by the methyltransferase (MTase) (writer) complex^{12,13}, the core components of which are METTL3 and METTL14. Biochemical reconstitution and structural studies have demonstrated that this heterodimeric METTL3-METTL14 complex is necessary and sufficient for m⁶A modification of substrate RNAs *in vitro*. METTL3 is the catalytic SAM-binding subunit but requires the METTL14 co-factor that likely mediates RNA-binding¹⁴. However, this core enzyme interacts with numerous other proteins, including VIRMA, ZC3H13, RBM15, HAKAI, and WTAP, as part of an ~1 MDa MTase complex in cells^{15,16}. The precise role of these accessory proteins is currently not well defined. METTL3/14 are essential for normal development^{17–21} and dysregulation can drive tumorigenesis and is implicated in numerous different cancer types²².

Transcriptome-wide mapping of m⁶A modification in different cell types using methylated RNA immunoprecipitation and sequencing (meRIP-seq), as well as more recently using antibody-independent approaches, has revealed that m⁶A occurs on large subsets of several thousand mRNAs within a particular cell, and that modification occurs at a consensus 'RRACH' motif (R = A/G, H = A/C/U)^{23,24}. Intrinsic substrate specificity of the METTL3-METTL14 MTase can explain the preferential modification of this consensus motif, yet only about 5% of all RRACH motifs are m⁶A-modified in cellular RNA²⁵. FTO and ALKBH5 have been identified as demethylases (erasers) capable of m⁶A removal^{11,26}. Importantly, m⁶A is not evenly distributed throughout mRNAs and is instead highly enriched in 3' UTRs at sites close to the stop codon, as well as within long internal exons.

Despite the major importance of m⁶A in post-transcriptional gene control, the mechanisms responsible for its distribution throughout the transcriptome remain largely unknown. m⁶A modification is considered a co-transcriptional event^{27–30}, and it is

¹Stem Cell Program, Division of Hematology/Oncology, Boston Children's Hospital, Boston, MA 02115, USA. ²Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, Boston, MA 02115, USA. ³Department of Pediatrics, Harvard Medical School, Boston, MA 02115, USA. ⁴Harvard Stem Cell Institute, Cambridge, MA 02138, USA. ⁵Harvard Initiative for RNA Medicine, Boston, MA 02115, USA. ⁶Present address: Twentyeight-Seven Therapeutics, Watertown, MA 02472, USA. ⁷Present address: Rice Research Institute, Guangdong Academy of Agricultural Sciences, Guangzhou 510640, China. ⁸Present address: Guangdong Key Laboratory of New Technology in Rice Breeding, Guangzhou 510640, China. ✉e-mail: rgregory@enders.tch.harvard.edu

therefore expected that m⁶A profiles are established in nascent RNA. However, the distribution of m⁶A in nascent RNA is controversial. Unlike mature mRNAs, m⁶A methylation is not restricted to the 3' UTR of nascent RNA transcripts^{29,31,32}. Ke et al. found ~93% of m⁶As in chromatin-associated RNA to be located in exons²⁷. Using TNT-seq to specifically examine m⁶A in newly synthesized RNAs, however, Louloup et al. found most m⁶A sites to be located in introns²⁸. Both studies found more m⁶A near splice junctions in nascent RNA compared with mature RNA, suggesting that mRNA maturation can reshape the m⁶A topology. So far, studies have implicated MTase recruitment by polymerase II and trans-acting factors, including transcription factors, RNA-binding proteins, or histone modifications, to promote localized m⁶A modification^{4,33–37}. For example, it has been proposed that H3K36me3 chromatin modification might help recruit METTL3-METTL14 and thereby enrich m⁶A towards the 3'-end of mRNAs³⁷. In addition, non-coding RNAs are involved in m⁶A modification deposition by recruiting m⁶A regulators to the UTR region³⁸. Increasing evidence indicates that m⁶A demethylases associate with chromatin and engage in RNA splicing regulation^{29,30,39,40}. However, none of these known links can adequately explain the characteristic m⁶A distribution across mRNAs with a strong peak in the 3' UTR, and enrichment of m⁶A in long internal exons.

In this study, we explore the mechanism responsible for the characteristic m⁶A profile that has been observed for over a decade. We find the core component of the exon junction complex, EIF4A3, shapes the m⁶A landscape by blocking METTL3-mediated m⁶A modification close to exon-exon boundaries within the mRNA coding sequence (CDS). Our conclusions are based on METTL3 eCLIP-seq and m⁶A MeRIP-seq data showing that METTL3 binding and m⁶A modification is restricted by EIF4A3, in particular within short internal exons and in the vicinity of exon-exon boundaries. We provide additional evidence using reporter gene assays to further support this conclusion. Our results explain how m⁶A topology is established and uncover a role of the EJC in shaping the m⁶A epitranscriptome.

Results

Dynamic m⁶A profiles during mRNA maturation

To begin to explore the mechanism(s) responsible for the characteristic m⁶A distribution on mature mRNAs, we isolated chromatin-associated RNA (caRNA; rRNA depleted and enriched for nascent transcripts), and polyA⁺ RNA (mature mRNA) from HeLa cells, and quantified m⁶A modification levels by HPLC-MS/MS. Global m⁶A levels were found to be similar between caRNA and polyA⁺ RNA (Supplementary Fig. 1a). Considering the presence of other types of non-coding RNA in the caRNA fraction that could influence global m⁶A measurements, we performed m⁶A methylated RNA immunoprecipitation sequencing (MeRIP-seq), and compared m⁶A modification profiles between caRNA and polyA⁺ mRNA. As expected, although the RRACH motif is quite evenly distributed across mRNAs (Fig. 1a), a strong peak of m⁶A close to the stop codon was observed in mature mRNAs (Fig. 1a). Distinct, however, from this characteristic m⁶A modification landscape in polyA⁺ RNA, m⁶A modification in caRNA was much more evenly distributed (Fig. 1a), a finding that is consistent with other publicly available MeRIP-seq datasets^{2,28,29}. Among the m⁶A peaks located in exons of protein-coding genes, 16,758 and 7766 peaks were exclusively detected in caRNAs or polyA⁺ RNAs, respectively, whereas 9942 peaks are common to both RNA samples (Fig. 1b). m⁶A occurred at the consensus motif in both polyA⁺ and caRNA (Supplementary Fig. 1b), and mainly within exons of protein-coding genes (Supplementary Fig. 1c). Notably, the m⁶A peaks in caRNA are mostly located in internal exons, whereas in polyA⁺ RNA the m⁶A peaks are strongly enriched in last exons (Fig. 1c). Furthermore, polyA⁺ RNA had higher overall m⁶A modification than caRNA (Supplementary Fig. 1d). More

detailed analysis showed relative enrichment of m⁶A close to splice sites in caRNA compared to polyA⁺ RNA (Fig. 1d). Moreover, the length of internal exons with specific m⁶A peaks in caRNAs was generally shorter than those of polyA⁺ RNAs (Fig. 1e). We then divided internal exons into groups based on their length and found that the m⁶A signal in short internal exons of caRNAs to be higher than in polyA⁺ RNAs. An opposite pattern was observed for long exons and regions near the stop codon with a higher m⁶A signal in polyA⁺ RNA (Fig. 1f). Together, these results support dynamic patterns of m⁶A modification during mRNA maturation, with relatively more m⁶A present in short internal exons of nascent mRNAs compared with mature mRNAs. These observations raise the possibility that m⁶A may be actively removed from regions close to exon-exon boundaries during mRNA maturation, and/or extra m⁶A is deposited more distal to exon-exon boundaries in spliced mRNAs (Supplementary Fig. 1e).

Minimal role of ALKBH5 demethylase during mRNA maturation

Since m⁶A demethylases (FTO and ALKBH5) have been reported to be involved in pre-mRNA splicing regulation^{11,40,41}, we first hypothesized that dynamic m⁶A distribution during mRNA maturation might involve removal of m⁶A in internal exons by FTO or ALKBH5 accompanied by splicing factors. To explore this, we performed pull-down assays coupled with protein mass spectrometry and found ALKBH5, but not FTO, interacts with numerous splicing factors (Supplementary Fig. 2a, b and Supplementary Data 1). Further analysis showed components of the Exon Junction Complex to be most highly enriched, and we verified this by immunoprecipitation and Western blots (Supplementary Fig. 2c, d and Supplementary Table 1). EJC is a multiprotein complex that deposited ~24 nucleotides upstream of exon-exon junctions during mRNA maturation. EJC assembly begins during splicing, and the first step involves the position-specific deposition of the DEAD-box protein EIF4A3 onto RNA by the spliceosome^{42,43}. We therefore next examined the role of ALKBH5, and the core component of the EJC, EIF4A3, in modulating the m⁶A epitranscriptome. We individually knocked down ALKBH5 and EIF4A3 using siRNAs and measured global changes in m⁶A levels of polyA⁺ RNA by HPLC-MS/MS. We found ALKBH5 depletion had a very minor effect on m⁶A levels. Strikingly, however, EIF4A3 depletion dramatically increased the global m⁶A level in polyA⁺ RNA (Supplementary Fig. 2e, f). These results pointed to a possible role of EIF4A3 and the EJC in shaping the m⁶A epitranscriptome.

