DART-seq: an antibody-free method for global m6A detection

Kate D. Meyer 1,2

6-methyladenosine (m6A) is a widespread RNA modification that influences nearly every aspect of the messenger RNA lifecycle. Our understanding of m6A has been facilitated by the development of global m6A mapping methods, which use antibodies to immunoprecipitate methylated RNA. However, these methods have several limitations, including high input RNA requirements and cross-reactivity to other RNA modifications. Here, we present DART-seq (deamination adjacent to RNA modification targets), an antibody-free method for detecting m6A sites. In DART-seq, the cytidine deaminase APOBEC1 is fused to the m6A-binding YTH domain. APOBEC1-YTH expression in cells induces C-to-U deamination at sites adjacent to m6A residues, which are detected using standard RNA-seq. DART-seq identifies thousands of m6A sites in cells from as little as 10 ng of total RNA and can detect m6A accumulation in cells over time. Additionally, we use long-read DART-seq to gain insights into m6A distribution along the length of individual transcripts.

The most abundant internal mRNA modification, N6-methyladenosine (m6A), plays diverse roles in RNA regulation. Recently, m6A has emerged as an important regulator of a variety of physiological processes; thus, detecting m6A sites in cells is critical for understanding how this modification impacts gene expression to contribute to cellular function and disease states.

To date, most methods for global m6A detection have relied on immunoprecipitation of methylated RNAs using m6A-recognizing antibodies in techniques such as MeRIP-seq or m6A-seq. Subsequent improvements to this method have come with the addition of ultraviolet crosslinking steps to identify m6A sites at single-nucleotide resolution. Although these methods have yielded unprecedented insights into the location and regulation of m6A in cellular RNAs, they suffer from several limitations. First, they require large amounts of input RNA, which makes global m6A detection prohibitive for limited-quantity samples. Second, m6A antibodies also recognize the structurally similar cap modification, m7Am, so immunoprecipitation of methylated RNAs does not exclusively enrich for m6A-containing RNA. Finally, antibody-based approaches are costly and the associated library preparation steps are time consuming, which can be major limiting factors for many experiments. Thus, there is a great need for a simple, sensitive, antibody-free method for global m6A detection.

We reasoned that a strategy that alters the sequence near methylation sites would enable m6A detection by standard RNA-seq and thus overcome the major limitations of current methods. APOBEC1 is a cytosine deaminase that targets DNA and RNA to induce cytosine-to-uracil (C-to-U) editing. Although initially discovered for its ability to edit the ApoB mRNA, APOBEC1 has since been utilized in CRISPR-Cas9-based genome editing approaches to induce C-to-U conversion at targeted single-stranded DNA sites. We speculated that a similar strategy could be used to edit m6A-adjacent cytidines in RNAs by fusing APOBEC1 to the m6A-binding YTH domain and detecting subsequent editing events with RNA-seq. Here, we demonstrate the utility of this approach for detecting m6A sites in cellular RNAs using transcriptome-wide mapping with as little as 10 ng of total RNA as input. Our strategy performs similarly to antibody-based approaches for methylated RNA detection and provides insights into clustering of m6A residues within individual transcript isoforms. This approach substantially improves the time and cost associated with global m6A detection and will enable transcriptome-wide mapping in limited RNA samples.

Results

Development of an antibody-free method for m6A detection. The preferred consensus sequence for m6A contains an invariable cytidine residue immediately following the m6A site (Rm6ACH, where R = A or G; H = A, C or U)1-4. Thus, we speculated that recruitment of APOBEC1 to m6A sites would enable deamination of the cytidine immediately following m6A residues. To test this, we fused APOBEC1 to the m6A-binding YTH domain of YTHDF2 (refs. 13,14; Fig. 1). The APOBEC1-YTH fusion protein was then incubated with a synthetic RNA containing a single internal adenosine. Reverse transcription and Sanger sequencing indicated frequent C-to-U mutation detection (Fig. 1, Methods). Comparison of APOBEC1-YTH in the absence and presence of a methylated RNA revealed preferential C-to-U conversion at single-stranded DNA sites. Next, we tested the deamination activity of APOBEC1-YTH on unmethylated RNA (Supplementary Fig. 1a). To confirm that the observed editing was caused by targeting of APOBEC1 to the m6A residue, we repeated the in vitro deamination assays using a mutant version of the APOBEC1-YTH fusion protein (APOBEC1-YTHmut) in which the m6A binding region of the YTH domain was deleted (Supplementary Fig. 1b). APOBEC1-YTHmut was impaired in its ability to bind m6A (Supplementary Fig. 1c) and failed to convert adjacent cytidines to uridines in m6A-containing RNA (Supplementary Fig. 1a), indicating that the deaminase activity of APOBEC1-YTH is directed by the m6A-binding activity of the YTH domain.