The core EJC subunit EIF4A3 influences m⁶A profiles

Based on our findings that EIF4A3- but not ALKBH5-depletion impacts global m⁶A levels, we next explored how the m⁶A epitranscriptome might be influenced by EIF4A3 and considered an alternative model whereby the methylation of internal exons might be hindered by the EJC rather than requiring ALKBH5-mediated demethylation of these sites (Supplementary Fig. 1e). In order to identify differentially modified sites more comprehensively, we utilized RADAR⁴⁴. Consistent with our HPLC-MS/MS results showing substantially increased global m⁶A level in polyA⁺ RNA in EIF4A3-depleted cells, MeRIP-seq results showed a remarkable induction of m⁶A modification sites upon EIF4A3- but not ALKBH5-depletion (Fig. 2a and Supplementary Fig. 2g, h). Further analysis revealed that EIF4A3 KD caused 43,744 and 8392 regions to be hypermethylated and hypomethylated, respectively (Fig. 2a), with most hypermethylated and hypomethylated m⁶A regions annotated to protein-coding genes (Supplementary Fig. 3a). Metaplot showed that hypermethylated regions upon EIF4A3 depletion are highly enriched in the CDS, while the pattern of hypomethylated sites was similar to overall m⁶A (Fig. 2b). Considering previous studies showing enrichment of EIF4A3 binding within short internal exons of spliced mRNAs^{43,45}, we next explored whether exons with hypermethylated m⁶A upon EIF4A3 depletion have similar features. Indeed, we found the

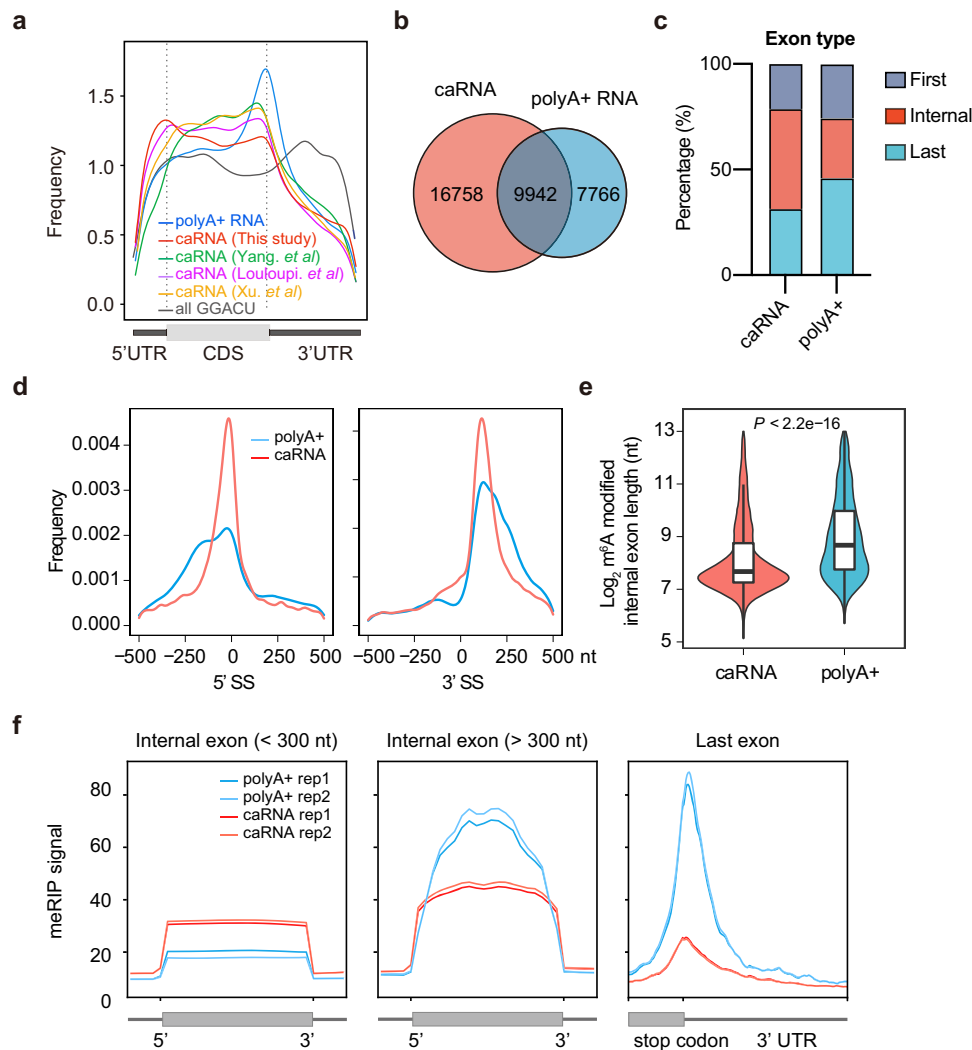


Fig. 1 | Nascent transcripts contain more m⁶A in short internal exons compared with mature mRNAs. **a** Metaplot of meRIP-Seq showing the m⁶A distribution in polyA⁺ RNAs compared with chromatin-associated RNAs (caRNAs). Also shown is the distribution of the GGACU sequence motif. **b** Venn diagram showing the overlap of m⁶A peaks identified in polyA⁺ RNAs and caRNAs. **c** Barplot showing the percentage of m⁶A peaks within the first, internal, and last exons of protein-coding

genes. **d** Distribution of the distance of m⁶A peaks to the closest splice site. **e** The length of internal exons with m⁶A peaks identified in caRNA, polyA⁺ RNA. The three lines inside the violin plots are the first quartile, median, and third quartile. $n = 9646; 3392$. Statistical analyses, unpaired two-tailed Student's *t*-test. **f** Aggregation plots of meRIP-Seq read signals in short-, or long- internal exons, and last exon. Internal exons were divided into two groups based on length.

proportion of internal exons with m⁶A modification was increased in EIF4A3-depleted cells, with the hypermethylated m⁶A sites being highly enriched in internal exons (Fig. 2c). We next analyzed the relative length of internal exons with altered m⁶A modification upon EIF4A3 knockdown and found a strong enrichment of hypermethylated sites localized within short internal exons (Fig. 2d). EIF4A3 depletion significantly increased m⁶A modification of short internal exons, but not long exons or the 3' UTR (Fig. 2e). For long exons, although the m⁶A signal was overall unchanged upon EIF4A3 knockdown, increased m⁶A was specifically observed at sites proximal to splice junctions (Supplementary Fig. 3b). Moreover, we found that hypermethylated m⁶A sites upon EIF4A3 depletion are enriched in longer mRNAs containing more exons (Fig. 2f, g), which is highly consistent with known features of EJC deposition^{45,46}. In addition, we analyzed the methylation status of single-exon (intronless) genes, that are expected to be substantially less or not associated with the EJC^{43,45,47}. We found that single-exon genes have a higher m⁶A level than the internal exons of multi-exon genes (Supplementary Fig. 3c), and that depletion of EIF4A3 did not impact m⁶A peak enrichment of single-exon genes (Supplementary Fig. 3d). Next, considering the

relative enrichment of m⁶A close to splice sites in caRNA compared to polyA⁺ RNA (Fig. 1d), as well as that the EJC resides -24 nucleotides upstream of exon-exon junctions^{43,45,46,48}, we next analyzed changes in m⁶A modification due to EIF4A3 deficiency at splice sites. EIF4A3 KD resulted in increased m⁶A at splice sites (Fig. 2h), suggesting changes in m⁶A could be linked to splicing. Further analysis identified a relatively small number of changes in exon usage, and that altered exon usage is positively correlated with changes in m⁶A modification upon EIF4A3 depletion (Supplementary Fig. 3e, f). Hypermethylated m⁶A sites might be recognized by certain m⁶A reader proteins to promote exon inclusion^{5,49}. Like protein-coding genes, long non-coding RNAs (lncRNAs) are transcribed by RNA polymerase II and have various lengths and exon numbers. The introns of lncRNAs are also spliced during transcription utilizing mechanisms shared with protein-coding genes. We, therefore, analyzed whether EIF4A3 influences the m⁶A modification of lncRNAs. We found 1155 hypermethylated and 669 hypomethylated regions on lncRNAs upon EIF4A3 depletion (Supplementary Fig. 4a), and like protein-coding genes, hypermethylated regions are enriched in longer genes with more exons (Supplementary Fig. 4b, c). These findings indicate that