DART-seq enables transcriptome-wide detection of m6A. We next sought to determine whether APOBEC1-YTH could be used to detect endogenous m6A sites in cells. To do this, we developed DART-seq (deamination adjacent to RNA modification targets), in which we transiently transfected APOBEC1-YTH into cells and then subjected total RNA to high-throughput sequencing followed by C-to-U mutation detection (Fig. 1, Methods). Comparison to antibody-based approaches for methylated RNA detection and provides insights into clustering of m6A residues within individual transcript isoforms. This approach substantially improves the time and cost associated with global m6A detection and will enable transcriptome-wide mapping in limited RNA samples.

1Department of Biochemistry, Duke University School of Medicine, Durham, NC, USA. 2Department of Neurobiology, Duke University School of Medicine, Durham, NC, USA. e-mail: kate.meyer@duke.edu

https://doi.org/10.1038/s41592-019-0570-0
Development of a targeted deamination strategy to detect m6A.

Schematic of the DART-seq method. APOBEC1 is fused to the YTH domain to guide C-to-U editing at cytidine residues adjacent to m6A sites. APOBEC1-YTH is expressed in cells and total RNA is isolated and subjected to RNA-seq. C-to-U mutations are then detected to identify sites of m6A.

Validation of the DART-seq approach. To validate individual DART-seq sites, we performed PCR with reverse transcription (RT–PCR) and Sanger sequencing to determine whether C-to-U editing events were detected in cells expressing APOBEC1 alone (Supplementary Fig. 3a). Together, these results suggest that the specificity of APOBEC1-YTH editing throughout the transcriptome depends on its ability to bind m6A.

To obtain a set of high-confidence editing sites, we filtered our list of APOBEC1-YTH sites to include only those with at least a 1.5-fold enrichment over APOBEC1-YTHmut samples. We also excluded all naturally occurring C-to-U mutations in HEK293T cells, as well as C-to-U editing events detected in cells expressing APOBEC1 alone (see Methods). This resulted in a list of 100,636 C-to-U editing sites in 9,793 RNAs that occurred in at least 5% of all reads. Of these, a stringent list of 40,263 editing events in 7,707 RNAs was observed in at least 10% of all reads (Supplementary Table 2).

Examination of sequences immediately surrounding DART-seq sites revealed enrichment of a GGACU-containing motif, which matches the preferred consensus sequence for m6A (Fig. 2c). In contrast, motifs detected in APOBEC1-YTHmut and APOBEC1 samples did not match the m6A consensus (Fig. 2c). Furthermore, DART-seq sites were highly enriched within 3′ UTRs and in the vicinity of the stop codon, as well as within long internal exons (Fig. 2b and Supplementary Fig. 3a,d), which matches the distribution of m6A. Comparison of methylated RNAs detected by MeRIP–seq and those identified by DART-seq showed a high degree of overlap, with 64% of m6A-containing RNAs detected by DART-seq (3,679 of 5,768 RNAs). Examination of individual RNAs showed that DART-seq editing events occurred at sites of MeRIP–seq enrichment (Fig. 2a). Furthermore, consistent with our in vitro deamination assays, we found that C-to-U editing events frequently occurred immediately downstream of known m6A sites in cellular RNAs (Fig. 2c).

We next assessed the ability of DART-seq to identify individual m5A sites compared to antibody-based approaches. Comparison of global m5A profiling datasets obtained by m5A immunoprecipitation showed that DART-seq performs similarly in its ability to detect m5A sites (Supplementary Fig. 5). We also observed an enrichment of DART-seq editing adjacent to m5A sites identified by single-nucleotide resolution m5A profiling (miCLIP/m5A-seq) (Supplementary Fig. 6a,b). Additionally, the majority (91.4%) of C-to-U editing sites in APOBEC1-YTH-expressing cells are preceded by an A, compared to only 67.9% of C-to-U editing sites in cells expressing APOBEC1 alone (Supplementary Fig. 6c), suggesting that APOBEC1-YTH deamination is directed specifically toward cytidines adjacent to m5A and that promiscuous editing of nonadjacent cytidines is rare. Further support for this comes from the finding that over 90% of DART-seq sites are greater than 10 nucleotides (nt) away from the closest editing event, which is similar to the distribution seen in miCLIP (Supplementary Fig. 5c). Collectively, these data indicate that DART-seq is capable of detecting m5A sites in cellular RNAs transcriptome wide.

Validation of the DART-seq approach. To validate individual DART-seq sites, we performed PCR with reverse transcription (RT–PCR) and Sanger sequencing to determine whether C-to-U editing occurs adjacent to m5A sites previously quantified by miCLIP or by SCARLET, another single-nucleotide resolution m5A identification method. We confirmed the presence of editing adjacent to known m5A sites in the BSG and ACTB mRNAs in cells expressing APOBEC1-YTH but did not observe robust editing in cells expressing APOBEC1 alone (Supplementary Fig. 7). To further validate that DART-seq editing depends on the presence of m5A, we

of three biological replicates indicated high reproducibility in C-to-U mutations in APOBEC1-YTH-expressing HEK293T cells (Supplementary Fig. 2), suggesting that APOBEC1-YTH targets specific RNAs for editing with high consistency across samples.

To determine whether DART-seq can identify m5A residues, we compared C-to-U editing sites from cells expressing APOBEC1-YTH and cells expressing APOBEC1 alone. DART-seq editing events from APOBEC1-YTH-expressing cells occurred primarily in the 3′ untranslated region (UTR) and coding sequence (CDS) (Fig. 2a) and were enriched in the vicinity of the stop codon (Fig. 2b and Supplementary Fig. 3a), which mirrors the distribution of m5A (refs. 34). In contrast, editing events from cells expressing APOBEC1 alone were located primarily in 3′ UTRs and intergenic regions and failed to show an enrichment near the stop codon (Fig. 2b and Supplementary Fig. 3b). Furthermore, there was little overlap in C-to-U editing between the two datasets, as 96% of edited sites from APOBEC1-YTH-expressing cells were not detected in cells expressing APOBEC1 alone (56,603 out of 59,246 sites) (Supplementary Table 1). To further ensure that C-to-U editing was caused by recruitment of APOBEC1-YTH to m5A, we performed RNA-seq on HEK293T cells expressing APOBEC1-YTHmut and carried out the same C-to-U editing analysis (Supplementary Table 1). C-to-U editing events in cells expressing APOBEC1-YTHmut showed a distinct distribution compared to those in cells expressing APOBEC1-YTH, characterized by an enrichment throughout the 3′ UTR as opposed to in the vicinity of the stop codon (Fig. 2a and Supplementary Fig. 3a,c). Together, these results suggest that the specificity of APOBEC1-YTH editing throughout the transcriptome depends on its ability to bind m5A.

Validation of the DART-seq approach. To validate individual DART-seq sites, we performed PCR with reverse transcription (RT–PCR) and Sanger sequencing to determine whether C-to-U editing occurs adjacent to m5A sites previously quantified by miCLIP or by SCARLET, another single-nucleotide resolution m5A identification method. We confirmed the presence of editing adjacent to known m5A sites in the BSG and ACTB mRNAs in cells expressing APOBEC1-YTH but did not observe robust editing in cells expressing APOBEC1 alone (Supplementary Fig. 7). To further validate that DART-seq editing depends on the presence of m5A, we
performed DART-seq using HEK293T cells depleted of the m^6^A methyltransferase, METTL3 (Supplementary Fig. 8). METTL3-depleted cells exhibited fewer DART-seq editing events in general and loss of the GGACU m^6^A consensus sequence surrounding DART-seq sites (Supplementary Fig. 9a and Supplementary Table 3). Furthermore, 97% of the DART-seq sites detected in wild-type cells were lost in METTL3-depleted cells (Supplementary Fig. 9b–e). These results further confirm that DART-seq editing depends on m^6^A.