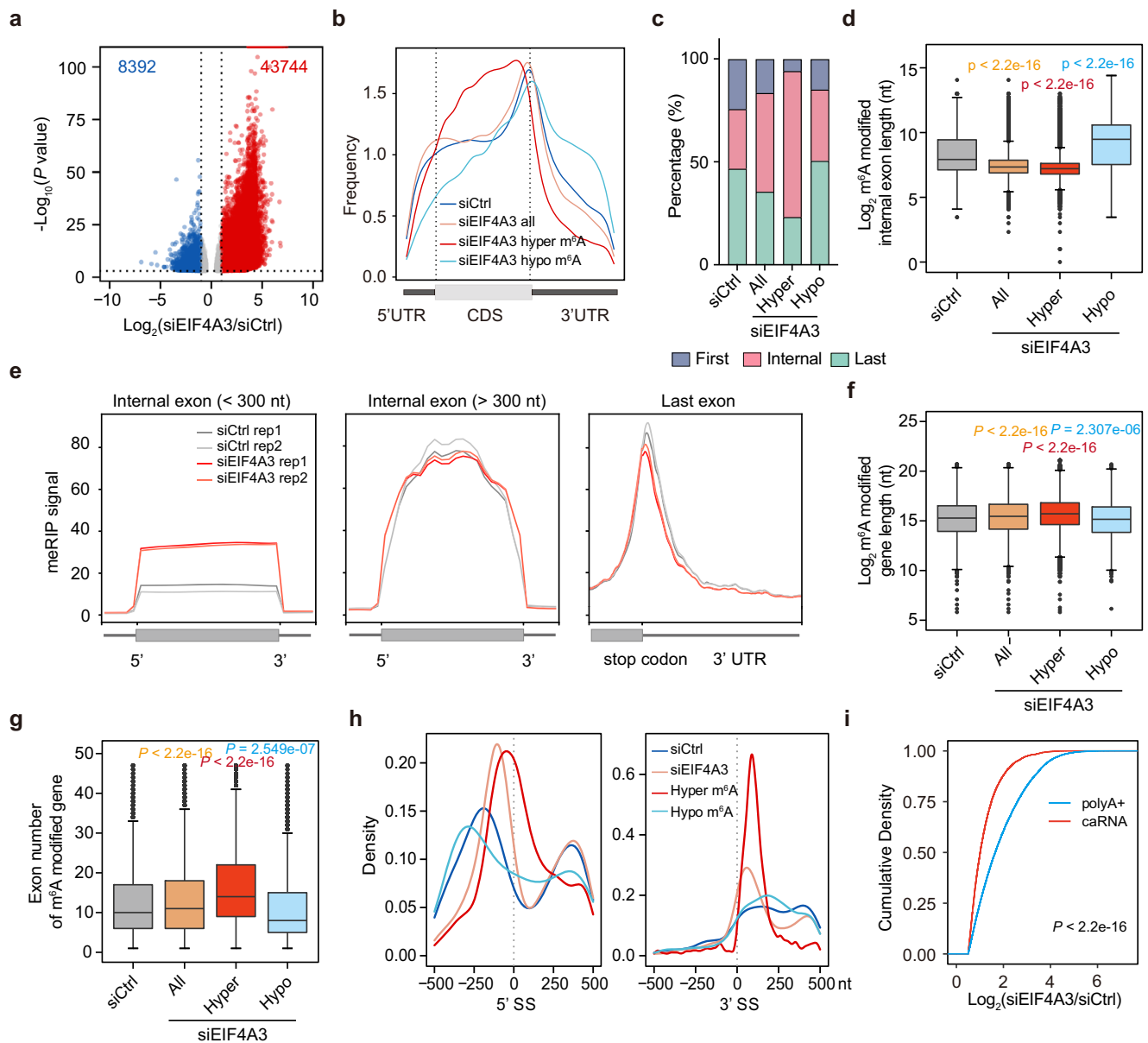


Fig. 2 | Depletion of EIF4A3 increases m⁶A modification of short internal exons.

a Differential m⁶A regions upon EIF4A3 depletion. Red dots represent increased m⁶A regions (hypermethyated m⁶A), Blue dots represent decreased m⁶A regions (hypomethyated m⁶A). **b** Metaplot showing the distribution of all m⁶A peaks in siCTRL and siEIF4A3, hypermethyated, hypomethyated m⁶A. **c** Barplot showing the percentage of different exons with m⁶A peaks in siCTRL and siEIF4A3, hypermethyated, hypomethyated m⁶A. **d** The length of m⁶A modified internal exons in siCTRL and siEIF4A3, and the length of internal exons with hypermethyated or hypomethyated m⁶A. Solid line represents median, with whiskers indicating minimum to maximum values. $n = 1743; 6643; 44,010$; and 1854. Statistical analyses, unpaired two-tailed Student's *t*-test. **e** Aggregation plots showing m⁶A enrichment in short, long and last exon. Internal exons were divided into two groups based on

length. **f** The length of m⁶A modified genes in siCTRL and siEIF4A3, and the length of genes with hypermethyated or hypomethyated m⁶A. Solid line represents median, with whiskers indicating minimum to maximum values. $n = 10,335; 12,291; 44,293$; and 8238. Statistical analyses, unpaired two-tailed Student's *t*-test. **g** The exon number of m⁶A modified genes in siCTRL and siEIF4A3, and the exon number of genes with hypermethyated or hypomethyated m⁶A. Solid line represents median, with whiskers indicating minimum to maximum values. $n = 10,335; 12,291; 44,293$; and 8238. Statistical analyses, unpaired two-tailed Student's *t*-test. **h** Distribution of the distance of m⁶A peak summits to the closest splicing sites. **i** Cumulative curves showing fold-change of hypermethyated m⁶A induced by EIF4A3 KD in polyA+ RNA and caRNA. *P* value was calculated using two-sided Wilcoxon and Mann–Whitney test.

EJC broadly influences m⁶A modification of short internal exons of polymerase II transcribed RNAs. In addition, we found that EIF4A3 depletion caused a more pronounced increase in m⁶A levels on polyA+ RNAs than on caRNAs (Fig. 2i). Collectively, our data support that EIF4A3, the core RNA-binding component of the EJC, influences m⁶A modification close to splice sites within the CDS of mRNAs to shape m⁶A distribution throughout the transcriptome during mRNA maturation.

EIF4A3 locally inhibits METTL3 binding and m⁶A modification
We postulated that EIF4A3 bound to exon-exon boundaries might locally occlude METTL3 mRNA association to shape the m⁶A epitranscriptome. To test this, we performed METTL3 eCLIP-seq in control and EIF4A3 depleted HeLa cells (Supplementary Fig. 5a–d). Further analysis identified 472 peaks from 400 genes that were significantly increased upon EIF4A3 depletion, while 231 peaks from 192 genes were decreased (Supplementary Fig. 5e).

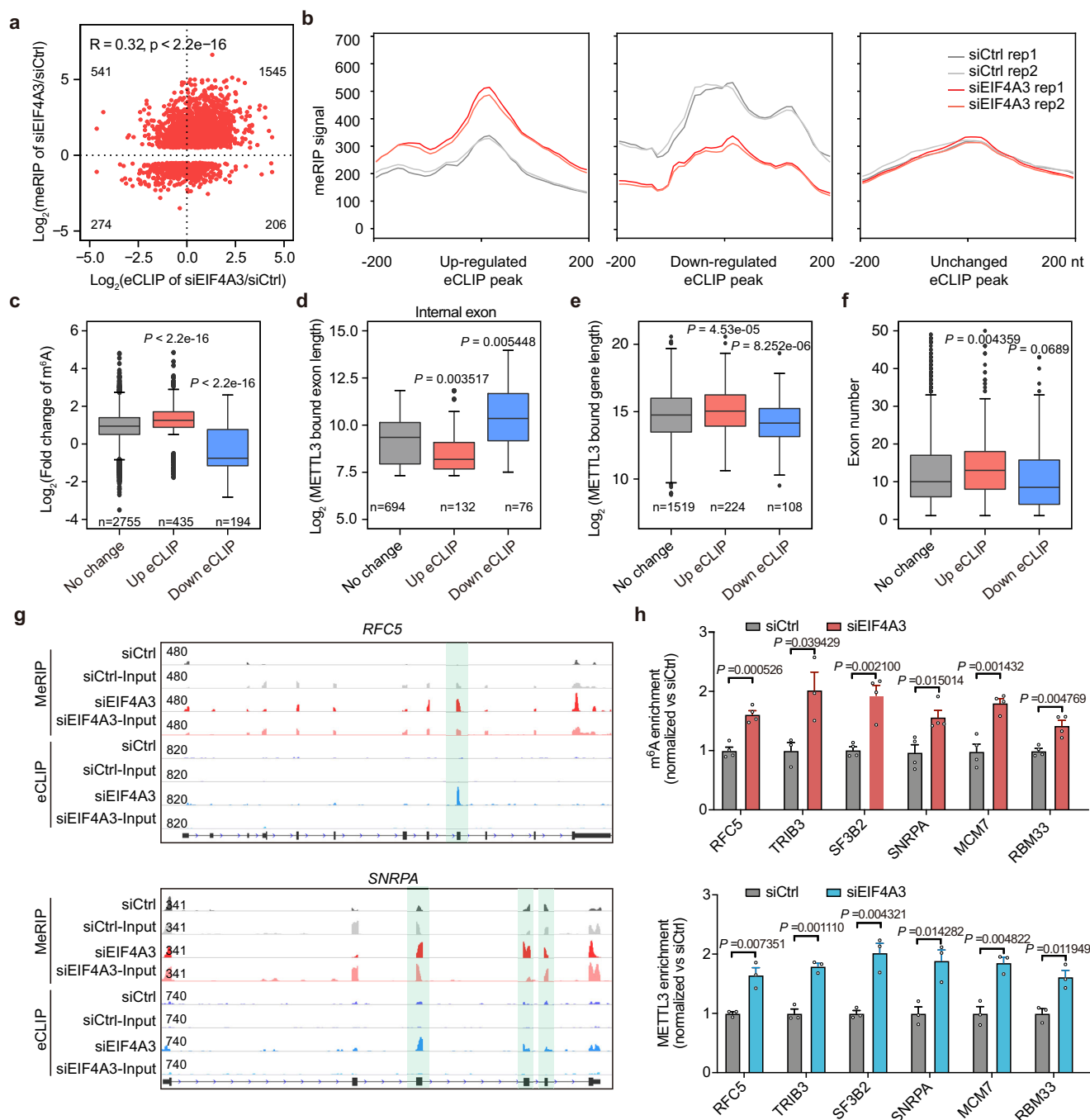


Fig. 3 | EIF4A3 locally inhibits METTL3 binding and m⁶A modification.