**DART-seq enables low-input global m^6^A profiling.** One of the biggest challenges for global m^6^A detection has been the large amount of input RNA required for effective immunoprecipitation and sequencing. Recent advances in library preparation have provided important improvements, with some studies reporting m^6^A profiling using as little as 150 ng of mRNA or 500 ng of total RNA^17^-19. However, even with such improvements, the requirement for high nanogram amounts of poly(A) or ribosomal RNA-depleted RNA can be limiting for certain cell or tissue types.

We therefore sought to determine whether DART-seq could be used to detect m^6^A in low-input RNA samples. Using as little as 10 ng of total RNA as input, we detected over 79% of the DART-seq edited mRNAs that were identified in our high input DART-seq library (Supplementary Fig. 10a,b and Supplementary Table 4).

Low-input DART-seq samples perform similarly to antibody-based approaches for m^6^A detection, albeit with slightly reduced efficiency compared to high input DART-seq samples (Supplementary Fig. 10c). In addition, low-input DART-seq sites are enriched for m^6^A consensus motifs and near the 5’ end of the 3’ UTR (Supplementary Fig. 10d,e). Thus, DART-seq is capable of detecting m^6^A sites from as little as 10 ng of total RNA.

**m^6^A detection using in vitro DART-seq.** APOBEC1-YTH-expressing cells exhibit normal levels of genes in the m^6^A regulatory pathway and show no alterations in cell viability, suggesting that prolonged APOBEC1-YTH expression does not alter the m^6^A landscape (Supplementary Fig. 11 and Supplementary Table 5).

Nevertheless, APOBEC1-YTH overexpression may not be possible or desirable in some cases, which would necessitate the use of in vitro deamination to perform DART-seq. To test the ability of this approach to detect m^6^A in cellular RNA, we performed in vitro
DART-seq using HEK293T cell RNA. We detected C-to-U editing at known m^6A sites, and global analyses revealed a distribution and motif enrichment similar to that of m^6A (Supplementary Figs. 12 and 13). Although the majority (91%) of methylated mRNAs identified with in vitro DART-seq were also identified using cellular DART-seq, in vitro DART-seq identified fewer methylated mRNAs than cellular DART-seq, suggesting reduced efficiency (Supplementary Table 6). Thus, in vitro DART-seq can reliably mark m^6A sites in cellular RNAs, although this approach will likely benefit from further optimization to increase identification of low-abundance m^6A sites.

DART-seq distinguishes m^6A from m^6Am. A limitation of antibody-based m^6A detection strategies is cross-reactivity of m^6A antibodies with m^6Am. Hydrogen bonding between the YTH domain and the 2'-OH of m^6A suggests that the YTH domain used in DART-seq may not recognize m^6Am. Furthermore, unlike m^6A residues, m^6Am is not invariably following by a cytidine.

**Fig. 3 | DART-seq monitors changes in m^6A over time.**

- **a**, Distribution of C-to-U mutations discovered by DArT-seq in HeK293T cells treated with CPT (n = 5,689) compared to untreated controls (UT; n = 40,594). **P < 0.0001.**
- **b**, IGV browser images showing C-to-U mutations within the CDS of the BPTF (b) and ATRX (c) transcripts following CPT treatment. C-to-U mutations found in at least 10% of reads are indicated by blue/red (C/U) coloring in the BPTF transcript (positive strand) and by gold/green (C/U) coloring in the ATRX transcript (negative strand). C-to-U sites enriched after CPT treatment are indicated by purple triangles. n = 2 independent samples. kb, kilobase.
- **d**, MeRIP–RT–qPCR confirms enrichment of m^6A in the BPTF and ATRX mRNAs following CPT treatment. n = 2 biological replicates; the box plot indicates mean and upper/lower limits. IP, immunoprecipitated fraction.
- **e**, RT–qPCR analysis using RNA from untreated and CPT-treated cells shows a decrease in abundance of the BPTF and ATRX transcripts following CPT treatment. n = 2 biological replicates; the box plot indicates mean and upper/lower limits.
which means that detection of m6Am by APOBEC1-YTH would require deamination of cytidines further away from the modified base. We therefore wondered whether DART-seq could be used to distinguish m6A from m6Am.