a Correlation between m⁶A fold change and METTL3 binding fold change upon EIF4A3 knockdown. Correlation coefficient (r) and P value was calculated by Pearson's correlation analysis. **b** Average distribution of MeRIP-seq signal is shown, aligned around increased, decreased, and unchanged METTL3 binding peaks. **c** Boxplot showing the fold change of m⁶A enrichment on METTL3 binding peaks upon EIF4A3 depletion. METTL3 binding peaks are divided into three groups: no change, up regulated and down regulated upon EIF4A3 KD. Solid line represents median, with whiskers indicating minimum to maximum values. Statistical analyses, unpaired two-tailed Student's t -test. **d** The length of internal exons with unchanged, increased or decreased METTL3 binding peaks. Solid line represents median, with whiskers indicating minimum to maximum values. Statistical analyses, unpaired

two-tailed Student's t -test. **e** The length of protein coding genes with unchanged, increased or decreased METTL3 binding peaks. The solid line represents median, with whiskers indicating minimum to maximum values. Statistical analyses, unpaired two-tailed Student's t -test. **f** The exon number of genes with unchanged, increased, or decreased METTL3 binding peaks. Solid line represents median, with whiskers indicating minimum to maximum values. $n = 1519; 224; 108$. Statistical analyses, unpaired two-tailed Student's t -test. **g** Genome Browser tracks of meRIP-seq and METTL3 eCLIP-seq read coverage at gene *RFC5* and *SNRPA* in siCtrl and siEIF4A3 HeLa cells. **h** meRIP-qPCR and METTL3 CLIP-qPCR analysis showing increased m⁶A modification and METTL3 binding ability upon EIF4A3 KD. Data are mean \pm S.E.M. of three or four independent experiments. Statistical analyses, two-tailed Student's t -test. Source data are provided as a Source Data file.

Importantly, a positive correlation was found between mRNAs with increased METTL3 occupancy and increased m⁶A modification upon EIF4A3 KD compared with control cells (Fig. 3a). Consistently, increased m⁶A modification was observed at sites with

increased, but not decreased or unchanged, METTL3 occupancy upon EIF4A3 depletion (Fig. 3b, c), and METTL3 occupancy was increased at sites of increased m⁶A modification when EIF4A3 was diminished (Supplementary Fig. 5f). Furthermore, 314 of the 472

increased METTL3 binding peaks upon EIF4A3 knockdown contained hypermethylated m⁶A regions (Supplementary Fig. 5g). A highly significant correlation between sites of hypermethylation and increased METTL3 binding upon EIF4A3 knockdown was observed (Supplementary Fig. 5h). Consistent with the dynamic changes in m⁶A profiles during mRNA maturation (Fig. 1) and with the changes in m⁶A profiles we identified with EIF4A3 depletion (Fig. 2), we found that EIF4A3 KD led to specifically increased METTL3 binding to short internal exons and longer mRNAs containing more exons (Fig. 3d–f). Genome Browser tracks of MeRIP-seq and METTL3 eCLIP-seq at selected genes are shown to exemplify the role of EIF4A3 in restricting METTL3 binding near splice site junctions to influence patterns of m⁶A mRNA modification (Fig. 3g), findings that we confirmed by m⁶A meRIP-qPCR and METTL3 CLIP-qPCR analysis for several individual genes that we identified with hypermethylated m⁶A and increased METTL3 occupancy upon EIF4A3 knockdown (Fig. 3h). These findings suggest that METTL3 recruitment to spliced mRNAs is hindered by the EJC complex, resulting in lower m⁶A modification within the CDS of mRNAs and in particular in short internal exons.

EJC blocks METTL3-mediated m⁶A modification during mRNA splicing

EIF4A3 is recruited to mRNA by the spliceosome during splicing and serves as a platform for the assembly of the EJC after exon ligation⁵⁰. We reasoned that splicing is required for EJC recruitment to locally restrict METTL3 accessibility and m⁶A mRNA modification. To further test this, we employed reporter assays. We individually cloned the exonic regions of several genes we found to be both m⁶A hypermethylated and with increased METTL3 binding in EIF4A3-deficient cells, into exon 2 of the intron-containing mouse beta-Globin gene (precursor reporters, pre-WT). As controls for each gene, a mature version of beta-Globin reporter without introns but containing the same exonic inserts was generated (mature reporters, Mature) (Fig. 4a). Cells were transfected with the different reporters, and relative levels of m⁶A modification or METTL3-association was measured by α -m⁶A meRIP-qPCR or α -METTL3 CLIP-qPCR, respectively. Compared with mature reporters, precursor reporters had overall lower m⁶A levels and decreased association with METTL3 (Fig. 4b, c and Supplementary Fig. 6a, b). Moreover, EIF4A3 depletion caused substantially increased m⁶A modification and METTL3 binding to the precursor reporters, but not to the mature reporters (Fig. 4b, c and Supplementary Fig. 6a, b). We further generated specific GU/AG splicing mutant reporters (pre-Mut) and found splicing deficiency led to increased m⁶A levels compared to the corresponding wild-type pre-reporters (pre-WT) (Fig. 4a, d and Supplementary Fig. 6c). Ectopic expression of EIF4A3 decreased m⁶A levels of pre-WT but not pre-Mut or mature reporters (Fig. 4d and Supplementary Fig. 6d). Together, these results directly support our model that EJC deposition during mRNA splicing inhibits METTL3 binding and m⁶A modification.

To further directly test the role of EIF4A3 in locally blocking m⁶A deposition, we performed tethering experiments using mature mRNA reporters containing two MS2-binding sites located in the exon junction regions together with ectopic expression of an MS2-EIF4A3 or MS2-GFP fusion proteins (Fig. 4e and Supplementary Fig. 6e). We found that directly tethering MS2-EIF4A3, but not MS2-GFP control, to the exon junction region suppressed the m⁶A modification level and METTL3 binding of several different gene reporters (Fig. 4f, g). In addition, the binding capacity of the m⁶A MTase complex component WTAP to the reporter transcripts was also found to be inhibited by increased EIF4A3 expression (Supplementary Fig. 6f). These findings further directly support that the EJC, and in particular EIF4A3, blocks METTL3-mediated m⁶A modification within the vicinity of exon-exon boundaries of certain intron-containing mRNAs to shape m⁶A distribution throughout the epitranscriptome.

Discussion

m⁶A is highly enriched in the 3' UTR of a large subset of mature mRNAs to influence mRNA stability and/or translation yet the mechanism responsible for the observed m⁶A distribution has remained enigmatic. In this study, we find the exon junction complex shapes the m⁶A landscape by blocking METTL3-mediated m⁶A modification close to exon junctions within the mRNA coding sequence (CDS). We propose a model in which EJC-binding at exon-exon boundaries reduces the local accessibility of the mRNA to METTL3 resulting in less m⁶A modification throughout the CDS region of mRNA. In contrast, EJC-depleted regions, including long internal exons, and terminal exons that typically comprise 3' UTRs, are accessible to METTL3 binding and m⁶A modification (Fig. 5).

Our model is supported by METTL3 eCLIP-seq and m⁶A meRIP-seq data showing that METTL3 binding and m⁶A modification is restricted by EIF4A3. Upon EIF4A3 knockdown, m⁶A levels globally increased (measured by mass spec), and at the individual gene level positively correlated with increased m⁶A modification and METTL3 binding to mRNA. Both hypermethylated m⁶A and increased METTL3 binding regions are enriched in short internal exons, suggesting EIF4A3 blocks METTL3 from binding to short exons to deposit m⁶A. Known features of EJC occupancy across the transcriptome help further support our model of localized EIF4A3-mediated suppression of m⁶A mRNA modification. We provide additional evidence using reporter gene assays to further support this conclusion. Although EIF4A3, but not GFP, hinders m⁶A deposition, it remains possible that other RBPs might also influence m⁶A modification by a similar steric hinderance mechanism. Although we initially identified EJC components as ALKBH5-interacting proteins, we did not find any evidence for a requirement of ALKBH5 in controlling global m⁶A levels or influencing m⁶A profiles during mRNA maturation. Thus, demethylation of internal exons is unlikely to be responsible for the dynamic profile of m⁶A distribution. Therefore the relevance of the interaction between ALKBH5 and the EJC currently remains unknown. Furthermore, closer investigation reveals the apparent caRNA-specific peaks that might be expected to be removed during mRNA maturation (Fig. 1b) are actually also typically modified in mature polyA⁺ mRNAs but are not routinely computationally identified using the standard bioinformatics tools since these sites are considerably less prominent than the more heavily modified 3' UTR peaks that exist in mature mRNAs (Fig. 1f and Supplementary Fig. 1d). Thus, we propose that m⁶A occurs throughout the mRNA during transcription albeit at a relatively low stoichiometry, and that subsequent and locally restricted METTL3/14 activity by the EJC leads to accumulation of m⁶A in the 3' UTR of mature mRNAs.

While our findings provide a clear molecular explanation for the long-standing question of how characteristic m⁶A profiles are established, it remains unclear why this particular m⁶A topology is needed. A previous study suggested that elevated m⁶A modifications located in CDS regions reduced translational efficiency³⁴. EJC-mediated blocking of m⁶A within ORFs may function as a protective mechanism against aberrant mRNA hypermethylation in the CDS region, thereby promoting translation. The exon density and exon lengths within individual mRNAs might be an important determinant of levels of gene expression through this mechanism, with implications for the design of cDNA expression constructs both in the research lab and potentially also for gene therapies in the clinic. Alternatively, given our previous findings that METTL3 at sites close to the stop codon facilitates mRNA circularization for ribosome recycling and enhanced mRNA translation, this EJC-mediated mechanism controlling METTL3 binding and m⁶A distribution might be important to prevent alternative and non-productive mRNA looping at sites within the CDS³. Moreover, considering the well-established role of the EJC, it will be interesting also to explore how

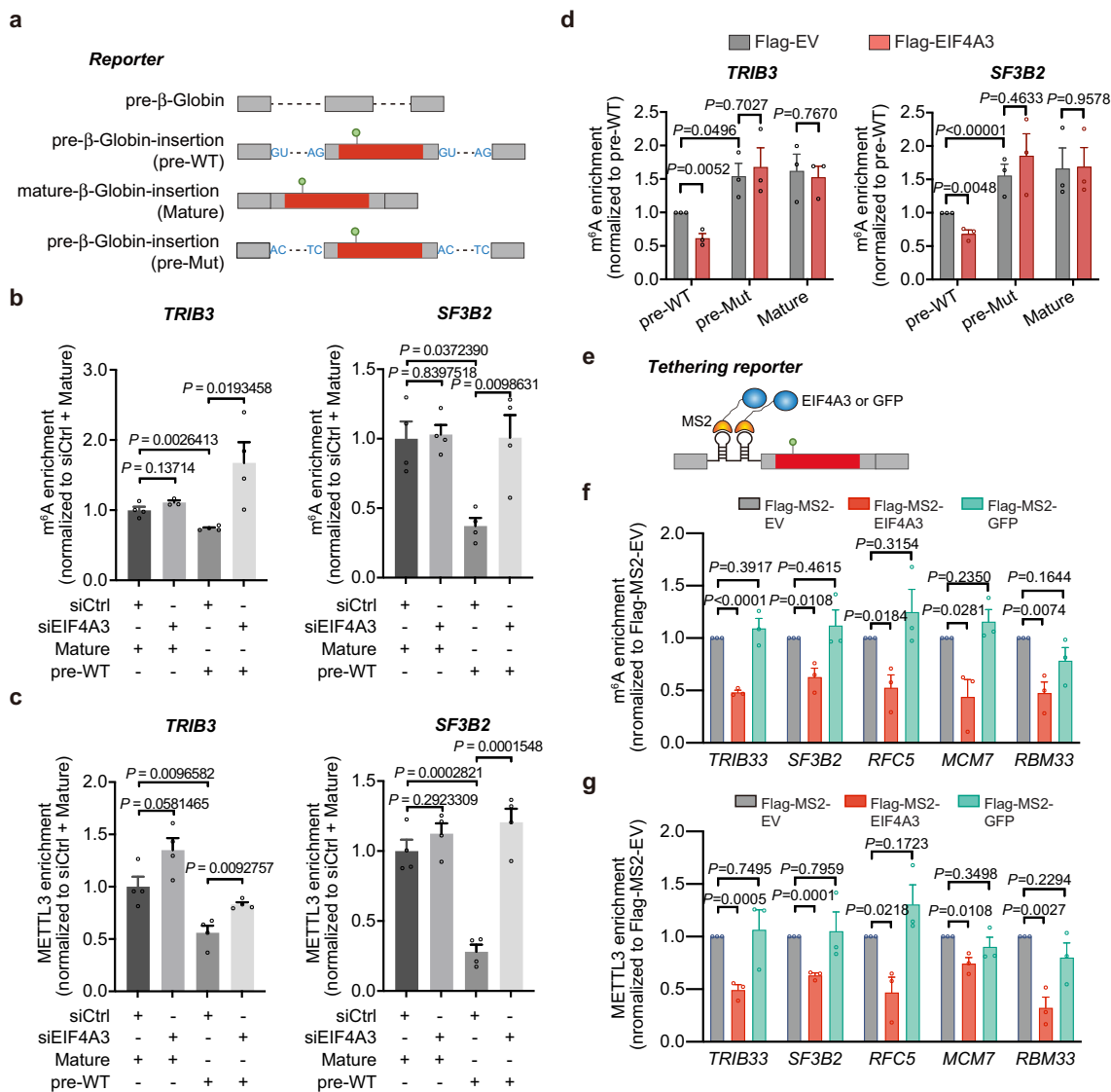


Fig. 4 | EIF4A3 blocks METTL3-mediated m⁶A modification during mRNA maturation. **a** Schematic of reporter constructs. Grey boxes and dash lines present the exons and introns of mouse beta-globin, respectively. Red boxes present the exons of candidate genes with hypermethylated m⁶A (green dot). GU/AG splicing mutation is indicated. **b, c** meRIP-qPCR (**b**) and METTL3 CLIP-qPCR (**c**) analysis of reporter constructs showing m⁶A and METTL3 enrichment upon EIF4A3 KD. Data are mean \pm S.E.M. of three independent experiments. Statistical analyses, two-tailed Student's *t*-test. **d** meRIP-qPCR analysis of pre-WT, pre-Mut, and Mature

reporter constructs showing m⁶A enrichment in Flag-tagged empty vector and EIF4A3 overexpressed HeLa cells. Data are mean \pm S.E.M. of three or four independent experiments. Statistical analyses, two-tailed Student's *t*-test. **e** Schematic of tethering reporter construct. **f, g** meRIP-qPCR (**f**) and CLIP-qPCR (**g**) analysis of tethering reporter constructs showing decreased m⁶A enrichment and METTL3 binding ability upon MS2-EIF4A3 but not MS2-GFP overexpression. Data are mean \pm S.E.M. of three independent experiments. Statistical analyses, two-tailed Student's *t*-test. Source data are provided as a Source Data file.

m⁶A modification might be related to nonsense-mediated decay (NMD) pathways.

Methods

Cell culture and antibodies

HeLa (ATCC, CCL-2) and HEK293T (ATCC, CRL-3216) were cultured in Dulbecco's Modified Eagle Medium (DMEM) (Gibco, 11965-092) supplemented with 10% fetal bovine serum (FBS) (Gemini Bio., 100-106), 1 mM sodium pyruvate (Gibco, 11360070), and 0.5% penicillin/streptomycin (ThermoFisher, 15140-122) at 37 °C in a humidified incubator with 5% CO₂. Cell lines were authenticated with morphology, karyotyping, and PCR-based approaches by ATCC.

The following antibodies were purchased from the indicated suppliers. Mouse anti-Flag M2 (sigma, F3165, 1:2000), anti-EIF4A3 (Abcam, ab32485, 1:2000), anti-Y14 (Abcam, ab5828, 1:2000), anti-MAGO3 (Abcam, ab186431, 1:2000), anti-PNN (Abcam, 244250,

1:2000), anti-ALKBH5 (Cell Signaling Technology, 80283, 1:1000), anti-ALYREF (Abcam, ab202894, 1:2000), anti-METTL3 (Abcam, ab195352, 1:2000 for West blotting and 1:100 for Immunoprecipitation), anti-WTAP (Proteintech, 60188, 1:2000 for West blotting and 1:100 for Immunoprecipitation), anti-m⁶A (Abcam, ab151230, 1:100), anti-beta-Actin (Abcam, 8226, 1:3000), anti-rabbit IgG, HRP-linked Antibody (Cell Signaling Technology, 7074, 1:3000), anti-mouse IgG, HRP-linked Antibody (Cell Signaling Technology, 7076, 1:3000).

PolyA + RNA purification

75 μ g of total RNA isolated using Trizol reagent were subjected to polyA+ RNA purification. Total RNA was denatured at 65 °C for 5 min, followed by standing on ice for 2 min. 200 μ l of oligo(dT)₂₅ magnetic beads (NEB, S1419S) were washed with 2 \times binding buffer (20 mM Tris-Cl, pH 7.5, 1 M LiCl, 2 mM EDTA), and resuspended with 100 ml of 2 \times binding buffer. Denatured total RNAs were mixed with washed beads,

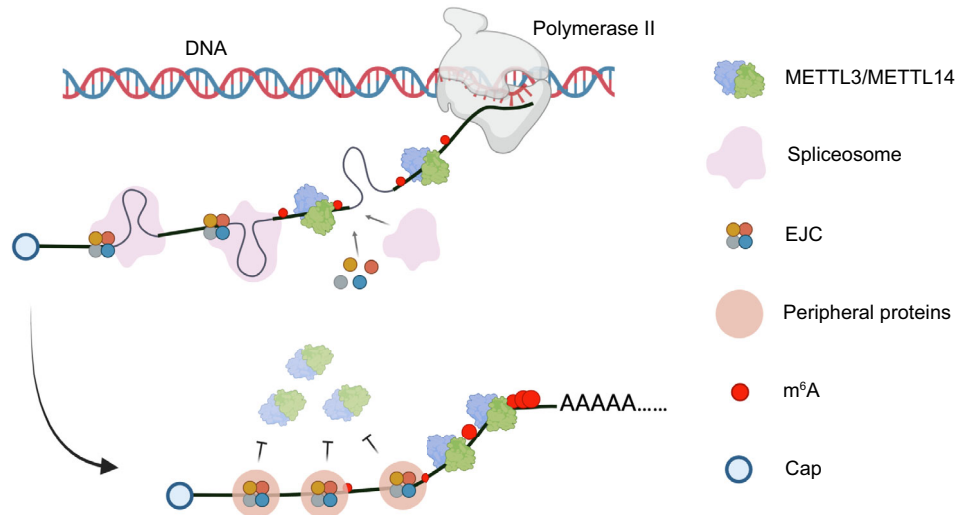


Fig. 5 | Blockade of METTL3 by the exon junction complex shapes the m⁶A epitranscriptome. Model for how the EJC locally controls METTL3-mediated m⁶A modification during mRNA processing. Created with BioRender.com.

and incubated for 15 min at R.T. with rotation. After discarding the unbound RNA, beads were washed twice with washing buffer (10 mM Tris-Cl, pH 7.5, 200 mM LiCl, 1 mM EDTA). Bound RNAs were eluted from beads with 100 μ l of TE buffer (5 mM Tris-Cl, pH 7.5, 1 mM EDTA). Binding and washing steps were repeated one more time, and elute the RNAs with 20 μ l of TE buffer.