To investigate this, we compared a list of m6Am sites in HEK293 cells to DART-seq datasets. Since m6A sites in 5′ UTRs may actually reflect m6Am residues at misannotated start sites, we extended our DART-seq sites to include regions 4 nt up- and downstream from the C-to-U editing site. We found only one RNA out of 3,431 with overlap between extended DART-seq sites and m6Am sites. On closer examination, the DART-seq editing site in this transcript is diminished in METTL3-depleted cells and is found internally within the 5′ UTR, suggesting that it is not an m6Am site (Supplementary Fig. 14). Thus, we conclude that DART-seq does not recognize m6Am and can be used to identify m6A residues independently of m6Am residues.

**Estimation of m6A abundance.** Determining m6A abundance within individual RNAs has been a major challenge to RNA methylation research. SCARLET enables quantitative measures of m6A in individual RNAs, but this approach is not amenable to transcriptome-wide measurements. m6A-LAIC-Seq uses immunoprecipitation of full-length transcripts to estimate methylation levels of individual mRNAs, but it does not account for multiple m6A sites or the presence of m6Am. Finally, peak over input can be used in MeRIP-seq, but these measures provide only a rough estimate of m6A abundance.

We speculated that the degree of methylation may correlate with APOBEC1-YTH binding and C-to-U editing to enable global estimates of m6A abundance in individual transcripts. To test this, we performed in vitro deamination assays using RNA with various amounts of m6A. We found that C-to-U editing was positively correlated with m6A levels at individual sites within an RNA (Supplementary Fig. 15a,b). Examination of DART-seq editing of cellular RNAs also showed a positive relationship between m6A abundance and editing efficiency (Supplementary Fig. 7a) Thus, DART-seq can be used as an indicator of m6A abundance in individual RNAs.

**DART-seq identifies m6A accumulation in cellular RNAs.** We next sought to determine whether changes in m6A can be detected by DART-seq. Previous studies have shown that treatment of cells with moderate concentrations of the topoisomerase inhibitor camptothecin (CPT) causes slowed transcription and an increase in m6A abundance in the CDS. We treated HEK293T cells expressing APOBEC1-YTH with CPT for 5 h and performed DART-seq to identify m6A sites. This led to 6,258 C-to-U sites that showed at least a twofold increase in editing compared to untreated cells (Supplementary Table 7). Metagene analysis of these sites indicated a slight enrichment in the CDS compared to untreated cells (Fig. 3a), and examination of individual mRNAs confirmed this analysis (Fig. 3b, c). We validated the increase in m6A within the CDS of select mRNAs using m6A immunoprecipitation followed by quantitative PCR with reverse transcription (MeRIP–RT–qPCR) (Fig. 3d). We also found that increased m6A in these RNAs negatively correlated with their abundance (Fig. 3e), similar to what has been previously observed. Thus, DART-seq can be used to detect accumulation of m6A in individual RNAs in response to changing cellular conditions.

**Long-read DART-seq reveals isoform-specific methylation patterns.** Immunoprecipitation-based m6A detection strategies have previously reported clustering of m6A sites. However, it remains unknown whether this reflects clustering of m6A on the same or distinct RNA molecules. Since DART-seq induces editing events in single transcripts, we reasoned that individual sequencing reads could be examined to determine whether m6A sites are found in the same RNA molecule. To investigate this, we performed long-read DART-seq using the PacBio platform. Examination of individual mRNAs showed that, although some transcripts exhibit isoform-specific regional editing, others contain DART-seq sites in the 5′ UTR, CDS and 3′ UTR (Fig. 4). In addition, 41% of reads spanning at least two

---

**Fig. 4 | Long-read DART-seq reveals m6A distribution within individual RNA molecules.** IGV browser tracks showing PacBio DART-seq data at two representative mRNAs (PABPC1, top and RCC1, bottom). C-to-U mutations found in at least 10% of reads are indicated by blue/red (C/U) coloring in the PABPC1 transcript (positive strand) and by gold/green (C/U) coloring in the RCC1 transcript (negative strand). Boxes show expanded views of C-to-U mutations within the same read spanning the 5′ UTR, CDS and 3′ UTR. n = 2 independent samples. bp, base pairs.
editing sites contain two or more C-to-U editing events. These data suggest that the majority of individual RNA molecules have just one m^6A site, but that many RNAs harbor multiple sites, which is consistent with previous reports from isoform-specific m^6A immunoprecipitation (m^6A-LAIC-Seq). Further studies will be needed to understand whether distinct m^6A residues on the same transcript work in a coordinated or competing manner. Additionally, although our data suggest that multiple C-to-U editing events caused by the same m^6A site are rare (Supplementary Fig. 5c and Supplementary Fig. 6), studies of clustered m^6A sites in individual transcripts may benefit from additional validation using miscLIP or SCARLET.

Discussion

DART-seq provides a novel strategy for global m^6A detection that overcomes many of the limitations of current approaches. A major advantage of DART-seq is its ability to map m^6A sites from low quantities of RNA, and we envision that DART-seq can potentially be coupled with single-cell isolation and library preparation methods to achieve single-cell m^6A detection. Although DART-seq is presented here using transient transfection of APOBEC1-YTH into mammalian cells, genomic integration or viral-mediated delivery into a cell population of interest will likely improve m^6A detection even further. In addition, in vitro DART-seq provides versatility to the approach by enabling m^6A detection without the need for APOBEC1-YTH expression in cells. In its current form, in vitro DART-seq shows reduced sensitivity for low-abundance m^6A sites, but this limitation is likely to be improved through purification of the DART-seq fusion proteins.

We find that DART-seq marks more sites than antibody-based approaches, which could speak to the differences between the two techniques: while m^6A immunoprecipitation takes a snapshot of m^6A at any given time and is prone to accessibility of the antibody to m^6A during immunoprecipitation, DART-seq irreversibly marks m^6A sites in cells over several hours. Thus, sites that would be otherwise buried in structure may be briefly accessible under physiological conditions and edited by APOBEC1-YTH. A potential future application of DART-seq is to express APOBEC1-YTH in animals to enable m^6A profiling in cell types of interest during distinct physiological states.

Modifications to the YTH domain that improve its affinity for m^6A will likely enable more precise and sensitive m^6A detection using the DART-seq approach. Similarly, DART-seq can potentially be used to detect other RNA modifications by fusing APOBEC1 to small proteins engineered to bind them. Finally, fusion of localization elements to APOBEC1-YTH could potentially facilitate compartmentalized m^6A detection. For instance, forcing a nuclear, cytoplasmic or mitochondrial localization could enable selective detection of methylated RNAs residing in distinct cellular compartments.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, statements of code and data availability and associated accession codes are available at https://doi.org/10.1038/s41592-019-0570-0.

Acknowledgements

We thank S.H. Choi for generating METTL3-depleted cells. We also thank S. Horner and members of the Meyer laboratory for comments and suggestions. This study was supported by the National Institutes of Health grant nos. RO0MH104742, R01MH118366 and DP1DA046584. K.D.M. is also supported by the Rita Allen Foundation, the Kinship Foundation Sears Scholars Program and the Klingenstein–Simons Fellowship in Neurosciences.

Author contributions

K.D.M. conceived of the project, collected and analyzed the data, and wrote the manuscript.

Competing interests

The author declares no competing interests.

Additional information

Supplementary information is available for this paper at https://doi.org/10.1038/s41592-019-0570-0.

Correspondence and requests for materials should be addressed to K.D.M.

Peer review information

Nicole Rusk was the primary editor on this article and managed its editorial process and peer review in collaboration with the rest of the editorial team.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher’s note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.
Methods

Antibodies. The following antibodies and concentrations were used: rabbit anti-
HA (Cell Signaling; 3724S; 1:1,000), rabbit anti-m6A (Abcam; ab151230; 1:1,000),
horseradish peroxidase (HRP)-conjugated goat anti-rabbit (Abcam; ab67621;
1:2,500). HRP-conjugated mouse anti-β-actin (GE Healthcare; 90107-554; 1:2,500),
mouse anti-β-actin (Genscript; A00702; 1:5,000), rabbit anti-METTL3 (Abcam;
ab195352; 1:1,000), rabbit anti-cleaved caspase 3 (Proteintect; 25546-1-AP;
1:1,000), AlexaFluor488-conjugated goat anti-rabbit (Thermo-Fisher;
A-21206; 1:1,000).