Chromatin associated RNA (caRNA) extraction

Cells were lysed with cell lysis buffer (10 mM Tris-HCl, pH7.5, 150 mM NaCl, 0.05% NP-40 and proteinase inhibitor cocktail (Roche, 11836170001)) and incubated on ice for 10 min. The suspension was carefully added at the top of sucrose buffer (10 mM Tris-HCl, pH7.5, 150 mM NaCl, 24% sucrose), and centrifuged at 3200 \times g for 10 min. Nuclear pellets were resuspended in Glycerol buffer (20 mM Tris-HCl, pH7.5, 75 mM NaCl, 0.5 mM EDTA, pH8.0, 50% glycerol), and mixed with same volume nuclear lysis buffer (10 mM Tris-HCl, pH7.5, 300 mM NaCl, 7.5 mM MgCl₂, 0.2 mM EDTA, pH8.0, 1% NP-40, 1 M urea) on ice for 2 min. The lysate was centrifuged at 13,000 \times g for 2 min to precipitate the chromatin-RNA complex. The chromatin pellets were briefly rinsed with PBS-EDTA. RNA was extracted with TRIzol reagent (Invitrogen, 15596018). Ribosomal RNAs were removed using RiboMinus™ Eukaryote System v2 (ThermoFisher, A15026).

HPLC-MS/MS analysis

500 ng RNA was digested with 100 U S1 nuclease (ThermoFisher, EN0321) at 37 °C for 2 h and dephosphorylated with 1 U rSAP (NEB # M0371S) at 37 °C for 1 h. The 100 μ l samples were filtered with Millex-GV 0.22 μ filters (Millipore Sigma # SLGV033RS). A total of 5–10 μ l from each sample was injected into the Agilent 6470 Triple Quad LC/MS instrument with Agilent Zorbax Eclipse C18 reverse phase HPLC column. The samples were run at 500 μ l/min flow rate in mobile phase buffer A (water with 0.1% Formic Acid) and 0–20% gradient of buffer B (methanol with 0.1% formic acid). MRM transitions are measured for N⁶-methyl adenosine (282.1–150.1), and adenosine (268.1–136.1). Standard compounds for m⁶A (Cayman Chemical #16111) were run on HPLC-MS/MS to optimize HPLC method and determine retention times for each nucleoside. For HPLC-MS/MS data collection and analysis, Agilent Mass Hunter LC/MS Data Acquisition Version B.08.00 and Quantitative Analysis Version B.07.01 software was used.

Plasmid construction and transfection

Human FTO, ALKBH5, EIF4A3, CASC3, MAGOH, Y14, and ALYREF were cloned into pFlag-CMV2 vector. Mouse beta-Globin gene with/without

introns was cloned into pBi-luciferase vectors. Exons with hypermethylated m⁶A of selected genes were inserted into second exons of beta-Globin. 2 \times MS2 binding site sequence was PCR-amplified from FLuc-MS2bs and inserted into the exon1-exon2 junction region of mature beta-Globin. Splicing signals GU-AG in introns of beta-Globin were mutated into AC-TC to generate pre-Mut constructs. MS2 coat protein sequence was inserted into the N terminal of EIF4A3 or GFP to generate pFlag-CMV2-MS2-EIF4A3 or pFlag-CMV2-MS2-GFP. Plasmids were transfected using Lipofectamine 2000 transfection reagent (Invitrogen, 1166809). The siRNA duplexes (Dharmacon) were transfected into cells at a final concentration of 50 nM using Lipofectamine RNAiMax (Invitrogen, 13778150) following the manufacturer's instructions. Cells were harvested 48 h after transfection. All primers and siRNAs used in this study are listed in Supplementary Tables 2 and 4.

Protein complex purification for MS analysis

Protein complexes were isolated from HEK293T clones expressing FLAG-ALKBH5 or FLAG-FTO as described before⁵¹, with a few exceptions: whole cell extracts were prepared by adding lysis buffer (20 mM Tris-HCl (pH 8.0), 137 mM NaCl, 1 mM EDTA, 1% Triton X-100, 10% glycerol, 1.5 mM MgCl₂, 0.2 mM PMSF, 0.5 mM dithiothreitol (DTT) and 1 \times protease inhibitor cocktail) directly to the cells, incubating on ice for 15 min and centrifuging lysates at 20,000 \times g for 15 min in a microcentrifuge at 4 °C. Protein complexes were isolated as described before⁵¹, with a few exceptions: BC100 buffer (20 mM Tris-HCl (pH 7.8), 100 mM KCl, 0.2 mM EDTA, 10% glycerol, 10 mM β -mercaptoethanol (pH 7.8), 0.2% NP40), was used for all the washing steps instead of BC500 buffer. Protein samples were separated by SDS-PAGE and Coomassie-stained gel was submitted to the Taplin Mass Spectrometry Facility at Harvard Medical School for protein identification in gel sections. The gel bands were cut into approximately 1 mm³ pieces. The samples were reduced with 1 mM DTT for 30 min at 60 °C and then alkylated with 5 mM iodoacetamide for 15 min in the dark at room temperature. Gel pieces were then subjected to in-gel trypsin digestion procedure and then washed and dehydrated with acetonitrile for 10 min. followed by the removal of acetonitrile. Pieces were then completely dried in a speed-vac. Rehydration of the gel pieces was with 50 mM ammonium bicarbonate solution containing 12.5 ng/ μ l modified sequencing-grade trypsin (Promega, V5111) at 4 °C. Samples were then incubated at 37 °C overnight. Peptides were extracted by removing the ammonium bicarbonate solution, followed by one

wash step using 50% acetonitrile/1% formic acid. The extracts were then dried in a speed-vac for 1 h. Samples were reconstituted in 10 μ l of solvent A (2.5% acetonitrile, 0.1% formic acid). A nano-scale reverse-phase HPLC capillary column was created by packing 2.6 μ m C18 spherical silica beads into a fused silica capillary (100 μ m inner diameter; 30 cm length) with a flame-drawn tip. After equilibrating the column, each sample was loaded via a Famos autosampler (LC Packings) onto the column. Peptides were eluted using a linear gradient of solvent B (97.5% acetonitrile, 0.1% formic acid). As peptides eluted, they were subjected to electrospray ionization and then entered into an LTQ Orbitrap Velos Pro ion-trap mass spectrometer (ThermoFisher). Eluting peptides were detected, isolated, and fragmented to produce a tandem mass spectrum of specific fragment ions for each peptide. Peptide sequences (and hence protein identity) were determined by matching the human proteome database with the acquired fragmentation pattern by the software program Sequest (ThermoFisher). All databases include a reversed version of all the sequences and the data was filtered with a cutoff at 1% peptide false discovery rate.

Co-immunoprecipitation

HeLa cells with overexpressed Flag-tag proteins were homogenized in 1 ml of lysis buffer (50 mM Tris-HCl, pH 7.5, 150 mM NaCl, 2 mM EDTA, 1% NP-40, cComplete™ EDTA-Free Protease Inhibitor Cocktail) for 15 min on ice. Cells were further lysed by sonication with low energy (Bioruptor 30 s, on; 30 s off; 6 cycles). The supernatant was collected after centrifugation (14,000 *g*, 10 min). A total of 50 μ l of lysate was saved as input, 500 μ l of lysate was incubated with 30 μ l of anti-Flag M2 magnetic beads overnight at 4 °C with rotation. After washing with lysis buffer for 5 times, the proteins were eluted in 1 \times LDS loading buffer and resolved on SDS-polyacrylamide gels and detected by western blot analysis.

m⁶A RNA and caRNA immunoprecipitation and sequencing

1 μ g of PolyA⁺ RNAs and caRNAs was fragmented into ~150 nt by RNA Fragmentation reagent (Ambion, AM8740), and then incubated with 5 μ g of anti-m⁶A antibody (Abcam, ab151230) in 500 μ l of IP buffer (50 mM Tris, pH 7.4, 100 mM NaCl, 0.05% NP-40 and RNase inhibitor) at 4 °C for 4 h. Thirty microlitres of magnetic protein A beads (Thermo) were added to samples and incubated at 4 °C for 2 h. Beads were washed twice with high-salt buffer (50 mM Tris, pH 7.4, 1 M NaCl, 1 mM EDTA, 1% NP-40, 0.1% SDS), twice with IP buffer, and once with PNK wash buffer (20 mM Tris, pH 7.4, 10 mM MgCl₂, 0.2% Tween 20). After the PNK treatment, RNA was purified and subjected to library construction by SMARTer smRNA-Seq Kit for Illumina (Clontech).