Constructs. YTH-HA was synthesized as a gene fragment (IDT), and YTH-HA
or YTHmut was subsequently amplified using the YTH Fwd/YTH-HA Rev or
YTHmut Fwd/YTH-HA Rev primers (see section ‘Primers (5′−3′)’). The YTH-
HA sequence comprised amino acids 385−579 of human YTHFD2 fused at its
C-terminal end to the HA tag (YPFDVPDYA). The YTHmut-HA fusion lacked
amino acids 385−499 comprising the m6A binding region. These YTH-HA fusions
were then inserted downstream of rat APOBEC1 in the pCMV-BE1 plasmid
(a gift from D. Liu; Addgene plasmid no. 73019; http://n2t.net/addgene:73019)
were then inserted downstream of rat APOBEC1 in the pCMV-BE1 plasmid
(a gift from F. Zhang; Addgene plasmid no. 62988; http://
NATURE METHODS
ARTICLES
Articles Nature Methods

Articles

hg19 annotation. Sequencing reads spanning individual exons were processed by Exon length was determined using the RefSeq

analysis using a m/k filtering threshold of 5–60%.

sites. In vitro DART-seq datasets were also subjected to the same C-to-U mutation

the reference sample. PacBio datasets were aligned to the human genome (hg19)

was used to find sites that were of the indicated fold enrichment greater than

further processed by removing sites detected in cells expressing APOBEC1 alone.

For determining enrichment of C-to-U editing between samples, a filter of m/k

had a minimum of two mutations, at least ten reads per replicate and a mutation/

are immediately preceded by an A; however, this could potentially exclude some

If desired, sites can be further filtered to include only C-to-U editing events that

number of editing events within these reads was then counted and summed.

Dataset comparisons. DART-seq and m6A immunoprecipitation (MeRIP–seq, miCLIP) datasets were analyzed using the ‘closest’ and ‘intersect’ features of the bedtools suite35. For comparison of methylated mRNAs, bed file coordinates were annotated using the annotation feature of the metagene analysis pipeline (see section ‘Metagene and motif analyses’) to give individual mRNAs, which were then compared between datasets.

Statistics. Statistical analysis of cell viability and western blot data were performed using a two-tailed t-test. Analysis of C-to-U editing enrichment in various transcript regions following CPT treatment, as well as analysis of the proportion of C-to-U sites following each nucleotide, was performed using a chi-squared test.

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

Data availability

The data that support the findings of this study have been deposited in NCBI's Gene Expression Omnibus under accession code GSE125780.

References

25. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 15, 530 (2014).

26. Shah, A., Qian, Y., Weyn-Vanhentenryck, S. M. & Zhang, C. CLIP Tool Kit (CTK): a flexible and robust pipeline to analyze CLIP sequencing data. Bioinformatics 33, 566–567 (2017).

27. Wu, T. D. & Watanabe, C. K. GMAP: a genomic mapping and alignment program for mRNA and EST sequences. Bioinformatics 21, 1859–1875 (2005).

28. Olarerin-George, A. O. & Jaffrey, S. R. MetaPlotR: a Perl/R pipeline for plotting metagenes of nucleotide modifications and other transcriptomic sites. Bioinformatics 33, 1563–1564 (2017).

29. 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).

30. Lindenbaum, P. JVarkit: java-based utilities for bioinformatics. Figshare https://doi.org/10.6084/m9.figshare.1425030.v1 (2015).

31. Quinlan, A. R. & Hall, I. M. BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics 26, 841–842 (2010).
Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see Authors & Referees and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

- The exact sample size (n) for each experimental group/condition, given as a discrete number and unit of measurement
- A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
- The statistical test(s) used AND whether they are one- or two-sided
- Only common tests should be described solely by name; describe more complex techniques in the Methods section.
- A description of all covariates tested
- A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
- A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
- For null hypothesis testing, the test statistic (e.g. F, t, r) with confidence intervals, effect sizes, degrees of freedom and P value noted Give P values as exact values whenever suitable.
- For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
- For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
- Estimates of effect sizes (e.g. Cohen’s d, Pearson’s r), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection

No software was used for data collection.