eCLIP-seq

METTL3 eCLIP was performed as described previously⁵². In brief, UV crosslinked HeLa cells were collected and resuspended in 1 ml of lysis buffer (50 mM Tris-HCl pH 7.4, 100 mM NaCl, 1% NP-40, 0.1% SDS, 0.5% sodium deoxycholate, cComplete™ EDTA-Free Protease Inhibitor Cocktail (Roche, 11873580001) and 10 μ l of Murine RNase inhibitor (NEB, M0314L)) for 15 min on ice and then sonicated with low energy (Bioruptor 30 s, on; 30 s off; 6 cycles). After RNase-I (Ambion, AM2294) and DNase digestion for 5 min at 37 °C, lysates were centrifuged at 13,000 \times *g* for 15 min. Lysates were incubated with antibody-coupled sheep anti-rabbit IgG magnetic beads (30 μ l, ThermoFisher, 11203D) at 4 °C overnight. Two percent of lysate was saved size-matched input. Immunoprecipitated (IP) samples were washed with high salt buffer (50 mM Tris-HCl pH 7.4, 1 M NaCl, 1 mM EDTA, 1% NP-40, 0.1% SDS, 0.5% sodium deoxycholate in nuclease-free water) then with wash buffer (20 mM Tris-HCl pH 7.4, 10 mM MgCl₂, 0.2% Tween 20, in nuclease-free H₂O). 5' and 3' RNA ends were repaired with FastAP (Fermentas, EF0651) and T4 PNK

(NEB, M0201L), followed by RNA adapter ligation. Protein-RNA complexes were run on SDS-PAGE gels and transferred to nitrocellulose membranes, and RNA was isolated from the membranes with proteinase K digestion at 37 °C for 20 min then 50 °C for 20 min with interval mixing at 1200 rpm. RNA was cleaned and concentrated using Zymo RNA Clean & Concentrator kit (Zymo Research, R1015) and was reverse transcribed with AffinityScript enzyme (Agilent, 600107) at 54 °C for 20 min. After cDNA end repair, a 3' ssDNA adapter was ligated. Libraries were amplified according to the Ct values obtained. PCR conditions consisted of 98 °C (30 s) followed by 6 cycles of (98 °C (15 s), 70 °C (30 s), 72 °C (40 s)), then (Ct-5) cycles of (98 °C (15 s), 72 °C (45 s) and 72 °C (1 min)). Libraries were loaded into a 3% agarose gel and regions between 175–350 bp were extracted and purified using MinElute Gel Extraction Kit (Qiagen, 28604). The quantity and quality of the final libraries were assessed using a Bioanalyzer (Agilent Technology Inc). All samples were multiplexed and sequenced by dual indexed run (PE150) on the Illumina NovaSeq 6000 sequencer. Two biological replicates were conducted for each experiment.

Cross-Linking Immunoprecipitation (CLIP)

UV crosslinked HeLa cells transfected with reporter constructs were collected and resuspended in 1 ml of lysis buffer (50 mM Tris-HCl pH 7.4, 100 mM NaCl, 1% NP-40, 0.1% SDS, 0.5% sodium deoxycholate, cComplete™ EDTA-Free Protease Inhibitor Cocktail (Roche, 11873580001) and 10 μ l of Murine RNase inhibitor (NEB, M0314L)) for 15 min on ice and then sonicated with low energy (Bioruptor 30 s, on; 30 s off; 6 cycles). Lysates were centrifuged at 13,000 \times *g* for 15 min at 4 °C. 10% percent of lysate was saved as input. The remaining lysates were incubated with antibody-coupled sheep anti-rabbit IgG magnetic beads (30 μ l, ThermoFisher, 11203D) at 4 °C overnight. Immunoprecipitated (IP) samples were washed with high salt buffer (50 mM Tris-HCl pH 7.4, 1 M NaCl, 1 mM EDTA, 1% NP-40, 0.1% SDS, 0.5% sodium deoxycholate in nuclease free water) then with wash buffer (20 mM Tris-HCl pH 7.4, 10 mM MgCl₂, 0.2% Tween 20, in nuclease-free H₂O). IP RNA was isolated from the beads with proteinase K (NEB, P8107S) digestion at 37 °C for 20 min then 50 °C for 20 min with interval mixing at 1200 rpm. RNA was cleaned and concentrated using Zymo RNA Clean & Concentrator kit. RNA was extracted with TRIzol reagent (Invitrogen, 15596018).

Reverse-transcriptase qPCR

For RT-qPCR, cDNA was generated using Superscript III reverse transcriptase (ThermoFisher, 18080044) according to the manufacturer's instructions. qPCR assay was performed with Taq Pro Universal SYBR qPCR Master Mix (Vazyme, Q712) according to the manufacturer's instructions and quantified by the StepOnePlus™ Real-Time PCR System (Applied Biosystem, 4376600). All primers were listed in Supplementary Table 3.

MeRIP-seq analysis

Raw reads were trimmed using Cutadapt (2.5)⁵³ and aligned to hg38 genome using hisat2 (2.1.0)⁵⁴. Peaks were called using exomepeaks (2.16.0)⁵⁵. Transcript GTF file was built for peak identification on caRNA and polyA⁺ RNA. For differential m⁶A regions of polyA⁺ RNA upon EIF4A3 depletion, we utilized RADAR software (0.2.4)⁴⁴ with parameters "fragmentLength = 150, binSize = 50, minCountsCutOff = 15, cutoff = 0.1, Beta_cutoff = 0.5". Differential m⁶A regions with |foldchange| > 2 were kept. The distribution of the m⁶A peaks was analyzed using MeRIPtools⁵⁶. Peaks were annotated using annotatePeak⁵⁷. Genome coverage bedGraph files were generated by deeptools (3.0.2)⁵⁸ bamCoverage with the parameters "-normalizeUsing RPKM - binSize 5" and visualized using Integrative Genomics Viewer (IGV)⁵⁹. Aggregation plots were generated using deeptools⁵⁸ computeMatrix and plotProfile.

eCLIP-seq analysis

eCLIP-seq analysis was followed the pipeline⁵². Differential binding peaks were identified using Diffbind (3.6.1)⁶⁰. The distribution of the METTL3 binding peaks was analyzed using deeptools⁵⁸. Genome coverage bedGraph files were generated by deeptools⁵⁸ bamCoverage with the parameters “-normalizeUsing RPKM – binSize 5” and visualized using Integrative Genomics Viewer (IGV)⁵⁹. Aggregation plots were generated using deeptools⁵⁸ computeMatrix and plotProfile.

mRNA-seq analysis

Raw reads were trimmed using Cutadapt (2.5)⁵³ and aligned to hg38 genome using hisat2 (2.1.0)⁵⁴. Differential expression levels of exons and genes were analyzed using featureCounts (2.0.0)⁶¹ followed by DEseq2 (1.32.0)⁶². Differential expressed exons and genes with |fold-change| > 2 were kept. Gene ontology analysis was performed using ClusterProfiler (4.0.5)⁶³.

Statistics and reproducibility

The experiments in this study were performed with three biological replicates and are presented as the mean ± S.D. or mean ± S.E.M. calculated by GraphPad Prism 8.0. Two-tailed unpaired Student's *t*-test, Mann–Whitney U test, and Fisher's exact test were performed in GraphPad Prism 8.0 and R Bioconductor. *P* values < 0.05 were considered statistically significant. All statistical tests, resulting *P* values, and numbers of observations are indicated in figure panels or legend.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

High-throughput sequencing data have been deposited in the Gene Expression Omnibus (GEO) under the accession number [GSE207663](https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE207663). The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE partner repository⁶⁴ with the dataset identifier [PXID038705](https://www.ebi.ac.uk/pride/archive/study/PXD038705). Source data are provided with this paper.

References

- Roundtree, I. A., Evans, M. E., Pan, T. & He, C. Dynamic RNA modifications in gene expression regulation. *Cell* **169**, 1187–1200 (2017).
- Yang, X. et al. m⁶A promotes R-loop formation to facilitate transcription termination. *Cell Res.* **29**, 1035–1038 (2019).
- Choe, J. et al. mRNA circularization by METTL3-eIF3h enhances translation and promotes oncogenesis. *Nature* **561**, 556–560 (2018).
- Barbieri, I. et al. Promoter-bound METTL3 maintains myeloid leukaemia by m⁶A-dependent translation control. *Nature* **552**, 126–131 (2017).
- Xiao, W. et al. Nuclear m⁶A reader YTHDC1 regulates mRNA splicing. *Mol. Cell* **61**, 507–519 (2016).
- Lin, S., Choe, J., Du, P., Triboulet, R. & Gregory, R. I. The m⁶A methyltransferase METTL3 promotes translation in human cancer cells. *Mol. Cell* **62**, 335–345 (2016).
- Wang, X. et al. N⁶-methyladenosine modulates messenger RNA translation efficiency. *Cell* **161**, 1388–1399 (2015).
- Wang, Y. et al. N⁶-methyladenosine modification destabilizes developmental regulators in embryonic stem cells. *Nat. Cell Biol.* **16**, 191–198 (2014).
- Wang, X. et al. N⁶-methyladenosine-dependent regulation of messenger RNA stability. *Nature* **505**, 117–120 (2014).
- Wang, Y. et al. N⁶-methyladenosine RNA modification regulates embryonic neural stem cell self-renewal through histone modifications. *Nat. Neurosci.* **21**, 195–206 (2018).
- Zheng, G. et al. ALKBH5 is a mammalian RNA demethylase that impacts RNA metabolism and mouse fertility. *Mol. Cell* **49**, 18–29 (2013).
- Liu, J. et al. A METTL3-METTL14 complex mediates mammalian nuclear RNA N⁶-adenosine methylation. *Nat. Chem. Biol.* **10**, 93–95 (2014).
- Ping, X. L. et al. Mammalian WTAP is a regulatory subunit of the RNA N⁶-methyladenosine methyltransferase. *Cell Res.* **24**, 177–189 (2014).
- Zaccara, S., Ries, R. J. & Jaffrey, S. R. Reading, writing and erasing mRNA methylation. *Nat. Rev. Mol. Cell Biol.* **20**, 608–624 (2019).
- Lee, Y., Choe, J., Park, O. H. & Kim, Y. K. Molecular mechanisms driving mRNA degradation by m⁶A modification. *Trends Genet.* **36**, 177–188 (2020).
- Wang, P., Doxtader, K. A. & Nam, Y. Structural basis for cooperative function of Mettl3 and Mettl14 methyltransferases. *Mol. Cell* **63**, 306–317 (2016).
- Lin, Z. et al. Mettl3-/Mettl14-mediated mRNA N⁶-methyladenosine modulates murine spermatogenesis. *Cell Res.* **27**, 1216–1230 (2017).
- Xu, K. et al. Mettl3-mediated m⁶A regulates spermatogonial differentiation and meiosis initiation. *Cell Res.* **27**, 1100–1114 (2017).
- Yoon, K. J. et al. Temporal control of mammalian cortical neurogenesis by m⁶A methylation. *Cell* **171**, 877–889 (2017).
- Zhang, C. et al. m⁶A modulates haematopoietic stem and progenitor cell specification. *Nature* **549**, 273–276 (2017).
- Zhao, B. S. et al. m⁶A-dependent maternal mRNA clearance facilitates zebrafish maternal-to-zygotic transition. *Nature* **542**, 475–478 (2017).
- Huang, H., Weng, H. & Chen, J. m⁶A modification in coding and non-coding RNAs: roles and therapeutic implications in cancer. *Cancer Cell* **37**, 270–288 (2020).
- Meyer, K. D. et al. Comprehensive analysis of mRNA methylation reveals enrichment in 3' UTRs and near stop codons. *Cell* **149**, 1635–1646 (2012).
- Dominissini, D. et al. Topology of the human and mouse m⁶A RNA methylomes revealed by m⁶A-seq. *Nature* **485**, 201–206 (2012).
- He, P. C. & He, C. m⁶A RNA methylation: from mechanisms to therapeutic potential. *EMBO J.* **40**, e105977 (2021).
- Jia, G. et al. N⁶-methyladenosine in nuclear RNA is a major substrate of the obesity-associated FTO. *Nat. Chem. Biol.* **7**, 885–887 (2011).
- Ke, S. et al. m⁶A 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).
- Louloupi, A., Ntini, E., Conrad, T. & Orom, U. A. V. Transient N⁶-Methyladenosine Transcriptome Sequencing Reveals a Regulatory Role of m⁶A in splicing efficiency. *Cell Rep.* **23**, 3429–3437 (2018).
- Xu, W. et al. Dynamic control of chromatin-associated m⁶A methylation regulates nascent RNA synthesis. *Mol. Cell* **82**, 1156–1168 (2022).
- Huang, H., Weng, H. & Chen, J. The biogenesis and precise control of RNA m⁶A methylation. *Trends Genet.* **36**, 44–52 (2020).
- Chelmicki, T. et al. m⁶A RNA methylation regulates the fate of endogenous retroviruses. *Nature* **591**, 312–316 (2021).
- Xu, W. et al. METTL3 regulates heterochromatin in mouse embryonic stem cells. *Nature* **591**, 317–321 (2021).

33. Akhtar, J. et al. m⁶A RNA methylation regulates promoter- proximal pausing of RNA polymerase II. *Mol. Cell* **81**, 3356–3367 (2021).
34. Slobodin, B. et al. Transcription impacts the efficiency of mRNA translation via Co-transcriptional N6-adenosine methylation. *Cell* **169**, 326–337 (2017).
35. Bertero, A. et al. The SMAD2/3 interactome reveals that TGFbeta controls m⁶A mRNA methylation in pluripotency. *Nature* **555**, 256–259 (2018).
36. Fish, L. et al. Nuclear TARBP2 drives oncogenic dYSREgulation of RNA splicing and decay. *Mol. Cell* **75**, 967–981 (2019).
37. Huang, H. et al. Histone H3 trimethylation at lysine 36 guides m⁶A RNA modification co-transcriptionally. *Nature* **567**, 414–419 (2019).
38. Lan, T. et al. KIAA1429 contributes to liver cancer progression through N6-methyladenosine-dependent post-transcriptional modification of GATA3. *Mol. Cancer* **18**, 186 (2019).
39. Wei, J. et al. FTO mediates LINE1 m⁶A demethylation and chromatin regulation in mESCs and mouse development. *Science* **376**, 968–973 (2022).
40. Zhao, X. et al. FTO-dependent demethylation of N6-methyladenosine regulates mRNA splicing and is required for adipogenesis. *Cell Res.* **24**, 1403–1419 (2014).
41. Tang, C. et al. ALKBH5-dependent m6A demethylation controls splicing and stability of long 3'-UTR mRNAs in male germ cells. *Proc. Natl Acad. Sci. USA* **115**, E325–E333 (2018).
42. Le Hir, H., Sauliere, J. & Wang, Z. The exon junction complex as a node of post-transcriptional networks. *Nat. Rev. Mol. Cell Biol.* **17**, 41–54 (2016).
43. Singh, G. et al. The cellular EJC interactome reveals higher-order mRNP structure and an EJC-SR protein nexus. *Cell* **151**, 915–916 (2012).
44. Zhang, Z. et al. RADAR: differential analysis of MeRIP-seq data with a random effect model. *Genome Biol.* **20**, 294 (2019).
45. Sauliere, J. et al. CLIP-seq of eIF4AIII reveals transcriptome-wide mapping of the human exon junction complex. *Nat. Struct. Mol. Biol.* **19**, 1124–1131 (2012).
46. Obrdlik, A., Lin, G., Haberman, N., Ule, J. & Ephrussi, A. The transcriptome-wide landscape and modalities of EJC binding in adult drosophila. *Cell Rep.* **28**, 1219–1236 e1211 (2019).
47. Viphakone, N. et al. Co-transcriptional loading of RNA export factors shapes the human transcriptome. *Mol. Cell* **75**, 310–323 (2019).
48. Zhang, X. et al. An atomic structure of the human spliceosome. *Cell* **169**, 918–929 (2017).
49. Zhou, K. I. et al. Regulation of co-transcriptional pre-mRNA Splicing by m⁶A through the low-complexity protein hnRNPG. *Mol. Cell* **76**, 70–81 e79 (2019).
50. Muller-McNicoll, M. & Neugebauer, K. M. How cells get the message: dynamic assembly and function of mRNA-protein complexes. *Nat. Rev. Genet.* **14**, 275–287 (2013).
51. Gregory, R. I., Chendrimada, T. P. & Shiekhattar, R. MicroRNA biogenesis: isolation and characterization of the microprocessor complex. *Methods Mol. Biol.* **342**, 33–47 (2006).
52. Van Nostrand, E. L. et al. Robust transcriptome-wide discovery of RNA-binding protein binding sites with enhanced CLIP (eCLIP). *Nat. Methods* **13**, 508–514 (2016).
53. Marcel, M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet. J.* **17**, 10–12 (2011).
54. Kim, D., Langmead, B. & Salzberg, S. L. HISAT: a fast spliced aligner with low memory requirements. *Nat. Methods* **12**, 357–360 (2015).
55. Meng, J. et al. A protocol for RNA methylation differential analysis with MeRIP-Seq data and exomePeak R/Bioconductor package. *Methods* **69**, 274–281 (2014).
56. Zhang, Z. et al. Genetic analyses support the contribution of mRNA N6-methyladenosine (m⁶A) modification to human disease heritability. *Nat. Genet.* **52**, 939–949 (2020).
57. Heinz, S. et al. Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. *Mol. Cell* **38**, 576–589 (2010).
58. Ramirez, F. et al. deepTools2: a next generation web server for deep-sequencing data analysis. *Nucleic Acids Res.* **44**, W160–W165 (2016).
59. Thorvaldsdottir, H., Robinson, J. T. & Mesirov, J. P. Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. *Brief. Bioinform.* **14**, 178–192 (2013).
60. Ross-Innes, C. S. et al. Differential oestrogen receptor binding is associated with clinical outcome in breast cancer. *Nature* **481**, 389–393 (2012).
61. Liao, Y., Smyth, G. K. & Shi, W. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* **30**, 923–930 (2014).
62. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 550 (2014).
63. 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).
64. Perez-Riverol, Y. et al. The PRIDE database resources in 2022: a hub for mass spectrometry-based proteomics evidences. *Nucleic Acids Res.* **50**, D543–D552 (2022).

Acknowledgements

R.I.G. is supported by an Outstanding Investigator Award (R35CA232115) and R01 (R01CA233671) from the National Cancer Institute (NCI) of the NIH.

Author contributions

X.Y., R.T., and R.I.G. conceived the project. X.Y. performed most of the experiments with contributions from R.T., X.Y., and Q.L. performed bioinformatics analysis of sequencing data. E.S. performed RNA mass spectrometry. R.I.G. supervised the project. All authors organized and analyzed data. X.Y. and R.I.G. wrote the manuscript with input from all authors.

Competing interests

R.I.G. is a co-founder, scientific advisory board member, and equity holder of 28/7 Therapeutics and Theonys. All other authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41467-022-35643-1>.

Correspondence and requests for materials should be addressed to Richard I. Gregory.

Peer review information *Nature Communications* thanks Ulf Ørom and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

Reprints and permissions information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2022