Data analysis

C to U mutation sites in sequencing data were identified using the CIMS package, which is part of the CLIP toolkit (CTK) suite of programs (version 1.1.2) and which is freely available: https://zhanglab.c2b2.columbia.edu/index.php/CTK_Documentation.

In addition, analysis of C to U editing sites on individual reads (PacBio and Illumina HiSeq4000) was done using the Sam2TSV code, which is also freely available: http://lindenb.github.io/jvarkit/

Analysis of motifs was performed using the findMotifsGenome.pl program of the HOMER suite (v4.9.1):

http://homer.ucsd.edu/homer/motif

Other analysis and processing steps used the following programs and versions: bedtools suite (v.2.19.1), samtools (v1.4), Novoalign (v3.09.00)

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

Data were deposited in NCBI's Gene Expression Omnibus (GEO) under accession number GSE125780.
Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

- [x] Life sciences
- [ ] Behavioural & social sciences
- [ ] Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

**Sample size**

No statistical analysis was performed to determine sample size for next-generation sequencing samples; however, for the main findings, we performed a minimum of 3 experiments (control versus sample condition). Supporting experiments were performed using a minimum of two biological replicate samples for each condition. These sample sizes were chosen based on standard practice for RNA-seq datasets as well as the number of samples needed to determine sample to sample variability.

**Data exclusions**

Data were not excluded.

**Replication**

Multiple replicate experiments were performed and all were reproducible and supported the central findings of the study.

**Randomization**

Next-generation sequencing samples were from independent experiments in which cells were randomly assigned to a given treatment or overexpression group.

**Blinding**

Sequencing datasets were blinded by the addition of unique barcodes to each individual sample.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

### Materials & experimental systems

| n/a | Involved in the study |
|-----|-----------------------|
| [ ] | Antibodies            |
| [x] | Eukaryotic cell lines |
| [ ] | Palaeontology         |
| [ ] | Animals and other organisms |
| [ ] | Human research participants |
| [ ] | Clinical data         |

### Methods

| n/a | Involved in the study |
|-----|-----------------------|
| [ ] | ChiP-seq              |
| [ ] | Flow cytometry        |
| [x] | MRI-based neuroimaging |

### Antibodies

**Antibodies used**

- rabbit anti-HA (1:1000; Cell Signaling; 37245),
- rabbit anti-m6A (1:1000; Abcam; ab151230),
- HRP-conjugated goat anti-rabbit (1:2500; Abcam; ab6721),
- HRP-conjugated sheep anti-mouse (1:2500; GE Healthcare; 95017-554),
- mouse anti-beta-actin (1:5000; Genscript; A00702),
- rabbit anti-METTL3 (1:1000; Abcam; ab195352),
- rabbit anti-cleaved caspase 3 (1:1000; Proteintech; 25546-1-AP),
- AlexaFluor 488-conjugated goat anti-rabbit (1:1000; Thermo-Fisher; A-21206; 1:1000).

**Validation**

Cells not expressing an HA-tagged protein were used as negative controls for western blot and immunofluorescence to confirm lack of immunoreactivity for the HA antibody. For the m6A antibody, we have previously validated this antibody for both blotting and IP applications in human cells (Meyer et al, Cell 2015). In addition, here, we use m6A methyltransferase-deficient cells to demonstrate loss of m6A immunoreactivity with this antibody.

### Eukaryotic cell lines

**Policy information about cell lines**

**Cell line source(s)**

HEK293T (ATCC)

**Authentication**

Cell lines were authenticated by visual inspection as well as RNA-seq. METTL3-depleted cells were validated with m6A immunoreactivity, anti-METTL3 western blot, and RNA-Seq.
Mycoplasma contamination

The cells used here tested negatively for mycoplasma contamination. Mycoplasma contamination was tested by DAPI staining.

Commonly misidentified lines
(See ICLAC register)

No commonly misidentified cell lines were used.